**Smartphone-based Applications for Atrial Fibrillation Detection: A Systematic Review and Meta-analysis of Diagnostic Test Accuracy**

Isabella Oliveira Freitas Barbosa, M.S.†; Beatriz Costa de Oliveira, M.S.†; Antônio da Silva Menezes Júnior, M.D., Ph.D.\*†

†Medical Department, Medical Sciences and Life School, Pontifical Catholic University of Goiás, Goiânia, Goiás, Brazil

Address for correspondence:

Antônio da Silva Menezes Júnior, M.D, PhD.

a.menezes.junior@uol.com.br

Medical Department, Medical Faculty, Federal University of Goiás.

5ª Avenida, s/n - Setor Leste Universitário, Goiânia - GO, 74605-050

**ABSTRACT**

**Background:** Atrial fibrillation (AF) burden is strongly associated with an increased risk of stroke, which, in most cases, could have been prevented through earlier detection of AF and the timely initiation of anticoagulation therapy. Smartphone devices can provide a simple, non-invasive, cost-effective early AF detection solution.

**Methods:** PubMed, Embase, and Scopus databases were searched for studies comparing smartphone-based photoplethysmography (PPG) with standard electrocardiogram (ECG) for AF detection. A bivariate random-effects model with a 95% confidence interval (CI) was applied to generate the summary receiver operating characteristic (SROC) curve.

**Results:** Fourteen studies were included, comprising 5,090 patients with an AF prevalence of 31.6%. The pooled sensitivity and specificity were 0.96 (95% CI, 0.93–0.97) and 0.97 (95% CI, 0.95–0.98). The area under the SROC curve was 0.98 (95% CI, 0.94–0.99). The diagnostic odds ratio was 960 (95% CI, 439–2,104), with significant heterogeneity (I² = 51%). The projected positive and negative predictive values were 66.5% and 99.7%, respectively, in the elderly population aged >65 years and 39.2% and 99.9% in the general population.

**Conclusion:** Smartphone-based PPG demonstrated relatively high sensitivity and specificity and appears capable of ruling out AF. Patients aged >65 are more likely to benefit from AF screening.

**Keywords:** Atrial Fibrillation. Meta-analysis. Photoplethysmography. Smartphone. Telehealth

|  |  |  |
| --- | --- | --- |
| Abbreviations | | |
| AF | - | Atrial Fibrillation |
| PPG | - | Photoplethysmography |
| ECG | - | Electrocardiogram |
| ESC | - | European Society of Cardiology |
| NNS | - | Number Needed to Screen |
| PRISMA-DTA | - | Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies |
| PROSPERO | - | International Prospective Register of Systematic Reviews |
| QUADAS-2 | - | Quality Assessment of Diagnostic Accuracy Studies 2 |
| DM | - | Diabetes Mellitus |
| HTN | - | Hypertension |
| CAD/VD | - | Coronary Artery Disease/Vascular Disease |
| HF | - | Heart Failure |
| TIA | - | Transient Ischemic Attack |
| CHA2DS2–VASc | - | Congestive Heart Failure, Hypertension, Age ≥75 years, Diabetes Mellitus, Stroke/Transient Ischemic Attack, Vascular Disease, Age 65–74 years, Sex category (female) |
| CI | - | Confidence Interval |
| AUC | - | Area Under the Curve |
| DOR | - | Diagnostic Odds Ratio |
| PLR | - | Positive Likelihood Ratio |
| NLR | - | Negative Likelihood Ratio |
| PPV | - | Positive Predictive Value |
| NPV | - | Negative Predictive Value |
| SROC | - | Summary Receiver Operating Characteristic |

**INTRODUCTION**

Atrial fibrillation (AF) is the most common cardiac arrhythmia, affecting over 33 million individuals globally, and it is strongly associated with increased morbidity and mortality.1 Age is a significant risk factor, with the prevalence of AF rising dramatically after the age of 65. This trend poses a growing challenge as the global proportion of adults over 65 is projected to double from 12% in 2010 to 22% by 2040.2

AF often remains asymptomatic, with its first manifestation frequently being a thromboembolic event such as a stroke. Alarmingly, over 25% of strokes are attributed to previously undiagnosed AF, underscoring the critical need for earlier detection. Timely identification and treatment, such as the initiation of anticoagulation therapy, could prevent many of these strokes. This highlights a pressing challenge for cardiology public health and telemedicine, where scalable, accessible, and cost-effective solutions are essential.3,4

The advent of telemedicine has opened new avenues for AF detection and management. Portable rhythm-monitoring devices, including smartphones and wearables, have significantly expanded access to AF detection.2,4 This innovation, particularly through photoplethysmography (PPG), has increased consumer-initiated monitoring. PPG technology utilizes light penetration into biological tissues, where it is absorbed by substances such as blood. During systole, when blood volume in the arteries is highest, light absorption increases, whereas it decreases during diastole. PPG provides a simple, low-cost optical method for monitoring pulse irregularities by measuring these fluctuations at the skin's surface. 5,6

Conversely, robust evidence supporting the clinical impact of consumer-based AF screening using PPG devices remains limited. 7,8 The European Society of Cardiology (ESC) guidelines recommend opportunistic AF screening via pulse checks or ECG rhythm strips in individuals over 65 (class I, level B recommendation) to prevent AF-related complications. However, this approach requires screening approximately 70 individuals to detect a single case of AF, resulting in a high number needed to screen (NNS = 70) and limited cost-effectiveness.9,10

Given that traditional ECG monitoring can be costly, using a smartphone’s built-in camera, PPG technology presents an attractive alternative for facilitating AF screening through telehealth monitoring. 11,12To bridge this gap between technological innovation and clinical implementation, we aimed to evaluate the diagnostic accuracy of smartphone-based PPG applications for AF detection, contextualizing their potential within the broader telemedicine framework and comparing their clinical utility and cost-effectiveness to other digital and traditional AF screening methods.

