

















## ORIGINAL RESEARCH

# Validation of a Deep Learning Algorithm for Continuous, Real-Time Detection of Atrial Fibrillation Using a Wrist-Worn Device in an Ambulatory Environment

Ming-Zher Poh , PhD\*; Anthony J. Battisti , PhD; Li-Fang Cheng , PhD; Janice Lin , PhD; Anil Patwardhan , PhD; Ganesh S. Venkataraman , MD; Charles A. Athill , MD; Nimesh S. Patel , MD; Chinmay P. Patel , MD; Christian E. Machado , MD; Jeffrey T. Ellis , PhD; Lori A. Crosson , PhD; Yuriko Tamura , MS; R. Scooter Plowman , MD, MBA, MHSA, MSc; Mintu P. Turakhia , MD, MS; Hamid Ghanbari , MD, MPH

**BACKGROUND:** Wearable devices may be useful for identification, quantification and characterization, and management of atrial fibrillation (AF). To date, consumer wrist-worn devices for AF detection using photoplethysmography-based algorithms perform only periodic checks when the user is stationary and are US Food and Drug Administration cleared for pre-diagnostic uses without intended use for clinical decision-making. There is an unmet need for medical-grade diagnostic wrist-worn devices that provide long-term, continuous AF monitoring.

**METHODS AND RESULTS:** We evaluated the performance of a wrist-worn device with lead-I ECG and continuous photoplethysmography (Verily Study Watch) and photoplethysmography-based convolutional neural network for AF detection and burden estimation in a prospective multicenter study that enrolled 117 patients with paroxysmal AF. A 14-day continuous ECG monitor (Zio XT) served as the reference device to evaluate algorithm sensitivity and specificity for detection of AF in 15-minute intervals. A total of 91 857 intervals were contributed by 111 subjects with evaluable reference and test data (18.3h/d median watch wear time). The watch was 96.1% sensitive (95% CI, 92.7%–98.0%) and 98.1% specific (95% CI, 97.2%–99.1%) for interval-level AF detection. Photoplethysmography-derived AF burden estimation was highly correlated with the reference device burden ( $R^2=0.986$ ) with a mean difference of 0.8% (95% limits of agreement, –6.6% to 8.2%).

**CONCLUSIONS:** Continuous monitoring using a photoplethysmography-based convolutional neural network incorporated in a wrist-worn device has clinical-grade performance for AF detection and burden estimation. These findings suggest that monitoring can be performed with wrist-worn wearables for diagnosis and clinical management of AF.

**REGISTRATION INFORMATION:** URL: <https://www.clinicaltrials.gov>; Unique identifier: NCT04546763.

**Key Words:** atrial fibrillation ■ atrial fibrillation burden ■ continuous monitoring ■ electrocardiogram ■ photoplethysmography

**W**earable device detection of atrial fibrillation (AF) episodes and AF burden may be useful for identification of undiagnosed AF and in

making clinical management decisions that include oral anticoagulation, restoration and maintenance of sinus rhythm, control of ventricular rate, and risk factor

Correspondence to: Hamid Ghanbari, MD, MPH, Department of Internal Medicine, Division of Cardiovascular Medicine, Cardiac Electrophysiology Services, University of Michigan, 1500 East Medical Center Dr, Ann Arbor, MI 48109-5853. Email: [gghamid@med.umich.edu](mailto:gghamid@med.umich.edu)

\*Dr. Poh is currently with Google, Cambridge, MA

This manuscript was sent to Luciano A. Sposato, MD, MBA, Associate Editor, for review by expert referees, editorial decision, and final disposition.

Supplemental Material is available at <https://www.ahajournals.org/doi/suppl/10.1161/JAHA.123.030543>

For Sources of Funding and Disclosures, see page 12.

