

REVIEW

Open Access



Systematic review of validation studies for the use of wearable smartwatches in the screening of atrial fibrillation

Muhammad Samsoor Zarak^{1*}, Sher Ali Khan², Harris Majeed³, Abdul Qahar Khan Yasinza⁴, Wadana Hamzazai⁵, Duy Chung¹, Gregory Koshkarian⁶ and Kevin S. Fleming¹

Abstract

Introduction Atrial fibrillation (AFib) is a common dysrhythmia and a risk factor for stroke and heart failure. Early detection and treatment are key to avoiding complications (especially in sustained AFib). Here, we systematically review the potential of wearable smartwatches (WSWs) to screen for AFib.

Method A literature search was conducted, and only those validation studies were shortlisted where the screening ability of WSWs was compared with EKG, and the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and Kappa statistic were provided.

Result Twelve studies were included with a combined sample size of 1,075,088. Most validation measures of WSWs were above 90% and comparable with KardiaBand by AliveCor (KB), which is an FDA-approved device to detect AFib.

Conclusion WSWs have the potential to reliably and continuously screen for AFib and detect it in a timely manner. The inconclusive results produced by WSWs are a significant problem. Once the inconclusive results are rectified, WSWs may be used for widespread screening of AFib in those people who are at high risk of developing AFib.

Keywords Atrial fibrillation, Wearable smartwatches, KardiaBand AliveCor, EKG

Introduction

Atrial fibrillation (AFib) is the most prevalent atrial arrhythmia, affecting 37.6 million people globally, and is only expected to rise as the population ages [1]. In the USA, the burden of AFib is projected to reach 12.1 million cases by 2030 [2]. This arrhythmia is associated with increased morbidity and mortality primarily characterized by thromboembolic events and cardiomyopathy. Prompt detection is crucial to managing these risks.

Widespread adoption of wearable smartwatches (WSW) has introduced an accessible method for detecting AFib. Specifically, WSW with electrocardiographic capability may be used to make a new diagnosis of AFib with clinician expertise. The recently updated 2023 ACC/AHA/ACCP/HRS Guideline for the Diagnosis and Management of AFib includes a class I recommendation allowing for the initial diagnosis of AFib to be made

*Correspondence:

Muhammad Samsoor Zarak
samzarak@gmail.com

¹ Department of Internal Medicine, Northwest Medical Centre, 6200 N La Cholla, Tucson, AZ 85741, USA

² Department of Family and Community Medicine, School of Medicine, University of New Mexico, Albuquerque, NM, USA

³ Department of Internal Medicine, School of Medicine, University of New Mexico, Albuquerque, NM, USA

⁴ Department of Medicine, Bolan University of Medical and Health Sciences, Quetta, Pakistan

⁵ Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ, USA

⁶ PIMA Heart and Vascular, Tucson, AZ, USA



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

based on a health care professional's examination of the EKG tracing from WSWs [3]. Of note, the most readily available WSWs use photoplethysmography (PPG)-based devices that are only useful for detecting irregular pulses and cannot be used to make a diagnosis of AFib—hence why we discuss the screening capability rather than the diagnostic potential of WSWs.

The screening capability of WSW in the detection of AFib is not entirely clear. Reported drawbacks of using WSW for detection of AFib, such as high false-positive results (i.e., low specificity) [4] unreliable readings with hairy wrists, or tattooed wrist skin, etc [5]. Although seven reviews on this topic exist, none are systematic or meta-analytical, often being narrative and limited in scope [4, 6–11]. Our study, encompassing eight varied studies, addresses this gap and responds to the evolving field of wearable technology, making our systematic review both timely and essential.

Our study aims to systematically review the effectiveness of these new screening methods compared to traditional ones in detecting AFib. We focus on evaluating the validity of smartwatches for the screening of AFib in comparison to EKG.

Methods

The systematic review included all published studies until December 30th, 2023. We utilized two databases, PubMed and OvidSP (outlined in Fig. 1), and searched the terms: “Smartwatch,” “Apple Smartwatch,” “Samsung Smartwatch,” “Huawei Smartwatch,” “Wearable Smartwatch,” “Atrial Fibrillation,” “AFib,” “Samsung,” “Apple,” “Huawei,” “Paroxysmal Atrial Fibrillation,” “Persistent Atrial Fibrillation,” “Long-Term Persistent Atrial Fibrillation,” “Permanent atrial fibrillation”. A manual search of references of key studies was also conducted. In the first phase, titles were screened and in the second phase, the full text of shortlisted articles was screened to shortlist studies. We followed the recommendation of the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) statement while conducting this systematic review [12].

Inclusion criteria for studies

We included research papers from English-language peer-reviewed journals that attempt to validate the detection of AFib by wearable smartwatches using a clinician-interpreted EKG strip as reference and provide validation measures.

Data extraction

The extracted data included characteristics of each study, types of smartwatches used and the technology employed, and the sources of reference EKG used

as reference. Additionally, validation measures such as sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and kappa statistics were also extracted. The same validation measures were also extracted for Kardiaband by AliveCor (KB) wherever possible, which is an FDA-approved device for the detection of AFib [13, 14].

Results

Literature search and data extraction were done by the author MSZ. The author SK double-checked the data extracted and disagreements were mutually resolved. Twelve studies were included with a combined sample size of 1,075,088. Table 1 shows the characteristics of the study, smartwatches, and the sources of reference EKG. Table 2 shows validation measures of smartwatches along with comments on the peculiarities of each of the included studies.