**METHODS**

This systematic review and meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies (PRISMA-DTA) reporting guidelines (Supplemental Methods 1 and 2).13 A protocol was registered in the International Prospective Register of Systematic Reviews (PROSPERO) under CRD42024526899.

**Eligibility Criteria**

We included (1) randomized trials and observational studies (2) that recruited participants aged ≥18 years, (3) assessed smartphone-based PPG applications for AF detection, (4) provided a reference standard with ECG interpreted by a competent professional, and (5) reported sufficient data to calculate diagnostic accuracy. We excluded: (1) Studies in which diagnostic accuracy data were not extractable or could not be calculated; (2) Wristband devices and smartwatches; (3) Systematic reviews, case series, conference abstracts, studies focused on algorithm training, and studies lacking a gold-standard reference.

**Search strategies**

A comprehensive literature search was conducted in the PubMed, Embase, and Scopus databases from inception to October 30, 2024. To maximize the sensitivity of the results, no filters were applied regarding language, year of publication, or geographical location. The search included the descriptors 'smartphone,' 'mobile application,' 'cellphones,' and 'atrial fibrillation,' along with Medical Subject Headings (MeSH)/Emtree terms and keywords, combined using Boolean operators. The specific search strategies are detailed in Supplemental Methods 3.

**Study selection and data extraction**

Two reviewers (I.O.F.B. and B.C.O.) independently performed a conventional two-step screening process. The retrieved articles were uploaded to Rayyan software. First, titles and abstracts were screened, followed by a review of the full texts for final inclusion. Any disagreements were resolved by a third reviewer (A.S.M.J.). The included studies extracted the following information: author, setting, study design, algorithm used by the application, diagnostic measures, reference standard, comorbidities, and baseline clinical characteristics. Each investigator independently reviewed the data extracted by the other to ensure accuracy.

**Quality assessment**

We used the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)14 tool to assess the risk of bias in the included studies. The risk of bias in a diagnostic accuracy study was evaluated across four domains: patient selection, index test, reference standard, and flow and timing. Each domain was graded as having a high, low, or unclear risk of bias. Publication bias was assessed using Deeks' funnel plot asymmetry test.

**Subgroups and Meta-Regression**

Given the significant risk of patient selection bias in the studies, we performed a subgroup analysis by stratifying the populations into two groups: (1) studies that included only patients with a prior history of AF, such as those scheduled for electrical cardioversion or AF ablation, and (2) studies that recruited patients both with and without a diagnosis of AF. Additionally, we conducted subgroup analyses based on individual smartphone PPG applications, ECG lead types, country of enrollment, risk of bias, and study sample size. The likelihood ratio test for bivariate meta-regression was used to compare the goodness of fit between subgroups. Pairwise comparisons of the areas under the SROC curves were also performed to assess differences in test accuracy across subgroups.

Meta-regressions were also conducted on risk factors to assess whether comorbidities and clinical characteristics were associated with diagnostic test performance. The following covariates were included in the analysis: female sex, age, diabetes mellitus (DM), hypertension (HTN), coronary artery disease/vascular disease (CAD/VD), heart failure (HF), prevalence of AF, CHA2DS2–VASc score, previous stroke or transient ischemic attack (TIA), and anticoagulant therapy. These findings should be interpreted cautiously due to the limited number of participants included in the meta-regression analyses.

**Statistical Analysis**

To calculate pooled sensitivity and specificity, we performed a bivariate random-effects meta-analysis. We then modeled the summary receiver operating characteristic (SROC) curve and calculated the area under the curve (AUC) with the corresponding 95% confidence interval (CI). The pooled diagnostic odds ratio (DOR), positive likelihood ratio (PLR), and negative likelihood ratio (NLR) were calculated using a random-effects model. Heterogeneity for the DOR was assessed using Cochran's Q test and Higgins's I² statistic.

We chose not to model the pooled positive predictive value (PPV) and negative predictive value (NPV) based on the summary AF prevalence of the included studies for the following reasons: (1) most studies reported prevalence estimates significantly higher than expected for a screening population, and (2) small sample sizes and potential selection bias in participant sampling may have led to inflated prevalence estimates, which could subsequently overestimate PPV. We modeled predictive values for AF detection by projecting the PPV and NPV under various assumptions across a prevalence range from 1% to 100% to address this potential bias. From the resulting distribution for each prevalence level, we reported the PPV and NPV for two AF prevalence estimates representing different target populations, as in a previous study.15: (1) a 2% expected prevalence for the general adult population and (2) a 6% expected prevalence for elderly population aged >65 years.

A Fagan nomogram was also used to determine the post-test probability of disease based on the pre-test probability and the pooled NLR and PLR. In the Fagan nomogram, the left axis represents the pre-test probability, the middle axis displays the likelihood ratio of the diagnostic test, and the right axis shows the post-test probability. We also assumed 2% and 6% pre-test probabilities to represent the target populations.

All analyses were performed using R statistical software, version 4.2.2 (R Foundation for Statistical Computing), with the following libraries: mada, dmetatools, and meta. Statistical significance was defined as p < 0.05.

**RESULTS**

A total of 1,828 articles were retrieved, of which 1,218 remained for screening after deduplication. Of these, 1,195 were excluded based on title and abstract, leaving 23 articles for full-text review. Ultimately, 14 studies6,16–28 met the inclusion criteria and were included in the final quantitative analysis (Figure 1). The included studies enrolled a total sample of 5,090 participants, 31.6% of whom had AF detected by ECG. Two studies evaluated the FibriCheck application (Qompium), three assessed the Cardiio Rhythm Mobile application (Cardiio, Inc.), three examined the Preventicus application (Preventicus), and one appraised a study version of the PULSE-SMART application. Participant ages ranged from 63 to 80 years, with the proportion of females varying from 24.5% to 53.4% (Table 1). Comorbidities and additional baseline clinical characteristics are detailed in Supplemental Table 1. Diagnostic performance measures are provided in Supplemental Table 2.