© 2023 The Authors and iRhythm Technologies, Inc. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

JAHA is available at: [www.ahajournals.org/journal/jaha](http://www.ahajournals.org/journal/jaha)

## CLINICAL PERSPECTIVE

### What Is New?

- The study demonstrates performance of a prescription wrist-worn lead-I ECG and continuous photoplethysmography device for continuous monitoring coupled with a photoplethysmography-based algorithm in detection of atrial fibrillation (AF) episodes and estimation of AF burden.
- Unlike consumer-facing wrist-worn devices, which are US Food and Drug Administration cleared for pre-diagnostic purposes and not intended for clinical decision making, the findings of this study suggest that a wrist-worn device can support diagnosis and clinical management of AF.
- This can be achieved through continuous, nonperiodic ambulatory monitoring performed in conjunction with event- or symptom-based ECG recording.

### What Are the Clinical Implications?

- The prescription Study Watch bridges the gap between long-term continuous ECG monitoring and the consumer space with a practical solution for AF detection, burden assessment, and on-demand lead-I ECG acquisition based on irregular rhythm notification or other patient-activated reasons.
- This device may complement existing modalities by enabling AF monitoring over extended durations not currently possible with available noninvasive technologies.
- The Study Watch could be implemented in clinical practice for both ECG event recording and continuous AF monitoring. It can aid in screening, diagnosis, burden assessment, medication optimization, and characterization of episodes in patients with AF.

### Nonstandard Abbreviations and Acronyms

<b>CCT</b>	certified cardiographic technician
<b>FDA</b>	US Food and Drug Administration
<b>FP</b>	false positive

modification. Treatment of AF is associated with improvements of quality of life, stroke prevention, and reduced risk of heart failure.<sup>1</sup>

Ambulatory ECG monitoring is commonly used to manage patients with AF. However, the intermittent nature of AF can lead to low detection rates and

inaccurate estimation of AF burden when using short-term noninvasive monitoring modalities.<sup>2,3</sup> Implantable monitoring devices can increase the diagnostic yield by increasing the monitoring period; however, they are costly and require invasive procedures.<sup>4</sup> Several consumer wrist-worn devices have been cleared by the US Food and Drug Administration (FDA) for detection of AF based on photoplethysmography.<sup>5,6</sup> These devices are cleared as over-the-counter devices but only check for AF periodically when the user is still and do not provide real-time prompts for ECG capture that is clinically important for management of patients with AF. There is an unmet clinical need for a physician-prescribed, FDA-cleared wearable device that can continuously monitor and detect episodes of AF in real-time, notify patients to record an ECG when AF is suspected, and accurately assess the burden of AF for clinical use.

The Study Watch is a prescription-only, medical-grade wrist-worn device that is cleared by the FDA for 2 functions: (1) recording single-channel ECGs that are event- or symptom-triggered for event-recording capability; and (2) continuous monitoring of AF using a photoplethysmography-based algorithm in professional health care facility environments. The purpose of this study was to validate the performance of a new photoplethysmography-based algorithm on the Study Watch for detection of AF and assessment of AF burden against a reference patch ECG monitoring device in a diverse population of patients in an ambulatory environment during a 14-day monitoring period.

## METHODS

### Study Design

We conducted a multicenter, prospective study to evaluate the performance of the Study Watch and its photoplethysmography algorithm in detecting irregular rhythms suggestive of AF in subjects at risk for AF events in an ambulatory environment. Patients wore the Study Watch wrist-worn monitor (Verily Life Sciences LLC, South San Francisco, CA) along with a continuous ECG monitoring device (Zio XT, iRhythm Technologies, Inc., San Francisco, CA) for 14 days, hereafter referred to as the reference patch ECG.

The Study Watch is a wrist-worn wearable device capable of user-initiated single-lead (lead I) ECG acquisition for symptom- or event-triggered recording. The Study Watch also measures physiological parameters using a variety of sensors including photoplethysmography and accelerometer. It also includes electrodes that enable single-lead ECG acquisition. This design provides the ability to supplement ECG event-recording capability to continuously monitor for suspected AF events via the photoplethysmography

sensor and to notify patients in real-time to acquire an ECG for subsequent clinical confirmation of AF. The Study Watch has been used in multiple clinical studies such as the Project Baseline Health Study, which included >2500 patients. It has been shown to reliably record continuous data with no reported major clinical adverse events related to the use of the device.<sup>7</sup>