In most of the included studies, the detection with WSWs was compared with results from a 12 or 1 lead EKG, which was interpreted by one or more clinicians who were sometimes blinded. In three studies, the validation measures for the KB were also reported [15–17].

In 10 out of 12 studies, the inconclusive diagnosis of AFib with smartwatches was excluded when calculating validation measures of sensitivity, specificity, PPV, NPV, accuracy, and Kappa statistic. The 2 studies, Mannhart et al. and Ford et al., presented these validation measures while including the inconclusive smartwatch diagnosis in calculating them. They considered the inconclusive results as incorrect, i.e., false negatives and false positives when EKG would show AFib and the lack of AFib, respectively. Such a method where they included the inconclusive readings of WSWs in calculating the validation measures is called “intention to diagnose/screen analysis” [15, 16].

Four studies evaluated the effectiveness of Apple WSWs in detecting AFib with a combined sample size of 1679 originating from the USA, Australia, and Switzerland.

Three studies assessed AFib using Samsung WSWs, with a combined sample size of 919 originating from Germany, Switzerland, and the USA.

Three studies assessed AFib using Huawei WSWs [18–20] with a combined sample size of 1413 originating from China, the United Kingdom, and Denmark. Four studies assessed AFib using other WSWs [15, 21–23] such as Fitbit, Verily Study Watch, and an unknown WSW by Nonogushi et al., with a combined sample size of 56,297 from the USA, Japan, and China.

The validation measures, i.e., sensitivity, specificity, PPV, NPV, accuracy, and Kappa statistic of Apple,

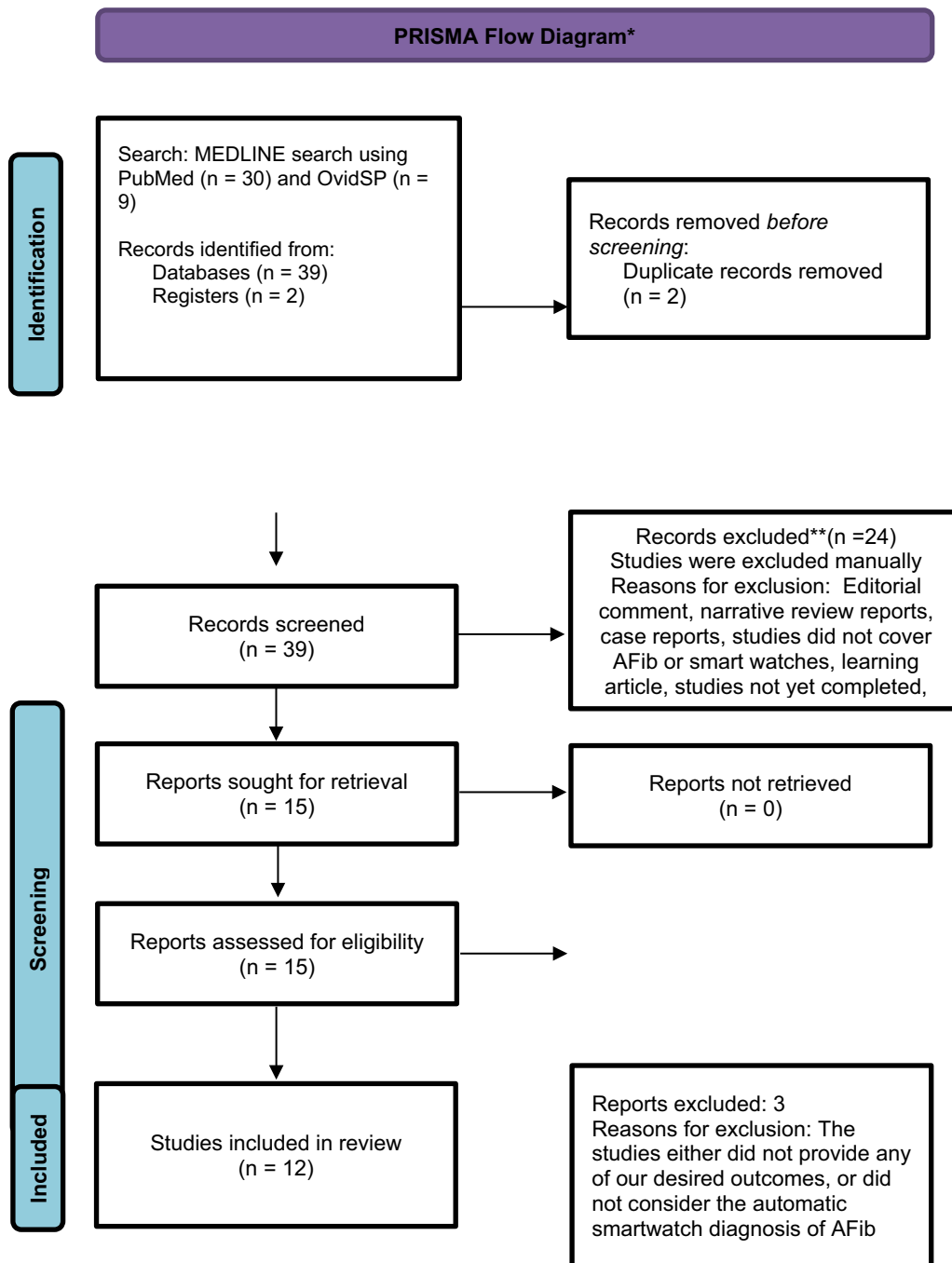


Fig. 1 Flowchart of literature search. *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews [12]

Samsung, and Huawei WSWs were comparable, while other WSWs showed comparatively lower values.