**Data Synthesis of Diagnostic Accuracy in the Detection of AF**

Among the 14 studies, the meta-analysis showed that smartphone PPG applications' pooled sensitivity and specificity for detecting AF were 0.96 (95% CI, 0.93−0.97) and 0.97 (95% CI, 0.95−0.98), respectively. The SROC curve showed an AUC of 0.98 (95% CI, 0.94−0.99) (Figure 2). The summary DOR was 960.0 (95% CI, 438.0−2104.0). Measuring inconsistency for DOR using Cochran's Q and Higgins' I² revealed significant between-study heterogeneity (p = 0.01, I² = 51%) (see Supplemental Results 1). The pooled PLR and NLR were 31.19 (95% CI, 19.89−48.92) and 0.05 (95% CI, 0.03−0.07), respectively (see Supplemental Results 2 and 3).

**Fagan Nomogram for the Prediction of AF**

We used a Fagan nomogram to calculate post-test probabilities based on expected AF prevalences of 2% and 6%, with a PLR of 31.19 and an NLR of 0.05. Assuming a 2% pre-test probability, smartphone PPG testing for AF detection increased the post-test probability to 38.9% for a positive result and decreased it to 0.1% for a negative result (Figure 3A). For a 6% pre-test probability, the Fagan nomogram showed that the post-test probability of AF was 66.56% for a positive result and 0.3% for a negative result (Figure 3B).

**Positive and Negative Predictive Values**

We calculated the projected PPV and NPV across a prevalence range from 1% to 100% (see Supplemental Results 4). The resulting distributions of the projected predictive values for each prevalence level within this range are presented in Supplemental Results 5. Assuming a 2% prevalence estimate, the projected PPV and NPV were 39.2% (95% CI, 28.4%−51.0%) and 99.9% (95% CI, 99.8%−99.9%), respectively (Figure 4). For an expected 6% prevalence estimate, the projected PPV and NPV were 66.5% (95% CI, 55.5%−76.5%) and 99.7% (95% CI, 99.6%−99.8%), respectively (Figure 4).

**Exploring Heterogeneity with Subgroups and Meta-regression**

*Analyses stratified by the history of AF*

We performed a subgroup analysis based on the history of AF (Figure 5A). In studies that enrolled only patients with a history of AF, the pooled sensitivity, specificity, and AUC were 0.961 (95% CI, 0.927–0.980), 0.941 (95% CI, 0.886–0.970), and 0.981, respectively. In studies that included patients with and without AF, the pooled sensitivity, specificity, and AUC were 0.954 (95% CI, 0.914–0.976), 0.978 (95% CI, 0.964–0.987), and 0.987, respectively. The likelihood ratio test showed a statistically significant difference in the goodness of fit between subgroups (p = 0.0402) (see Supplemental Results 6 and 12).

*Analyses stratified by algorithm application*

We performed a subgroup analysis based on smartphone PPG applications (Figure 5B). For the Preventicus application, the pooled sensitivity, specificity, and AUC were 0.976 (95% CI, 0.828–0.997), 0.991 (95% CI, 0.977–0.997), and 0.990, respectively. For the FibriCheck application, the pooled sensitivity, specificity, and AUC were 0.956 (95% CI, 0.899–0.982), 0.967 (95% CI, 0.923–0.986), and 0.968, respectively. For the Cardiio Rhythm Mobile application, the pooled sensitivity, specificity, and AUC were 0.936 (95% CI, 0.890–0.963), 0.968 (95% CI, 0.941–0.983), and 0.944, respectively. The likelihood ratio test revealed no significant difference in the goodness of fit between subgroups (p = 0.277) (see Supplemental Results 7 and 12).

*Analyses stratified by reference standard*

We performed a subgroup analysis based on the reference standard (Figure 5C). For studies comparing with the 12-lead ECG, the pooled sensitivity, specificity, and AUC were 0.942 (95% CI, 0.911–0.963), 0.955 (95% CI, 0.924–0.974), and 0.976, respectively. In comparisons with the single-lead ECG, the pooled sensitivity, specificity, and AUC were 0.966 (95% CI, 0.908–0.988), 0.983 (95% CI, 0.969–0.990), and 0.989, respectively. The likelihood ratio test revealed no significant difference in the goodness of fit between subgroups (p = 0.114) (see Supplemental Results 8 and 12).

*Analyses stratified by country*

We performed a subgroup analysis stratified by country (Figure 5D). For Switzerland/Germany, the pooled sensitivity, specificity, and AUC were 0.978 (95% CI, 0.811–0.998), 0.985 (95% CI, 0.956–0.998), and 0.989, respectively. For Belgium, the pooled sensitivity, specificity, and AUC were 0.956 (95% CI, 0.899–0.982), 0.967 (95% CI, 0.923–0.986), and 0.968, respectively. For Hong Kong, the pooled sensitivity, specificity, and AUC were 0.948 (95% CI, 0.893–0.975), 0.979 (95% CI, 0.948–0.986), and 0.965, respectively. For the USA, the pooled sensitivity, specificity, and AUC were 0.950 (95% CI, 0.888–0.979), 0.934 (95% CI, 0.751–0.985), and 0.977, respectively. The likelihood ratio test revealed no significant differences in the goodness of fit across subgroups (p = 0.425) (see Supplemental Results 9 and 12).