### Watch Photoplethysmography Algorithm Development

An on-watch algorithm was developed using deep convolutional neural networks trained to classify 15-minute photoplethysmography intervals into 4 categories (suspected AF, suspected non-AF, unanalyzable, and not enough data/off-wrist) and incorporated into the device firmware. The algorithm was trained using simulated photoplethysmography data derived from single-lead ECGs annotated by certified cardiographic technicians (CCTs). Suspected AF events detected by photoplethysmography in real-time generate irregular pulse notifications on the watch screen along with watch vibrations, prompting the user to acquire a single-lead ECG using the Study Watch.

The algorithm for estimation of AF burden uses a residual neural network trained using the same dataset as described for the on-watch algorithm but is run on server-side using encoded photoplethysmography signal features transferred from the watch. This approach allows for a more complex model to be used. Further details regarding model architecture and algorithm development are described in Data [S1](#).

### Eligibility Criteria

The study included adult participants (age  $\geq 22$  years) from 5 centers across the United States with a history of paroxysmal AF meeting 1 of the following criteria: planned AF ablation procedure, documented AF burden  $\geq 25\%$  in the last 3 months documented by implantable loop recorder, implantable cardioverter defibrillator, permanent pacemaker, Holter monitor, or adhesive monitoring patch results,  $CHA_2DS_2VAsc \geq 3$ , or left atrial diameter  $\geq 4.4$  cm. We excluded patients with implantable neurostimulators, cardiac implantable device in an atrial or ventricular paced rhythm documented on a 12-lead ECG during the initial study visit, or those patients successfully treated with antiarrhythmic medications or AF ablation with no documented AF of >30 seconds since the ablation or initiation of medications. Patients with known severe allergy to nickel, metal jewelry, adhesives, hydrogels, family history of adhesive skin allergies, open injury or rash at the device sites were also excluded. Subjects meeting eligibility criteria were required to provide informed consent.

### Study Procedures

Participants were provided the Study Watch and reference patch ECG, which were worn together for the entire duration of the 2-week ambulatory monitoring. The Study Watch was worn daily, including during sleep, except during showering, bathing, and water activities. It continuously recorded photoplethysmography signals from participants during the study period. The Study Watch was equipped with an investigational firmware that was designed to deliver notifications asking the user to take an ECG if an irregular rhythm suggestive of AF was detected (irregular pulse notifications). The notifications consisted of a series of vibration pulses on the watch lasting for 1 minute, in conjunction with the message “Please Take ECG” on the Study Watch screen. Participants were also asked to acquire once-daily ECG when no symptoms or notifications were present and could acquire additional ECGs for any symptomatic episodes.

At the completion of the 14-day wear period, the Study Watch and reference patch ECG were removed. Subjects completed a follow-up study visit between 14 and 20 days following enrollment to return the study devices and document changes in medical history, medication use, or vital signs. Institutional Review Board (IRB) approvals were obtained at all participating clinical sites and all subjects were asked to provide informed consent. The reviewing IRBs included Western IRB (Puyallup, WA), University of Texas Southwestern Medical Center IRB (Dallas, TX), and University of Pittsburgh Medical Center Pinnacle IRB (Harrisburg, PA).

### Study Visits and Data Collection

Demographics, relevant medical history information, medication use, and vital signs were documented at baseline using an electronic data capture system. Additionally, the subjects' skin tone and arm hair were assessed using the Fitzpatrick Skin Scale<sup>8</sup> and Arm Hair Index.<sup>9</sup> Following enrollment in the clinic, subjects were provided instructions for optimally taking ECGs with the Study Watch throughout the study period. An automated cellular-enabled network access point, which also functioned as a charging device for the watch, was provided to each subject for uploading device data to the cloud server. Reference patch ECG monitors were also provided to each enrolled subject along with instructions for wearing the patch.

### Statistical Analysis

The analysis population included subjects enrolled in the study who completed some portion of the 2-week follow-up period in whom at least 1 evaluable interval by the reference device was obtained. Subjects were

excluded from analyses if no evaluable reference intervals were obtained during the entire observation period.