Table 3 summarizes the validation measures of WSWs and KB across all included studies and breaks them down according to whether the inconclusive readings of WSWs were included or not. The

validation measures of smartwatches were above 90% and Kappa above 0.85 when not including the inconclusive readings, while for KB they were above 95% and Kappa above 0.85. When inconclusive readings were included (i.e., intention to diagnose/screen analysis) then the validation measures significantly reduced.

Table 1 Characteristics of studies

References	WSW	Funding source	Study location	Sample size	Age [Yr, Mean \pm SD, or median (IQR)]	Male %	Comparator EKG
Mannhart [15]	Apple, Samsung, Withings Scan-watch, Fitbit	University of Basel, the Stiftung für Herzschrittmacher und Electrophysiologic, the Freiwillige Akademische Gesellschaft Basel, and Johnson & Johnson,	Switzerland	201	66.7 (58–75)	69%	12 lead EKG
Niu [24]	Huawei Watch GT2 Pro	National Natural Science Foundation of China, Tsinghua University Spring Breeze Fund	China	628	64 \pm 14	46%	12 lead EKG
Poh [25]	Verily Study Watch	Verily Life Sciences	USA	111	65 \pm 11	50%	Ambulatory EKG patch
Lubitz [20]	Fitbit	Fitbit	USA	455,699	47 (35–58)	27%	Ambulatory EKG patch
Nonoguchi [26]	Unknown WSW	Japan Agency for Medical Research and Development	Japan	286	66 \pm 12	68%	Telemetry 1-lead EKG
Ford [16]	Apple Watch 4	Eastern Health Foundation	Australia	125	76 \pm 7	62%	12-lead EKG
Perez [23]	Apple Watch	Apple	USA	419,297	41 \pm 13	57%	Wearable EKG patch (ePatch)
Bashar [27]	Samsung Simband	NIH	USA	46*			7 lead Holter EKG
Guo [18]	HuaweiWatch GT, Honor Watch, and Honor Band 4	National Natural Science Foundation of China (H2501) Health and Family Planning Commission of Heilongjiang Province, China	China, United Kingdom, Denmark	187,912	34.7 \pm 11.5	86.7%	24-hour Holter EKG
Zhang [19]	Huawei WatchGT, Honor Watch, and the Honor Band4	Chinese PLA Healthcare Foundation and National Natural Science Foundation of China	China	361	50 (36–62)	50.7%	12-lead ECG,
Tison [22]	Apple Watch	Cardiogram Inc	USA	9750	42 \pm 12	63%	12-lead ECG
Dörr [17]	Samsung Gear Fit 2	Preventicus GmbH and the University Hospital Basel, Switzerland	Germany, Switzerland	672	76.4 \pm 9.5	55.7%	Single lead iECG of Kardiaband AliveCor

PAC: Premature Atrial Contraction, PVC: Premature Ventricular Contraction. ECG: Electrocardiogram, USA: United States of America, PPG: Photoplethysmography. NIH: National Institutes of Health, Yr: year/s, SD: standard deviation, IQR: interquartile range, WSW: wearable smartwatch

*37 patients from the University of Massachusetts, 9 patients from Connecticut

Discussion

Our systematic review shows that WSW has a similar screening accuracy to the FDA-approved KB [8, 9]. WSWs might offer a better alternative for AFib detection as compared to the standard of care.

False positive is a concern with such devices where patients get false alerts causing significant anxiety and

leading to unwarranted emergency department visits and ultimate loss of healthcare resources. False positive results can be diminished by raising the specificity to a level that renders tolerable false negatives. Ford et al. achieved 100% specificity but at the cost of sensitivity (i.e., 50%) [16].