*Analyses stratified by risk of bias*

We performed a subgroup analysis stratified by risk of bias (Figure 5E). For studies with at least one high-risk domain, the pooled sensitivity, specificity, and AUC were 0.957 (95% CI, 0.932–0.973), 0.971 (95% CI, 0.953–0.982), and 0.985, respectively. For studies classified as having a low risk of bias, the pooled sensitivity, specificity, and AUC were 0.931 (95% CI, 0.763–0.983), 0.937 (95% CI, 0.818–0.975), and 0.950, respectively. The likelihood ratio test revealed no significant difference in the goodness of fit between subgroups (p = 0.645) (see Supplemental Results 10 and 12).

*Analyses stratified by sample size*

We performed a subgroup analysis based on sample size (Figure 5F). In studies enrolling more than 100 patients, the pooled sensitivity, specificity, and AUC were 0.962 (95% CI, 0.934–0.979), 0.976 (95% CI, 0.960–0.986), and 0.987, respectively. For studies with fewer than 100 patients, the pooled sensitivity, specificity, and AUC were 0.930 (95% CI, 0.878–0.961), 0.927 (95% CI, 0.854–0.965), and 0.947, respectively. The likelihood ratio test comparing these subgroups' goodness of fit was insignificant (p = 0.078) (see Supplemental Results 11 and 12).

*Analyses of risk factors between studies*

Additionally, we performed meta-regression analyses on risk factors to evaluate the impact of clinical characteristics on diagnostic performance. The likelihood ratio test for meta-regression indicated that DM (p = 0.0405), anticoagulant therapy (p = 0.0438), and AF prevalence (p = 0.012) accounted for some of the heterogeneity observed among the primary studies (see Supplemental Results 13).

**Quality assessment and publication bias**

The QUADAS-2 evaluation revealed twelve studies.6,18–28 presented at least one domain with a high risk of bias, whereas two studies16,17 demonstrated an overall low risk of bias (see Supplemental Results 14). Deeks’ funnel plot test identified statistically significant asymmetry (p = 0.0013), indicating potential publication bias among studies assessing smartphone PPG applications for AF detection (Supplemental Results 15).

**DISCUSSION**

This systematic review and diagnostic test accuracy meta-analysis of 14 studies involving 5,090 patients found that smartphone-based PPG applications for AF detection, compared to standard ECG, demonstrated the following: (1) sensitivity of 96% and a specificity of 97%, corresponding to a false-positive rate of only 3%; (2) post-test probabilities of 38.9% (for a positive test) and 0.1% (for a negative test) in the general population, and 66.56% (for a positive test) and 0.3% (for a negative test) in the population aged >65 years; (3) PPV and NPV of 39.2% and 99.9%, respectively, in the general population, and 66.5% and 99.7%, respectively, in the population aged >65 years; and (4) a significant difference in diagnostic accuracy when the test was performed in populations with a known diagnosis of AF versus those being screened for AF, including individuals without a history of the condition. These findings align with previous studies, such as those by Prasitlumkum et al., further highlighting the value of PPG in AF detection as an advancement in telemedicine. 29

Smartphone app-based PPG devices have proven effective for screening or prolonged asymptomatic and symptomatic AF monitoring. These devices use the finger and camera light to accurately capture the rhythm and calculate variations in local arteriole blood volume pulses.30 Previous meta-analyses have reported high diagnostic performance for smartphone-based PPG, with sensitivity ranging from 94% to 94.7% and specificity from 95.8% to 97.6%.29–32 The discriminative ability presented in Gill et al. and Prasitlumkum et al., quantified by the AUC of the SROC curve, ranges between 0.96 and 0.98.29,31 Despite the accuracy of these tools, their use requires proper patient education, supervision, and guidance. Patients using smartphone PPG apps without sufficient support may experience heightened anxiety due to false-positive results. Furthermore, misinterpretation of PPG signals can lead to overtreatment. Current guidelines emphasize the importance of confirming AF detected by PPG devices through evaluation by a physician experienced in ECG rhythm interpretation. While PPG monitors effectively prompt individuals to seek further diagnostics, they are unreliable enough to establish a definitive AF diagnosis.10,33

Wearable smartwatches have introduced another accessible method for AF detection. Prasitlumkum et al. have demonstrated high diagnostic accuracy for smartwatches, with a sensitivity of 94%, a specificity of 97%, and an AUC of 0.99.29 The WATCH-AF trial highlighted comparable performance using smartwatch-derived PPG signals, reporting a sensitivity of 93.7%, a specificity of 98.2%, and an overall accuracy of 96.1%.34 In addition to PPG-based devices, these technologies widely incorporate single-lead ECG-based devices due to their accessibility. These devices, such as handheld finger or chest electrodes and adhesive patches, demonstrate a broad range of sensitivity (55–100%) and specificity (60–100%).35,36 Further head-to-head comparisons of these novel digital devices are needed to establish their comparative effectiveness in clinical settings, accounting for variations across different populations and environments.10

The probability of patients with AF having a normal rhythm on PPG is very low, at only 4%. The odds of an irregular R-R interval on PPG indicate that a patient is 960 times more likely to have AF. Additionally, we obtained a PLR of 31.19 and a NLR of 0.05, leading to a post-test probability of AF of 38.9% in the general population and 66.56% in the population aged >65. As shown by Sattar et al., the post-test probability, with a predicted general prevalence of AF in 3.03 million individuals in the USA, was 73% for those with an irregular AF rhythm detected via smartwatch and smartphone PPG.37 We modeled the PPV, NPV, and post-test probability for a population in which the detection of AF could lead to a change in management and help prevent cryptogenic stroke, such as in 65-year-old subjects.10,38 We obtained a moderate PPV and high NPV. Considering the NPV, a diagnosis of sinus rhythm on PPG strongly suggests the absence of AF, making it effective in ruling out AF. However, the same conclusion cannot be drawn from a positive result. We observed a better PPV in a population with a higher AF burden (>65 years), as it is well-known that PPV increases with a higher disease prevalence. Therefore, it would be worthwhile to investigate whether the PPV improves if these applications were used to screen a more selective, high-risk population, such as patients with a CHA2DS2-VASc score of at least 3.