The clinical outcome of interest was suspected AF occurring within distinct contiguous 15-minute time intervals. We used a 15-minute time interval to strike a balance between specificity and sensitivity. This approach minimizes false positives (FPs), because the watch is designed to be integrated into a process that includes human expert review, further reducing FPs and the likelihood of missed diagnoses. The 15-minute interval also enables timely AF episode detection and prompts the user to obtain an on-demand ECG for AF confirmation, which may not be feasible with longer time intervals. This optimized strategy can effectively identify clinically significant AF episodes, accurately estimate AF burden, and better delineate between AF and non-AF rhythms. It can also reduce the burden on health care professionals within the target population. For each subject, multiple intervals were available for analysis and were treated as repeated measurements. For each interval, the reference patch ECG identified whether AF was present (positive classification) or not present (negative classification) and this was considered ground-truth reference data for calculating accuracy statistics. Similarly, the test device labeled intervals as positive or negative.

Sensitivity and specificity for detection of AF episodes during a 15-minute interval served as co-primary end points and were based on the subset of intervals receiving positive and negative classification from both the reference and test devices (ie, either AF or non-AF). For the primary end point, we only included the intervals that were determined as analyzable. Intervals were excluded from the primary analysis if they were determined by the device to be unanalyzable or if the intervals were collected off-wrist, resulting in insufficient data. End points were evaluated based on tests of sensitivity and specificity, where the successful rejection of null hypotheses was based on prespecified thresholds for sensitivity ( $H_0 \leq 67\%$ ) and specificity ( $H_0 \leq 90\%$ ). A 1-sided test ( $\alpha=2.5\%$ ) was used to evaluate each of these co-primary end points.

The accuracy of the Study Watch photoplethysmography-based algorithm in detecting suspected AF was further investigated in subgroups defined by race, age group, activity level, Fitzpatrick scale score, and AF burden. Accuracy was also reported in terms of positive predictive value (PPV) and negative predictive value based on the overall AF burden observed in the study. Activity levels were defined based on the wrist-based actigraphy counts as light ( $<3.0$  metabolic equivalents, such as slow walking), moderate (3.0–5.9 metabolic equivalents, such as fast walking and yoga), and vigorous activity ( $\geq 6.0$  metabolic equivalents, such as running and high impact

aerobic dancing). A sensitivity analysis was performed to evaluate the impact of unevaluable and off-wrist device data by imputing missing test device data using best-case/worst-case scenarios (ie, all test-positive, all test-negative, all matching-reference, and all not matching reference). The percentage of evaluable intervals for each activity level was used to assess the impact of increased activity on signal quality and was determined as the number of evaluable test device intervals out of the total number of available reference device intervals.

Estimates of accuracy were calculated across intervals, while accounting for clustering of intervals within-subject. Logistic mixed-effect regression models, with a random effect added to account for the effect of clustering, were used to estimate sensitivity, specificity, PPV, and negative predictive value. 95% CIs were provided using bootstrap methods (cluster or block bootstrap) where the correlation structure is preserved, taking into account the fact that repeat intervals are nested within subject. Further details regarding CI estimation and logistic mixed-effect regression models are provided in Data S1.

A patient-level analysis was also performed to assess whether subjects with any AF during the period of follow-up could be correctly identified (ie, patient-level sensitivity); and similarly if patients who did not have any AF throughout the follow-up period could be correctly identified (ie, patient-level specificity). The unit of analysis here was the patient, and patients were assumed to be independent. CIs for patient-level analyses were calculated using Clopper-Pearson (exact) intervals.

AF burden was defined as the fraction of all intervals determined to be reference-positive in the study. We also defined photoplethysmography-based AF burden, which was the fraction of all evaluable test device intervals that were determined to be AF. Photoplethysmography-based estimates of AF burden were calculated as the proportion of AF-positive photoplethysmography intervals to analyzable intervals. We excluded unanalyzable intervals from the calculation of the photoplethysmography-derived AF burden. The reference AF burden was determined from the reference patch ECG that is calculated as the proportion of total time in AF to total analyzable time over the wear period. The Pearson correlation coefficient with mean square error and Bland–Altman methods were used to compare AF burden measurements between the reference and test device at the patient level.