Table 2 Results of studies

Study	WSW Manufacturer	Sensitivity, % (95 CI)	PPV, % (95 CI)	Specificity, % (95 CI)	NPV, % (95 CI)	Accuracy, % (95 CI)	Kappa	Comments		
Mannhart [15]	Apple	93 (81–99)	91 (78–98)	96 (89–99)	97 (91–99)	95	0.88	The participants of the study included patients who were scheduled for catheter ablation, electrical cardioversion, placement of implantable cardioverter-defibrillator (ICD) or a pacemaker. Before the procedure, all patients had 12 lead EKG followed by readings of 30 s each from 5 different WSWs that were worn on left hand. WSWs did not use PPG modality; therefore, one lead ECG was obtained by placing finger either on the crown, 2 diagonal corners, or stainless-steel ring. The study focused on AFib and did not merge at atrial flutter with AFib. The inconclusive readings were 1.8%, 1.7%, 2.4%, 2.1%, and 2.6% of the times for the Apple Watch, Samsung Galaxy Watch 3, Withings ScanWatch, Fitbit Sense, and KB, respectively, and were considered incorrect in the intention to diagnose/screen analysis.		
	Apple*	85 (72–94)	58 (46–70)	75 (67–83)	93 (86–97)	78 (71–84)	0.53			
	Withings Scanwatch	87 (70–96)	82 (67–93)	94 (87–98)	96 (89–99)	92 (86–96)	0.79			
	Withings Scanwatch*	58 (42–72)	48 (35–62)	75 (67–83)	82 (73–88)	70 (63–77)	0.31			
	Fitbit	89 (73–97)	94 (80–99)	98 (93–100)	96 (90–99)	95 (90–98)	0.88			
	Fitbit*	66 (51–79)	55 (42–69)	79 (70–86)	85 (77–91)	75 (68–82)	0.42			
	Samsung	95 (84–99)	85 (72–94)	93 (85–97)	98 (92–100)	93 (88–97)	0.85			
	Samsung*	85 (72–94)	57 (45–69)	75 (67–83)	93 (86–97)	78 (70–84)	0.52			
	KB	98 (89–100)	96 (86–100)	96 (90–99)	98 (92–100)	96 (91–99)	0.91			
	KB*	79 (64–89)	50 (38–62)	69 (60–77)	89 (81–95)	72 (64–78)	0.40			
	Niu [24]	Huawei GT2 Pro		95			95.8			The participants were of age ≥ 22 years from China without history of implantable heart devices. WSW and 12 lead EKG both were applied simultaneously to detect AFib for 3 min. WSW was worn on left wrist and subjects touched WSW with right hand. Accuracy is indicated by macro F1-score
Poh [25]	Verily Study Watch	96 (93–98)	81 (69–90)	98 (97–99)	99.7 (99.5–99.8)	98.6		The participants were of age ≥ 22 years with history of paroxysmal AFib and were from 5 ambulatory centers from USA. WSW and patch EKG were worn simultaneously for 2 weeks. Twenty-one percent of the WSW readings were inconclusive. We considered the Pearson correlation of 0.986 as accuracy		

Table 2 (continued)

Study	WSW Manufacturer	Sensitivity, % (95 CI)	PPV, % (95 CI)	Specificity, % (95 CI)	NPV, % (95 CI)	Accuracy, % (95 CI)	Kappa	Comments
Lubitz [20]	Fitbit	68 (62–73)	98.2 (95.5–99.5)	98 (96–99.5)	86.5 [#]		0.76 [#]	The participants were of age ≥ 22 years old U.S. residents who never had AFib. Those who had irregular heart rhythm detected with WSW were then monitored with ambulatory patch for 7 days. Both AFib and atrial flutter were considered as AFib
Nomoguchi [26]	Unknown WSW	98	69.4	90.6	99.5			Participants were recruited from a university hospital in Japan, 123 had known AFib while 163 were high risk for it The WSW and telemetry I-EKG were applied simultaneously for 3 days to detect AFib
Ford [16]	Apple Apple* KB KB*	50 19 96 89	100 100 84 77	100 100 93 93	92 82 99 97	93 65 94 74	0.64 0.24 0.82 0.56	The participants were from outpatient cardiology clinic. The 30 s tracing from WSW and KB were obtained from each individual. The 12 lead ECG were also simultaneously obtained AFib and atrial flutter were merged for analysis In the intention to diagnose/screen analysis 30% of the readings of WSWs and 20% of the readings of KB were inconclusive and were considered incorrect
Perez, 2019 [23]	Apple		84 (0.76 to 0.92)			84		Community-based participants aged ≥ 22 years without AFib used WSW to detect AFib. The “irregular pulse” notification was then confirmed with wearing an EKG patch for 7 days
Bashar, 2019 [27]	Samsung (Simband)	98.18		97.43		97.54		Participants came from two samples, i.e., UMass database where 37 participants out of which 10 with AFib were in outpatient setting, and Chonlab database which had 9 healthy, male participants, none had AFib The smartwatch used special technology to detect AFib in background of noise and motion artifact A 7 lead Holter EKG monitor was used as reference

Table 2 (continued)

Study	WSW Manufacturer	Sensitivity, % (95 CI)	PPV, % (95 CI)	Specificity, % (95 CI)	NPV, % (95 CI)	Accuracy, % (95 CI)	Kappa	Comments
Guo, 2019 [18]	Honor Band 4, Honor watch and Huawei watch	100	91.6 (91.5–91.8)	99	99	99		Participants were from community, aged ≥ 18 years and monitored for 14 days with the WSW. Those who received “possible AFib” notification was confirmed with 24-h Holter EKG monitoring. The validation measures reported for three types of watches were similar, hence we considered one.
Zhang, 2019 [19]	Honor Watch 1 AFib in Honor Watch 2 Smart Band	100 (85.75–100) 100 (84.56–100) 100 (87.23–100)	92.31 (74.87–99.05) 91.67 (73.00–98.97) 93.10 (77.23–99.15)	99.16 (97.01–99.90) 98.93 (96.19–99.87) 99.15 (96.97–99.90)	100 (98.46–100) 100 (98.03–100) 100 (98.44–100)		0.96 (0.90–1) 0.95 (0.88–1) 0.96 (0.91–1)	Participants from community aged ≥ 18 years and outpatient setting were recruited and screened automatically and periodically with WSW for AFib for 14 days. Detection of “suspected AFib” was compared with 12 lead EKG.
Tison, 2018 [22]	Apple Watch	98.0	90.9	90.2	97.8			Initially, the neural network of WSW was trained on a cohort of 9750 participants who were enrolled in the Health eHeart Study. Then, the performance of WSW in screening AFib was validated against 12 lead EKG, among 50 participants, who had visited UCSF hospital for cardioversion. The 12 lead EKG was obtained followed by application of WSW to the wrist for 20 min before and after cardioversion. The ability of smartwatch to screen AFib was compared with 12 lead EKG.