Most consumer-based devices use PPG, and several large studies have typically been conducted in low-risk individuals10. However, the primary goal of screening and early detection of AF in patients is to reduce the risk of stroke through anticoagulation therapy. Previous randomized trials have assessed the incremental benefits of AF screening using handheld ECG devices in ambulatory patients aged ≥65 years who are at increased risk of stroke.39–41 The STROKESTOP study, which involved mass screening for AF using intermittent ECG recordings in the elderly, demonstrated that initiating stroke prophylaxis was highly successful in individuals with undiagnosed AF.40 Among patients aged ≥55 years with a recent cryptogenic stroke or TIA, ambulatory smartphone ECG recording with Kardio Mobile can significantly improve the detection of AF and the prescription of anticoagulation therapy.42

The smartphone-based AF screening strategies were evaluated in the eBRAVE-AF trial. Compared to routine symptom-based screening, they were found to significantly increase the detection rate of AF that requires oral anticoagulation therapy.43 In a prospective cohort study involving 60,629 individuals from the general population who underwent AF screening using a smartphone application based on PPG FibriCheck technology, 1.3% of participants tested positive for AF. The NNS was used to detect one new case of AF, which was 133. This strategy also positively influenced the uptake of oral anticoagulation therapy in individuals with clinical AF who indicated anticoagulation.44 We suppose that smartphone PPG devices can yield higher true positives when used as a screening tool for populations with a higher AF burden and could be considered for high-risk patients in a cost- and time-effective manner.

Besides being a screening tool, these applications were also applied for rhythm monitoring. Patients with an established diagnosis of paroxysmal AF may use these smartphone camera applications to monitor relapses of AF.10,30 A retrospective study comparing smartphone-based FibriCheck PPG monitoring of patients’ post-cardioversion with 12-lead ECG follow-up found that the smartphone application was highly effective, offering significant cost, time, and environmental savings. This approach could also be expanded to other settings, such as post-catheter ablation.45 In our meta-analysis, we performed a subgroup analysis separating the populations of patients with a diagnosed AF (post-AF ablation or post-cardioversion). In this population, smartphone PPG applications for rhythm monitoring showed a sensitivity of 96% and specificity of 94%. In the screening population, including patients with and without a diagnosis of AF, the sensitivity and specificity were 95% and 98%, respectively. The likelihood ratio test confirmed a statistically significant difference between the subgroups, likely driven by instability in specificity across the groups.

The substantial between-study heterogeneity observed in our meta-analysis warrants further examination. A key source of heterogeneity stemmed from differences in the selected populations due to potential selection bias and small sample sizes in single-center studies, leading to inflated AF prevalence estimates. Furthermore, each study employed different standard comparisons, and the test applications were not standardized, which could have been influenced by distinct signal strengths or technical issues that may contribute to the variations in sensitivity and specificity. To explore and address the potential impact of these factors, we conducted several subgroup analyses and meta-regressions to explore these methodological differences. This study underscores the consistent performance of smartphone-based PPG devices while addressing new technological advancements, such as integration with AI-driven analytics. Previous reviews often focused on the diagnostic accuracy of single-lead ECG devices, which remain highly effective but less scalable than PPG-based technologies. 6,41 By broadening the scope to include consumer devices, our analysis highlights the added value of democratizing AF detection through telemedicine solutions.

Opportunities presented by smartphone-based AF detection include improved accessibility, cost-effectiveness, and scalability. However, challenges persist, such as the need for standardized protocols, patient education, and integration with clinical workflows to minimize false positives and ensure follow-up care. Future recommendations include the development of hybrid models that combine PPG monitoring with AI-driven predictive algorithms and physician oversight to enhance reliability and clinical adoption. Future developments should focus on expanding the use of smartphone-based PPG technologies in high-risk populations, such as those with CHA2DS2-VASc scores ≥3, and refining algorithms to reduce false-positive rates. Large-scale, randomized trials are needed to establish the long-term benefits and cost-effectiveness of these tools.

This meta-analysis has several limitations that should be considered: 1) Between-study heterogeneity driven by various confounding factors associated with small sample size retrospective studies; 2) Many studies were conducted at single centers, limiting the generalizability of the findings; 3) The risk of selection bias cannot be excluded, as the absence of randomized data makes the observed differences in PPG sensitivity and specificity susceptible to confounding variables; 4) Selection bias also complicates the interpretation and application of the findings to real-world practice, potentially introducing spectrum bias; 5) Publication bias was evident, and the extent to which relevant studies remain unpublished is unclear.

**CONCLUSION**

Smartphone-based PPG devices have shown high accuracy in detecting AF compared to standard ECG, offering a reliable method for ruling out AF with strong negative predictive value. However, positive findings still require confirmatory ECG for diagnosis. The technology demonstrates greater utility in populations aged >65 years, where the higher prevalence of AF enhances its positive predictive value. To fully integrate PPG technology into telehealth monitoring, particularly for high-risk groups, randomized controlled trials focusing on stroke prevention and relevant clinical outcomes are essential.

**REFERENCES**

1. Andersen JH, Andreasen L, Olesen MS. Atrial fibrillation—a complex polygenetic disease. European Journal of Human Genetics 2021;29(7):1051–1060; doi: 10.1038/s41431-020-00784-8.

2. Kornej J, Börschel CS, Benjamin EJ, et al. Epidemiology of Atrial Fibrillation in the 21st Century. Circ Res 2020;127(1):4–20; doi: 10.1161/CIRCRESAHA.120.316340.

3. Freedman B, Potpara TS, Lip GYH. Stroke prevention in atrial fibrillation. The Lancet 2016;388(10046):806–817; doi: 10.1016/S0140-6736(16)31257-0.