## Sample Size Estimation

Previous clinical studies (NCT04074434) were used to inform the expected average patient wear time, test device performance, and AF burden distribution

across patients for the present study. These parameters allowed us to estimate the variance inflation factor (ie, design effect) necessary to derive sample size estimates in a repeated-measures setting.

A sample size of 110 subjects was necessary to provide 80% power to evaluate our primary end points. Additional subjects were enrolled to account for potential loss to follow-up. We used an adaptive design, where AF burden by the reference device was assessed at prespecified times during the study to ensure that the number of patients would provide sufficient power to detect our primary end points. Detailed sample size estimate calculations are documented in Data S1.

### Additional Analysis of Watch ECG Performance

A deep neural network capable of detecting arrhythmias from single-lead ECGs was used to label AF episodes using the Study Watch lead-I ECG data.<sup>10</sup> Given that each subject recorded a variable number of ECG samples and the resulting pool was large, a sample of ECGs from all the obtained Study Watch ECGs were selected to maximize the number of ECGs from distinct patients and to yield an approximately equivalent number of ECGs with and without AF. Two CCTs independently annotated AF episodes within sampled ECGs. These were validated by an electrophysiologist to produce a triple-reviewed reference dataset. Sensitivity and PPV of the ECG-based algorithm for detection of AF episodes within the sampled ECG data were determined using the clinician-annotated episodes as the reference standard.

## RESULTS

Study subjects were enrolled over the period between September 9, 2020 and May 14, 2021. Of 117 enrolled participants, there were 111 that contributed data to the analysis per protocol. Reasons for the 6 participants not included in the analysis included 3 patients who did not wear the ECG patches, 1 patient lost to follow-up and ECG patch was not returned, 1 with <15 minutes of Study Watch wear time, and 1 patient who withdrew during the Study Watch device training due to hand tremor preventing ECG acquisition. A total of 91857 intervals with simultaneous analyzable reference and test data were used to assess primary end points (per-interval sensitivity and specificity).

### Clinical Characteristics

Of the 111 patients included in the study, 50 (45%) were women, and the mean age was 65±11 years (Table 1). Median watch wear time was 18.3h/d

**Table 1. Baseline Characteristics**

Characteristics	n=111
Hypertension, n (%)	74 (66.7%)
Coronary artery disease, n (%)	15 (13.5%)
Congestive heart failure, n (%)	13 (11.7%)
Peripheral vascular disease, n (%)	2 (1.8%)
TIA or stroke, n (%)	16 (14.4%)
Obstructive sleep apnea, n (%)	24 (21.6%)
Chronic kidney disease, n (%)	3 (2.7%)
Diabetes	16 (14.4%)
Female, n (%)	50 (45.0%)
Age, mean±SD, y	65±11
BMI, mean±SD, kg/m <sup>2</sup>	29±6
Fitzpatrick skin tone n, (%)	
I	10 (9.0%)
II	17 (15.3%)
III	42 (37.8%)
IV	26 (23.4%)
V	11 (9.9%)
VI	5 (4.5%)
Black/African American, n (%)	13 (11.7%)
CHA2DS2-VASc score, mean±SD	2.6±1.6
History of atrial fibrillation ablation, n (%)	2 (1.8%)
Antiplatelet therapy, n (%)	2 (1.8%)
Oral anticoagulants, n (%)	18 (16.2%)

BMI indicates body mass index; and TIA, transient ischemic attack.

(25th percentile=14.4 h/d; 75th percentile=23.4 h/d; interquartile range [IQR] range=9.1 h/d). Participants obtained 1.9±2.1 ECGs per day using the Study Watch (median: 1.0; IQR, 1.0 [25th percentile=1.0; 75th percentile=2.0]) for a total of 3448 watch ECGs during the study period. The median analyzable photoplethysmography recording time was 13.4 h/d (25th percentile=7.6 h/d; 75th percentile=19.8 h/d; IQR range=12.2 h/d).