Table 2 (continued)

Study	WSW Manufacturer	Sensitivity, % (95 CI)	PPV, % (95 CI)	Specificity, % (95 CI)	NPV, % (95 CI)	Accuracy, % (95 CI)	Kappa	Comments
Dörr, 2018 [17]	Samsung Gear Fit 2	93.7 (89.8–96.4)	97.8 (94.9–99.3)	98.2 (95.8–99.4)	94.7 (91.4–97.0)	96.1 (94.0–97.5)		It was a case–controlled study with inpatient sample who were matched on sex and age; cases were the ones who had history of AFib and controls without AFib. Single lead IECG, produced by KB and interpreted by cardiologists; was used as reference. Performance of WSW to detect AFib at 1 min, 3 min and 5 min intervals was assessed. The results were almost same at all 3 intervals. We considered the results of 1 min. The prevalence of AFib in this sample was above 40% which is much higher than average healthy population. Hence, authors adjusted the PPV, NPV, and accuracy with lower prevalence of 5%, 10%, and 14%. That changed PPV 72.7% to 90%, NPV 99.7% to 98.9% for WSW, and PPV 66.8% to 87.1% for KB. The NPV and accuracy were above 97% for both WSW and KB
	KB	99.5 (97.5–99.9)	97.6 (95.0–99.0)	97.4 (94.7–98.9)	99.2 (97.3–99.9)	98 (96.9–99.2)		

PPV: Positive Predictive Value, NPV: Negative Predictive Value, AFib: Atrial Fibrillation, SR: Sinus Rhythm, AFSW: AFib Sensing Watch, PPG: photoplethysmography; AUC: Area Under the Curve, AW: Apple Watch, NSR: Normal Sinus Rhythm, PAC: Premature Atrial Complex, PVC: Premature Ventricular Complex, UCSF: University of California San Francisco, KB: KardiaBand AliveCor
 *intention to diagnose/screen measures

Diagnostic measure calculated by authors based on the provided data

Zhang et al. achieved high sensitivity and specificity (i.e., 100% and 99%, respectively), diminishing both false positives and false negatives [19]. The overall median sensitivity and specificity above 95% and median PPV and NPV above 90% in this study might provide a good balance between false-negative and false-positive alerts. We also know that PPV increases when the prevalence of a disease increases. Hence, if people with a higher risk of AFib use these WSWs, their PPV would increase, and the false positive alert rates would decline. The risk factors for AFib include advanced age, high blood pressure, obesity, European ancestry, diabetes, heart failure, ischemic heart disease, hypothyroidism, chronic kidney disease, moderate-to-heavy alcohol use, smoking, and enlargement of left heart chambers [28].

Ford et al. and Mannhart et al. brought up a significant problem when using WSWs to detect AFib, and that's inconclusive results. They noted that WSWs produced inconclusive readings up to 30% of the time and KB up to 25% of the time [15, 16]. When these inconclusive results were compared with the EKG, it was discovered that some of these were AFib and some were not. Hence, both false negatives and false positives were hidden behind the inconclusive label.

Even though the frequency of getting inconclusive results was comparable across WSWs and KB, it is still a significant number. This frequency also varied between WSWs ranging from 17 to 30%. Ford et al. and Mannhart et al. incorporated these for calculating the diagnostic measures (i.e., intention to diagnose/screen analysis) and showed that it significantly and negatively impacted the validation measures [15, 16]. In the real world when people use these devices for continuous detection of AFib, they get inconclusive results up to 30% of the time, which may cause significant anxiety and loss of resources. That means we still have a long way to improve the performance of these devices.

Cohen's Kappa statistic informs us about how reliable the WSW is in detecting AFib. In other words, how much agreement is there between the WSW and comparator EKG utilized by a clinician? Overall median kappa statistic of above 0.8 in WSWs and KB shows almost perfect agreement with the comparator EKG. However, when inconclusive results are taken into account then it drops to 0.4 for both WSWs and KB, showing just moderate agreement. Some of the included studies that have not provided Kappa statistics have provided accuracy. All accuracy measures are above 90% which drops to 70% when inconclusive results are added.

PPG technology is not very new and has been used in various medical purposes for detecting oxygen saturation, measuring blood pressure, and cardiac output, evaluating autonomic function, and detecting peripheral

vascular disease. PPG-equipped WSWs will be more beneficial compared to non-PPG since they have better outcomes and can be passively used for the detection of AFib. However, there are certain limitations associated with the utilization of the technology since PPG sensors are impacted by skin pigmentation, the color of the LED, the contact force between the site and sensor, ambient temperature, motion artifacts, and ambient light interference. The battery life can also bring limitations in terms of missing the arrhythmias, particularly in paroxysmal AFib [21].

As mentioned in the introduction, the updated 2023 ACC/AHA/ACCP/HRS Guideline for the Diagnosis and Management of AFib strongly recommends for initial diagnosis of AFib to be made based on a health care professional's examination of the EKG tracing from WSWs. Some of the WSWs assessed in the included studies have the capability to produce such EKG tracings. Having said that, once the PPG-equipped WSW notifies the user about AFib detection, the user can then touch the dial or the crown of the WSW with the opposite hand and immediately record an EKG tracing. This EKG tracing can be reviewed by clinician either in person or remotely to diagnose AFib. This way the asymptomatic or paroxysmal AFib can be caught in a timely manner to avoid medical consequences.