4. Ashburner JM, Lee PR, Rivet CM, et al. The Implementation and Acceptability of a Mobile Application for Screening for Atrial Fibrillation at Home. Telemedicine and e-Health 2021;27(11):1305–1310; doi: 10.1089/tmj.2020.0427.

5. Sijerčić A, Tahirović E. Photoplethysmography-Based Smart Devices for Detection of Atrial Fibrillation. Tex Heart Inst J 2022;49(5); doi: 10.14503/THIJ-21-7564.

6. Mutke MR, Brasier N, Raichle C, et al. Comparison and Combination of Single-Lead ECG and Photoplethysmography Algorithms for Wearable-Based Atrial Fibrillation Screening. Telemedicine and e-Health 2021;27(3):296–302; doi: 10.1089/tmj.2020.0036.

7. Knight S, Lipoth J, Namvari M, et al. The Accuracy of Wearable Photoplethysmography Sensors for Telehealth Monitoring: A Scoping Review. Telemedicine and e-Health 2023;29(6):813–828; doi: 10.1089/tmj.2022.0182.

8. Avram R, Ramsis M, Cristal AD, et al. Validation of an algorithm for continuous monitoring of atrial fibrillation using a consumer smartwatch. Heart Rhythm 2021;18(9):1482–1490; doi: 10.1016/j.hrthm.2021.03.044.

9. Hill NR, Groves L, Dickerson C, et al. Identification of undiagnosed atrial fibrillation using a machine learning risk-prediction algorithm and diagnostic testing (PULsE-AI) in primary care: a multi-centre randomized controlled trial in England. European Heart Journal - Digital Health 2022;3(2):195–204; doi: 10.1093/ehjdh/ztac009.

10. Van Gelder IC, Rienstra M, Bunting K V, et al. 2024 ESC Guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2024;45(36):3314–3414; doi: 10.1093/eurheartj/ehae176.

11. Orchard J, Neubeck L, Freedman B, et al. eHealth Tools to Provide Structured Assistance for Atrial Fibrillation Screening, Management, and Guideline‐Recommended Therapy in Metropolitan General Practice: The AF‐SMART Study. J Am Heart Assoc 2019;8(1); doi: 10.1161/JAHA.118.010959.

12. Linz D, Gawalko M, Betz K, et al. Atrial fibrillation: epidemiology, screening and digital health. The Lancet Regional Health - Europe 2024;37:100786; doi: 10.1016/j.lanepe.2023.100786.

13. McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies. JAMA 2018;319(4):388; doi: 10.1001/jama.2017.19163.

14. Whiting PF. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med 2011;155(8):529; doi: 10.7326/0003-4819-155-8-201110180-00009.

15. Gruwez H, Evens S, Proesmans T, et al. Accuracy of Physicians Interpreting Photoplethysmography and Electrocardiography Tracings to Detect Atrial Fibrillation: INTERPRET-AF. Front Cardiovasc Med 2021;8; doi: 10.3389/fcvm.2021.734737.

16. Couderc J-P, Page A, Lutz M, et al. Assessment of facial video-based detection of atrial fibrillation across human complexion. Cardiovasc Digit Health J 2022;3(6):305–312; doi: 10.1016/j.cvdhj.2022.08.003.

17. Gruwez H, Ezzat D, Van Puyvelde T, et al. Real-world validation of smartphone-based photoplethysmography for rate and rhythm monitoring in atrial fibrillation. Europace 2024;26(4); doi: 10.1093/europace/euae065.

18. Yan BP, Lai WHS, Chan CKY, et al. Contact‐Free Screening of Atrial Fibrillation by a Smartphone Using Facial Pulsatile Photoplethysmographic Signals. J Am Heart Assoc 2018;7(8); doi: 10.1161/JAHA.118.008585.

19. Rozen G, Vaid J, Hosseini SM, et al. Diagnostic Accuracy of a Novel Mobile Phone Application for the Detection and Monitoring of Atrial Fibrillation. Am J Cardiol 2018;121(10):1187–1191; doi: 10.1016/j.amjcard.2018.01.035.

20. Poh M-Z, Poh YC, Chan P-H, et al. Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms. Heart 2018;104(23):1921–1928; doi: 10.1136/heartjnl-2018-313147.

21. McMANUS DD, CHONG JW, SONI A, et al. PULSE‐SMART: Pulse‐Based Arrhythmia Discrimination Using a Novel Smartphone Application. J Cardiovasc Electrophysiol 2016;27(1):51–57; doi: 10.1111/jce.12842.

22. Fan Y-Y, Li Y-G, Li J, et al. Diagnostic Performance of a Smart Device With Photoplethysmography Technology for Atrial Fibrillation Detection: Pilot Study (Pre-mAFA II Registry). JMIR Mhealth Uhealth 2019;7(3):e11437; doi: 10.2196/11437.

23. Chan P, Wong C, Poh YC, et al. Diagnostic Performance of a Smartphone‐Based Photoplethysmographic Application for Atrial Fibrillation Screening in a Primary Care Setting. J Am Heart Assoc 2016;5(7); doi: 10.1161/JAHA.116.003428.

24. Brasier N, Raichle CJ, Dörr M, et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). EP Europace 2019;21(1):41–47; doi: 10.1093/europace/euy176.

25. Krivoshei L, Weber S, Burkard T, et al. Smart detection of atrial fibrillation. Europace 2016;euw125; doi: 10.1093/europace/euw125.

26. Fernstad J, Svennberg E, Aberg P, et al. Validation of a novel smartphone-based photoplethysmographic method for ambulatory heart rhythm diagnostics. EP Europace 2022;24(Supplement\_1); doi: 10.1093/europace/euac053.578.