There were 34186 hours of ECG recording obtained from the reference patch ECG devices. There were 34 of 111 (30.6%) participants who had at least 1 analyzable reference positive AF interval; 12 subjects had an AF burden of ≥25%. The average duration of an AF episode was 109 minutes. The median AF burden during the study period in patients with analyzable AF intervals was 8.9% (25th percentile=1.1%; 75th percentile=29.9%; IQR=28.8%).

### Performance of Photoplethysmography-Based Algorithm

Of all photoplethysmography intervals, 7.2%, 70.0%, and 22.8% were classified by the Study Watch algorithm as suspected AF, not suspected AF, and unanalyzable. FP and false negative intervals comprised 1.7%

and 0.3%, respectively, of total intervals. The primary end point analysis resulted in an interval-level sensitivity of 96.1% (95% CI, 92.7%–98.0%) and a specificity of 98.1% (95% CI, 97.2%–99.1%) for AF detection. In the worst-case scenario where the unanalyzable Study Watch intervals were assumed to not match reference patch ECG device labels, estimated sensitivity was 75.5%. The PPV and negative predictive value for the Study Watch algorithm were 81.4% (95% CI, 68.8%–89.9%) and 99.7% (95% CI, 99.5%–99.8%), respectively, based on the observed burden of AF (7.9%).

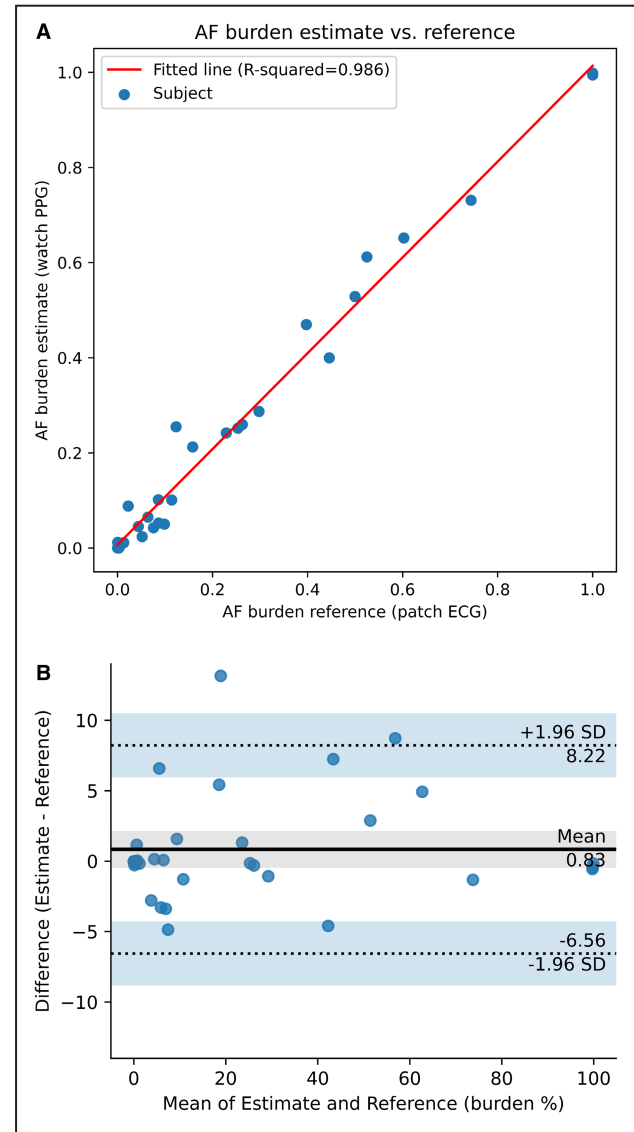
Two subjects were responsible for 41% (651 out of 1599) of the FP intervals in our study. The predominant rhythm from patch ECG for these 2 subjects was normal sinus rhythm. However, we observed higher interbeat interval variability for these 2 subjects, and the FP intervals generally, relative to the variability of true negative intervals (Table S1). We observed unlabeled ectopic atrial beats as the major reason for FPs in these individuals (Figure S1). In the majority of FP intervals, subjects had a low burden of AF, because only 25.9% of the false negative intervals had an aggregated AF duration exceeding 7.5 minutes. Sinus rhythm was the dominant rhythm in 72.7% of these intervals (Figure S2).