The economic burden of AFib is a significant concern in healthcare systems. AFib is associated with increased healthcare costs, hospitalizations, and resource utilization. The total estimated annual healthcare cost associated with AFib in Canada exceeds \$800 million and €20,403–€26,544 (\$22,340–\$29,064) per patient in Denmark [29, 30]. The annual rise in AFib incidence is closely linked to cumulating risk factors, notably advancing age, obesity, hypertension, and type 2 diabetes. The prevalence of AFib in adults ranges between 2 and 4%, with a pronounced surge among individuals over 65 years of age [31].

AFib is the leading cause of stroke [22]. This challenge is compounded by AFib's ability to remain asymptomatic or sporadic, revealing itself only over time. Such hidden nature exacts a substantial economic toll, accounting for 1%–2% of healthcare expenditures [32]. One study highlighted the escalating burden, with AFib hospitalizations surging from 288,225 in 2007 to 333,570 in 2014, attributed to an increase in total annual emergency department visits during the study period [33]. The adjusted annual charges for admitted AFib patients soared by 37%, from \$7.39 billion in 2007 to \$10.1 billion in 2014 [34]. AFib increases the risk of ischemic stroke and heart failure (both heart failure with reduced ejection fraction and heart failure with preserved ejection fraction) by fivefold [35, 36]. In terms of mortality, the age-adjusted

Table 3 Comparison of diagnostic measures of WSWs with KB

Diagnostic measures	WSW				KB			
	Inconclusive readings excluded		Inconclusive readings included ^a		Inconclusive readings excluded		Inconclusive readings included ^a	
	Total values	%, median (IQR)	Total values	%, median (IQR)	Total values	%, median (IQR)	Total values	%, median (IQR)
Sensitivity	15	96 (89, 100)	5	66 (58, 85)	3	98 (96, 99.5)	2	84 (79, 89)
Specificity	15	98 (94, 99)	5	75 (75, 79)	3	96 (93, 97.4)	2	81 (69, 93)
PPV	16	91.6 (85, 95)	5	57 (55, 58)	3	96 (84, 97.6)	2	64 (50, 77)
NPV	13	97.8 (96, 100)	5	85 (82, 93)	3	99 (98, 99)	2	93 (89, 97)
Accuracy	12	95 (93, 97)	5	75 (70, 85)	3	96 (94, 98)	2	73 (72, 74)
Kappa (in fraction of one)	8	0.88 (0.82, 0.95)	5	0.42 (0.31, 0.52)	2	0.86 (0.82, 0.91)	2	0.48 (0.4, 0.56)

IQR: Interquartile range, %: percentage, WSW: Wearable Smartwatch, KB: KardiaBand AliveCor

^aAlso called intention to diagnose/screen analysis

mortality rate with AFib soared from 18/100,000 in 2011 to 22.3/100,000 in 2018 [37].

The use of smartwatches to detect AFib has the potential to reduce medical costs, directly and indirectly. Currently, most of the symptomatic AFib cases are present in the emergency department. Detection of AFib with smartwatches can reduce direct medical costs due to AFib by diverting the flow of symptomatic AFib from the emergency department to outpatient care. It can also reduce the indirect costs by reducing the rates of AFib-related consequences with timely and effective treatment of AFib. Those consequences include embolic stroke, falls, heart failure, etc.

We also understand the updated 2023 perspective of ACC/AHA/ACCP/HRS for management of AFib states that it is unusual to detect AFib using WSW for mass screening of the asymptomatic population. They maintain that the data demonstrating improved outcomes, including stroke, is lacking even when AFib is detected in an asymptomatic population. However, they do not provide any specific recommendation for routine AFib screening, which is in concordance with the US Preventive Task Force [38].

It is prudent to mention here that we are not advocating for everyone to wear a WSW to detect AFib. That would be a significant economic burden and can lead to high false-positive and false-negative rates. WSWs used for the detection of AFib can be considered for those who are at higher risk for AFib. It can also be considered for those patients with AFib who are on rhythm control therapy for the prevention of stroke and are not candidates for anticoagulation. In that case, a cardiologist or a primary care physician can adjust the dose of the anti-arrhythmic agent as per the number of AFib alerts on a WSWs. We recommend that such AFib alerts

produced by PPG technology should then be confirmed by EKG tracing produced by the same WSW and read by a clinician.

Currently, the detection of AFib is primarily achieved through intermittent methods such as office visits, emergency department visits, or incidental findings, all of which are characterized by periodicity and a heightened likelihood of missing an AFib potentially leading to missed diagnosis and progression to fatal events [39]. Moreover, the STROKESTOP study demonstrated that conducting several intermittent short EKG recordings over an extended duration led to a fourfold enhancement in sensitivity for detecting AFib when compared to single-time measurements [40]. In cases where AFib is suspected but the diagnosis can't be confirmed, more advanced measures such as 24-h Holter monitoring or implantable cardioverter defibrillation are employed which bears more discomfort and psychological stress for the patients compared to WSWs. Furthermore, the introduction of WSWs provides a novel avenue for the identification of irregular pulses, offering users the ability to promptly detect such arrhythmias using WSWs, eliminating the constraints of periodic monitoring. These WSWs present a convenient, non-invasive, and easily accessible alternative to traditional EKG monitoring methods.