27. Mol D, Riezebos RK, Marquering HA, et al. Performance of an automated photoplethysmography-based artificial intelligence algorithm to detect atrial fibrillation. Cardiovasc Digit Health J 2020;1(2):107–110; doi: 10.1016/j.cvdhj.2020.08.004.

28. Proesmans T, Mortelmans C, Van Haelst R, et al. Mobile Phone–Based Use of the Photoplethysmography Technique to Detect Atrial Fibrillation in Primary Care: Diagnostic Accuracy Study of the FibriCheck App. JMIR Mhealth Uhealth 2019;7(3):e12284; doi: 10.2196/12284.

29. Prasitlumkum N, Cheungpasitporn W, Chokesuwattanaskul A, et al. Diagnostic accuracy of smart gadgets/wearable devices in detecting atrial fibrillation: A systematic review and meta-analysis. Arch Cardiovasc Dis 2021;114(1):4–16; doi: 10.1016/j.acvd.2020.05.015.

30. O’Sullivan JW, Grigg S, Crawford W, et al. Accuracy of Smartphone Camera Applications for Detecting Atrial Fibrillation. JAMA Netw Open 2020;3(4):e202064; doi: 10.1001/jamanetworkopen.2020.2064.

31. Gill S, Bunting K V, Sartini C, et al. Smartphone detection of atrial fibrillation using photoplethysmography: a systematic review and meta-analysis. Heart 2022;108(20):1600–1607; doi: 10.1136/heartjnl-2021-320417.

32. Yang TY, Huang L, Malwade S, et al. Diagnostic Accuracy of Ambulatory Devices in Detecting Atrial Fibrillation: Systematic Review and Meta-analysis. JMIR Mhealth Uhealth 2021;9(4):e26167; doi: 10.2196/26167.

33. 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); doi: 10.1161/CIR.0000000000001193.

34. 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; doi: 10.1016/j.jacep.2018.10.006.

35. Babar F, Cheema AM, Ahmad Z, et al. Sensitivity and Specificity of Wearables for Atrial Fibrillation in Elderly Populations: A Systematic Review. Curr Cardiol Rep 2023;25(7):761–779; doi: 10.1007/s11886-023-01898-3.

36. Hermans ANL, Gawalko M, Dohmen L, et al. Mobile health solutions for atrial fibrillation detection and management: a systematic review. Clinical Research in Cardiology 2022;111(5):479–491; doi: 10.1007/s00392-021-01941-9.

37. Sattar Y, Song D, Sarvepalli D, et al. Accuracy of pulsatile photoplethysmography applications or handheld devices vs. 12-lead ECG for atrial fibrillation screening: a systematic review and meta-analysis. Journal of Interventional Cardiac Electrophysiology 2022;65(1):33–44; doi: 10.1007/s10840-021-01068-x.

38. Friberg L, Engdahl J, Frykman V, et al. Population screening of 75- and 76-year-old men and women for silent atrial fibrillation (STROKESTOP). EP Europace 2013;15(1):135–140; doi: 10.1093/europace/eus217.

39. Halcox JPJ, Wareham K, Cardew A, et al. Assessment of Remote Heart Rhythm Sampling Using the AliveCor Heart Monitor to Screen for Atrial Fibrillation. Circulation 2017;136(19):1784–1794; doi: 10.1161/CIRCULATIONAHA.117.030583.

40. Svennberg E, Friberg L, Frykman V, et al. Clinical outcomes in systematic screening for atrial fibrillation (STROKESTOP): a multicentre, parallel group, unmasked, randomised controlled trial. The Lancet 2021;398(10310):1498–1506; doi: 10.1016/S0140-6736(21)01637-8.

41. Lubitz SA, Atlas SJ, Ashburner JM, et al. Screening for Atrial Fibrillation in Older Adults at Primary Care Visits: VITAL-AF Randomized Controlled Trial. Circulation 2022;145(13):946–954; doi: 10.1161/CIRCULATIONAHA.121.057014.

42. Koh KT, Law WC, Zaw WM, et al. Smartphone electrocardiogram for detecting atrial fibrillation after a cerebral ischaemic event: a multicentre randomized controlled trial. EP Europace 2021;23(7):1016–1023; doi: 10.1093/europace/euab036.

43. Rizas KD, Freyer L, Sappler N, et al. Smartphone-based screening for atrial fibrillation: a pragmatic randomized clinical trial. Nat Med 2022;28(9):1823–1830; doi: 10.1038/s41591-022-01979-w.

44. Gruwez H, Verbrugge FH, Proesmans T, et al. Smartphone-based atrial fibrillation screening in the general population: feasibility and impact on medical treatment. European Heart Journal - Digital Health 2023;4(6):464–472; doi: 10.1093/ehjdh/ztad054.

45. Calvert P, Mills MT, Howarth K, et al. Remote rhythm monitoring using a photoplethysmography smartphone application after cardioversion for atrial fibrillation. European Heart Journal - Digital Health 2024;5(4):461–468; doi: 10.1093/ehjdh/ztae028.