We evaluated subject-level sensitivity and specificity where a participant is considered reference positive if any intervals are observed to have AF by the reference patch ECG device throughout the study period. Subject-level sensitivity was 90.6% (95% CI, 74.9%–98%) and specificity was 60.8% (95% CI, 49.1%–71.6%). Additional sensitivity analysis was performed to evaluate the impact of unanalyzable and off-wrist test intervals by imputing missing test device data (Table S2). In the worst-case scenario where the unanalyzable and off-wrist test intervals were assumed to not match reference patch ECG device labels, estimated sensitivity was 69.2%. However, overall subject-level sensitivity remains high (90.6%). The photoplethysmography-based algorithm identified all but 3 subjects with at least 1 reference-positive interval during the 14-day study period. All 3 false-negative patients had a very low AF burden (0.09%–0.25%). Two subjects only had a single episode of AF (116 seconds and 530 seconds) while the third subject had 4 very short episodes of AF (31–104 seconds) during the study period, translating into a single analyzable AF interval per subject. All 3 false-negative intervals occurred during light activity.

### Performance of AF Burden Estimation

A total of 62952 30-minute photoplethysmography intervals were obtained for AF burden algorithm analysis, of which 79.6% were analyzable. For subjects with 1 or more AF episodes confirmed by reference patch ECG (34/111, 30.6%), the photoplethysmography-derived

AF burden estimation was highly correlated with the AF burden obtained from the reference patch ECG device via Pearson correlation coefficient ( $R^2=0.986$ ; Figure 1 A). Additionally, Bland–Altman analysis indicated agreement between the photoplethysmography and the reference patch ECG estimates of AF burden with a mean difference of 0.8% and 95% limits of agreement



**Figure 1. Correlation and agreement of AF burden methods.** **A**, Scatter plot and regression line of the PPG-derived AF burden estimates vs AF burden determined by the reference patch ECG. The Pearson correlation coefficient ( $R^2$ ) was 0.986. **B**, Bland–Altman analysis shows agreement between quantification of AF burden as estimated by the PPG-based algorithm and as determined from the reference patch ECG. Blue dots represent the difference in percent AF burden between methods for each AF-positive subject. The continuous line is the average value of difference between AF burden calculation methods (0.8%), and the dotted lines represent the upper and lower 95% limits of agreement (ie, the interval of 2 SDs of the measurement differences on either side of the average difference; 8.2%, –6.6%). AF indicates atrial fibrillation; and PPG, photoplethysmography.

of -6.6% to 8.2% (Figure 1B). The total off-wrist time for the Study Watch, when reference patch ECG recordings were available, was 3337.8 hours. During this off-wrist period, the reference patch ECG detected 244.2 hours of AF, representing 7.3% of the total off-wrist duration, consistent with the observed overall AF burden. We did not observe a significant difference between an estimated AF burden when patch ECG was used as the denominator versus analyzable photoplethysmography intervals (7.2% versus 7.9%).

### Performance of Algorithm for Classification of Study Watch ECGs

There were a total of 2448 notifications generated (median number of notifications was 2 per day [25th percentile=1 per day; 75th percentile=5 per day; IQR range=4 per day]). The median rate of recording a watch ECG in response to an irregular rhythm notification per subject day was 85.7% (25th percentile=38.5%; 75th percentile=100%; IQR range=61.5%). We sampled 371 ECG strips acquired from all available ECGs (3510; 10.6%). The ECGs included those that were self-initiated or in response to daily prompts, as well as those that were triggered by irregular rhythm notifications. The ECG classification algorithm correctly identified all AF episodes as annotated by the clinician panel, indicating sensitivity of 100% for the sampled ECG data. Episode-level PPV was 54.9%.