The treatment of AFib can become effective with the use of a WSW since the combination of machine learning with WSW has the potential to bypass error and fatigue which are embedded in human efforts. It can result in early detection and management of AFib resulting in reduced risk of serious consequences (especially in sustained AFib). Furthermore, unlike traditional devices, which are constrained by their intermittent monitoring and the requirement to be returned to the offices for subsequent analysis,

the data storage on WSWs, capable of being stored on the cloud, enables remote access. This holds the promise of streamlining access to essential health information, facilitating telemedicine and remote monitoring, which fosters seamless connectivity between patients and healthcare providers. Integrating real-time cardiac data with AI-driven algorithms empowers remote monitoring, enabling timely interventions and personalized treatment.

Our study has some strengths. All of the included studies are validation studies comparing the detection capabilities of smartwatches with the gold standard screener, i.e., EKG. Most of the studies have reported multiple validation measures that are necessary for the evaluation of a screening device such as sensitivity, specificity, PPV, NPV, and Kappa statistic. A consistency was noted across most studies when comparing above mentioned measures which reflects the precision across the studied WSWs in detecting AFib. The combined sample size was large, in thousands. Our study has some limitations. There was some heterogeneity between studies, such as in some studies participants were coming from cardiology procedures, while others had no history of AFib or cardiac conditions and were comparatively healthy coming from the community. Additionally, some studies had small sample sizes.

Conclusion

This systematic review presents evidence of the potential of WSWs in detecting AFib, which is comparable with KB, an FDA-approved device for AFib screening. There is room for improvement to rectify the inconclusive readings. This review underscores the emerging role of digital health applications in modern healthcare, especially in cardiovascular monitoring. Cost-utility analysis is needed to know the quality-adjusted life-years gained against the monetary expense. Another systematic review, assessing the diagnostic rather than screening potential of those WSWs which are equipped with producing EKG tracings is needed.

Abbreviations

AFib	Atrial fibrillation
PAC	Premature atrial contraction
PVC	Premature ventricular contraction
ECG	Electrocardiogram
USA	United States of America
PPG	Photoplethysmography
NIH	National Institutes of Health
Yr	Year/s
SD	Standard deviation
WSW	Wearable smartwatch
PRISMA	Reporting Items for Systematic Review and Meta-analysis
FDA	Food and Drug Administration
ACC/AHA/ACCP/HRS	American College of Cardiology/American Heart Association Joint Committee on Clinical Practice, Heart Rhythm Society

PPV	Positive predictive value
NPV	Negative predictive value
SR	Sinus rhythm
AFSW	AFib Sensing Watch
AUC	Area under the curve
AW	Apple Watch
NSR	Normal sinus rhythm
UCSF	University of California San Francisco
KB	KardiaBand AliveCor
IQR	Interquartile range
%	Percentage

Acknowledgements

Not applicable.

Author contributions

Concept and design of the study was performed by MSZ. Literature review was provided by MSZ and WH. Literature writing was done by MSZ, SAK, AQKY, and DC. Manuscript Review was presented by SAK, HM, GK, and KF.

Funding

There is no source of funding for this study.

Availability of data and materials

It will be provided as needed.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

All authors bear no competing interests.

Received: 13 March 2024 Accepted: 18 June 2024

Published online: 26 June 2024

References

- Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: an increasing epidemic and public health challenge. *Int J Stroke*. 2021;16(2):217–21.
- Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the US adult population. *Am J Cardiol*. 2013;112(8):1142–7.
- Joglar JA, Chung MK, Armbruster AL, et al. 2023 ACC/AHA/ACCP/HRS guideline for the diagnosis and management of atrial fibrillation: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation*. 2024;149(1):e1–156.
- Mohamoud A, Jensen J, Buda KG. Consumer-grade wearable cardiac monitors: what they do well, and what needs work. *Clev Clin J Med*. 2024;91(1):23–9.
- Kim S, Xiao X, Chen J. Advances in photoplethysmography for personalized cardiovascular monitoring. *MDPI*; 2022. p. 863.
- Raja JM, Elsagr C, Roman S, et al. Apple watch, wearables, and heart rhythm: where do we stand? *Ann Transl Med*. 2019;7(17).
- Pay L, Yumurtas AÇ, Satti DI, et al. Arrhythmias beyond atrial fibrillation detection using smartwatches: a systematic review. *Anatol J Cardiol*. 2023;27(3):126.
- Sajeev JK, Koshy AN, Teh AW. Wearable devices for cardiac arrhythmia detection: a new contender? *Intern Med J*. 2019;49(5):570–3.
- Carpenter A, Frontera A. Smart-watches: a potential challenger to the implantable loop recorder? *Europace*. 2016;18(6):791–3.