Table 1. Baseline Characteristics of Included Studies.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Study Design** | **Contry** | **Sample, N** | **AF rhythm diagnosed by ECG, %** | **Setting** | **Application** | **Gold Standard** | **Female sex, N (%)** | **Age, years mean ± SD or median (IQR)** |
| Mol D et al. 2020 | Validation study | Netherlands | 149 | 69.8 | Patients electing for ECV | Smartphone-based PPG algorithm iPhone 8 (Apple Inc., Cupertino, CA) | Standard ECG | 64 (42.9) | 69.0 ± 9.0 |
| Fernstad et al. 2024 | Prospective validation study | Sweden | 280 | 82.1 | Patients undergoing DCCV for treatment of persistent or recent-onset AF or AFL | CORAI Heart Monitor PPG software application iPhone 7 (Apple Inc., Cupertino, CA, USA) Corai Medicinteknik AB (Stockholm, Sweden) | Single-lead ECG | 86 (30.7) | 69.0 (13.6) |
| Krivoshei et al. 2017 | Case-control study | Switzerland | 80 | 50.0 | Inpatients and outpatients aged 18 years or older | RMSSD and PPAf Algorithm iPhone 4S (Apple, Inc., Cupertino, CA, USA) | Recorded ECG | 23 (28.8) | SR group: 75 ± 7 AF group: 80 ± 8 |
| Brasier et al. 2019 | Prospective clinical validation study | Switzerland and Germany | 592 | 41.9 | Hospitalized patients aged 18 years or older | PreventicusRhythmus, V1.0 iPhone 4S - Preventicus | Single-lead ECG | 268 (45.3) | 78.0 (13.0) |
| Chan et al. 2016 | Prospective screening study | Hong Kong | 1013 | 2.8 | Patients with history of hypertension and/or  diabetes mellitus or were ≥65 years of age | Cardiio Rhythm - iPhone | Single-lead ECG | 539 (53.2) | 68.4 ± 12.2 |
| Fan et al. 2019 | Prospective cross-sectional | China | 108 | 48.1 | Hospitalized patients aged 18 years or older | Preventicus - Mobile Phones (Huawei Technologies) | 12-lead ECG | SR group: 26 (46.0)  AF group: 19 (37.0) | SR group: 58.0 (14.8) AF group: 66.56 (13.17) |
| McManus et al. 2016 | Validation study | USA | 121 | 81.0 | Participants with AF who were scheduled to undergo elective cardioversion | PULSE-SMART - iPhone 4S | 12-lead ECG or 3-lead ECG | 28 (28.6)\* | 65.9 (12.2)\* |
| Mutke et al. 2021 | Prospective, blinded analysis used data from the previously published WATCHAF and DETECTAF pro trials | Switzerland and Germany | 1101 | 44.2 | Hospitalized patients aged 18 years or older | Preventicus - iPhone 4S | Single-lead ECG | NI | NI |
| Poh et al. 2018 | Retrospective cohort | Hong Kong | 1013 | 2.8 | Patients from MOBILE-SCREEN-AF (clinical validation) data set | DCNN model learnt in PPG waveforms - iPhone 4S | Single-lead ECG | 539 (53.2) | 68.4 (12.2) |
| Proesmans et al. 2019 | Prospective diagnostic accuracy study | Belgium | 223 | 45.7 | Patients aged 65 years or older with known paroxysmal or persistent AF, as well as those without a history of AF | FibriCheck - iPhone 5S | 12-lead ECG | 119 (53.4) | 77.0 (8.0) |
| Rozen et al. 2018 | Prospective, single-center study | USA | 98 | 88.8 | Consecutive patients diagnosed with AF who were scheduled for DCCV | Cardiio Rhythm Mobile Application (iPhone) | 12-lead ECG | 24 (24.5) | 67.7 ± 10.5 |
| Yan et al. 2018 | Prospective cohort | Hong Kong | 217 | 34.6 | Patients admitted to the cardiology ward | Cardiio Rhythm - iPhone 6S - Cardiio Rhythm app | 12-lead ECG | 62 (28.6) | 70.3 ± 13.9 |
| Henri Gruwez et al. 2024 | Pragmatic, prospective, blinded validation study | Belgium | 50 | 21.2 | Patients aged 18 years or older scheduled for AF ablation | Fibricheck (FC), Qompium NV, Hasselt, Belgium | Single-lead ECG | 16 (32.0) | 63.0 ± 11.0 |
| Jean-Philippe Couderc et al. 2022 | Validation study | NI | 45 | 41.2 | Patients aged ≥40 years with a diagnosis of paroxysmal, persistent, or permanent AF | VPG-based application - HealthKam AFibAndroid smartphone device (Samsung Galaxy S10) | Single-lead ECG | 13 (28.9) | 67.0 ± 9.7 |

*\*Atrial Fibrillation cohort; AF: Atrial Fibrillation; AFL: Atrial Flutter; SR: Sinus Rhythm; USA: United States of America; ECG: Electrocardiogram; PPG: Photoplethysmography; VPG: Videoplethysmography; ECV: Electrical Cardioversion; DCCV: Direct Current Cardioversion; RMSSD: Root Mean Square of Successive Differences; PPAf: Prediction Algorithm for AF; DCNN: Deep Convolutional Neural Network; IQR: Interquartile Range; SD: Standard Deviation; NI: Not Informed.*

**FIGURE LEGENDS**

**Figure 1.** PRISMA Flow Diagram of the Study Selection Process.

**Figure 2.** Summary Receiver Operating Characteristic Curve for the Diagnostic Performance of Smartphone-Based PPG in AF Detection.

*AF: Atrial Fibrillation; PPG: Photoplethysmography; AUC: Area Under the Curve; Se: Sensitivity; Sp: Specificity; SROC: Summary Receiver Operating Characteristic.*

**Figure 3.** Fagan Nomogram for Predicting AF Using Smartphone-Based PPG with Prevalence Estimates of 2% and 6%.

*AF: Atrial Fibrillation; PPG: Photoplethysmography; PLR: Positive Likelihood Ratio; NLR: Negative Likelihood Ratio.*

**Figure 4.** Projected Predictive Values of Smartphone-Based PPG for AF Detection with Prevalence Estimates of 2% and 6%.

*PPV: Positive Predictive Value; NPV: Negative Predictive Value.*

**Figure 5.** Bivariate Meta-Regression and Subgroup Analyses of Smartphone-Based PPG for AF Detection.

*\* Significance at p < 0.05; AF: Atrial Fibrillation; PPG: Photoplethysmography; ECG: Electrocardiogram; USA: United States of America*