10. Khurshid S, Healey JS, McIntyre WF, Lubitz SA. Population-based screening for atrial fibrillation. *Circ Res*. 2020;127(1):143–54.
11. Santala OE, Lipponen JA, Jäntti H, et al. Novel technologies in the detection of atrial fibrillation: review of literature and comparison of different novel technologies for screening of atrial fibrillation. *Cardiol Rev*. 2023;10:1097.
12. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Int J Surg*. 2021;88:105906.
13. Bumgarner JM, Lambert CT, Hussein AA, et al. Smartwatch algorithm for automated detection of atrial fibrillation. *J Am Coll Cardiol*. 2018;71(21):2381–8.
14. Wasserlauf J, You C, Patel R, Valys A, Albert D, Passman R. Smartwatch performance for the detection and quantification of atrial fibrillation. *Circ Arrhythmia Electrophysiol*. 2019;12(6):34.
15. Mannhart D, Lischer M, Knecht S, et al. Clinical validation of 5 direct-to-consumer wearable smart devices to detect atrial fibrillation: BASEL wearable study. *Clin Electrophysiol*. 2023;9(2):232–42.
16. Ford C, Xie CX, Low A, et al. Comparison of 2 smart watch algorithms for detection of atrial fibrillation and the benefit of clinician interpretation: SMART WARS study. *Clin Electrophysiol*. 2022;8(6):782–91.
17. Dörr M, Nohturfft V, Brasier N, et al. The WATCH AF trial: SmartWATCHes for detection of atrial fibrillation. *JACC Clin Electrophysiol*. 2019;5(2):199–208.
18. Guo Y, Wang H, Zhang H, et al. Mobile photoplethysmographic technology to detect atrial fibrillation. *J Am Coll Cardiol*. 2019;74(19):2365–75.
19. Zhang H, Zhang J, Li H-B, 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;21(12):e14909.
20. Lubitz SA, Faranesh AZ, Selvaggi C, et al. Detection of atrial fibrillation in a large population using wearable devices: the Fitbit Heart Study. *Circulation*. 2022;146(19):1415–24.
21. Allen J. Photoplethysmography and its application in clinical physiological measurement. *Physiol Meas*. 2007;28(3):R1.
22. Tison GH, Sanchez JM, Ballinger B, et al. Passive detection of atrial fibrillation using a commercially available smartwatch. *JAMA Cardiol*. 2018;3(5):409–16.
23. Perez MV, Mahaffey KW, Hedlin H, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med*. 2019;381(20):1909–17.
24. Niu Y, Wang H, Wang H, Zhang H, Jin Z, Guo Y. Diagnostic validation of smart wearable device embedded with single-lead electrocardiogram for arrhythmia detection. *Digit Health*. 2023;9:20552076231198680.
25. Poh MZ, Battisti AJ, Cheng LF, et al. Validation of a deep learning algorithm for continuous, real-time detection of atrial fibrillation using a wrist-worn device in an ambulatory environment. *J Am Heart Assoc*. 2023;12(19):e030543.
26. Nonoguchi NM, Soejima K, Goda A, et al. Accuracy of wristwatch-type photoplethysmography in detecting atrial fibrillation in daily life. *Eur Heart J Digit Health*. 2022;3(3):455–64.
27. Bashar SK, Han D, Hajeb-Mohammadalipour S, et al. Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches. *Sci Rep*. 2019;9(1):15054.
28. Benjamin EJ, Levy D, Vaziri SM, D'Agostino RB, Belanger AJ, Wolf PA. Independent risk factors for atrial fibrillation in a population-based cohort: the Framingham Heart Study. *JAMA*. 1994;271(11):840–4.
29. O'Reilly DJ, Hopkins RB, Healey JS, et al. The burden of atrial fibrillation on the hospital sector in Canada. *Can J Cardiol*. 2013;29(2):229–35.
30. Johnsen SP, Dalby LW, Täckström T, Olsen J, Fraschke A. Cost of illness of atrial fibrillation: a nationwide study of societal impact. *BMC Health Serv Res*. 2017;17:1–8.
31. Hindricks G, Potpara T, Dagres N, et al. 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*. 2020;42(5):373–498. <https://doi.org/10.1093/eurheartj/ehaa612>.
32. Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Kardiol Pol (Pol Heart J)*. 2016;74(12):1359–469.
33. Rozen G, Hosseini SM, Kaadan MI, et al. Emergency department visits for atrial fibrillation in the United States: trends in admission rates and economic burden from 2007 to 2014. *J Am Heart Assoc*. 2018;7(15):e009024.
34. Mei DA, Romiti GF, Boriani G. Prepare our healthcare systems to manage complexity: the case of atrial fibrillation. *Acta Cardiol*. 2023;78(9):1062–4.
35. Tsao CW, Aday AW, Almarzooq ZI, et al. Heart disease and stroke statistics—2022 update: a report from the American Heart Association. *Circulation*. 2022;145(8):e153–639.
36. Vermond RA, Geelhoed B, Verweij N, et al. Incidence of atrial fibrillation and relationship with cardiovascular events, heart failure, and mortality: a community-based study from the Netherlands. *J Am Coll Cardiol*. 2015;66(9):1000–7.
37. Tanaka Y, Shah NS, Passman R, Greenland P, Lloyd-Jones DM, Khan SS. Trends in cardiovascular mortality related to atrial fibrillation in the United States, 2011 to 2018. *J Am Heart Assoc*. 2021;10(15):e020163.
38. Khidhir A, Curtis AB. Perspective on ACC/AHA/ACCP/HRS for the management of atrial fibrillation. *Circulation*. 2024;149(1):4–6.
39. Steinhubl SR, Waalen J, Edwards AM, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mStoPS randomized clinical trial. *JAMA*. 2018;320(2):146–55.
40. Svennberg E, Friberg L, Frykman V, Al-Khalili F, Engdahl J, Rosenqvist M. Clinical outcomes in systematic screening for atrial fibrillation (STROKESTOP): a multicentre, parallel group, unmasked, randomised controlled trial. *The Lancet*. 2021;398(10310):1498–506.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.