### Subgroup Analyses

We evaluated the Study Watch AF detection performance during different levels of activity. Vigorous activity was observed in 55 subjects during the study period, representing 1.7% of total intervals. Light and moderate activity was observed in 111 and 107 patients, respectively, and they represented 98.5% of overall intervals. A lower sensitivity for detection of AF during vigorous activity compared with light and moderate activity was observed (Table 2). There were only 4 reference positive AF intervals available for calculating sensitivity during vigorous activity, 2 of which were detected by the Study Watch. The specificity for detection of AF was comparable during low, moderate, and vigorous activity.

The Study Watch had comparable performance for detection of AF in patients with high burden (>25%) of

**Table 3. Subgroup Interval-Level Sensitivity and Specificity by AF Burden**

Subgroup	Sensitivity (95% CI)	Specificity (95% CI)
AF burden decile		
Highest decile (≥25.4% burden) (N=12)	98.8 (97.8–99.5)	90.3 (77.4–99.8)
Not highest decile (<25.4% burden) (N=99)	81.6 (69.1–90.6)	98.5 (97.8–99.3)
Age group		
<65 y (N=45 subjects)	94.9 (84.1–98.7)	99.1 (98.5–99.9)
65–74 y (N=43 subjects)	95.3 (83.7–98.1)	98.4 (97.5–99.6)
≥75 y (N=23 subjects)	98.4 (95.8–99.4)	95.4 (92.2–98.9)
Racial group		
Asian (N=6)	98.7 (97.7–99.4)*	100.0 (98.7–100.0)*
Black/African American (N=13)	67.5 (55.4–83.5)	99.1 (98.8–99.5)
Native Hawaiian or other Pacific Islander (N=1)	NA	100.0 (99.6–100.0)*
White (N=86)	97.2 (95.0–98.4)	97.8 (96.7–99.0)
Other (N=5)	NA	100.0 (99.9–100.0)*

AF indicates atrial fibrillation; and NA, not assessable.

\*Due to absent or low error counts, CIs for sensitivity and/or specificity for some groups (Asian, Native Hawaiian/Pacific Islander groups, or Other) are reported based on the exact binomial rather than cluster bootstrap.

AF relative to patients with low burden (<25%) of AF (Table 3). There were no substantial differences in Study Watch AF detection performance in patients age <65, 65 to 74, or ≥75 years (Table 3).

Interval-level sensitivity was lower in the Black/African American subgroup, which included only 3 patients with episodes of AF (Table 3). We explored the impact of AF duration on interval-level sensitivity and specificity of AF across different racial groups. Higher sensitivity was observed for longer-duration AF (>6 minutes) as compared with short-duration AF (≤6 minutes), regardless of racial group. There was an increase in the percentage of intervals with short-duration AF (≤6 minutes) among the Black/African American subgroup (Table 4). Independent of the duration of AF, at the patient level, the device correctly identified all Black/African American patients with AF events during the monitoring period.

The proportion of analyzable observed test intervals were reduced with increasing Fitzpatrick scale (Table 4); however, interval-level sensitivity and

**Table 2. Sensitivity and Specificity by Interval-Level Activity as Determined by Study Watch Actigraphy Data**

Activity level	Total intervals with actigraphy	Number of evaluable intervals by study watch (%)	Sensitivity (95% CI)	Specificity (95% CI)
Low (observed in 111 subjects)	107 309 (87.9% of total intervals)	88 068 (82.1%)	96.2 (93.1–98.1)	98.2 (97.4–99.1)
Moderate (observed in 107 subjects)	12 818 (10.5% of total intervals)	3586 (28.0%)	94.3 (85.4–97.8)	96.6 (94.6–99.2)
Vigorous (observed in 55 subjects)	2022 (1.7% of total intervals)	203 (10.0%)	*NA	95.7 (92.1–98.2)

NA indicates not assessable.

\*There were 4 reference-positive atrial fibrillation intervals available for calculating sensitivity. Two of these were detected by the test device.

**Table 4. Interval-Level Sensitivity by Race and AF Interval Duration**

Racial group*	Number of analyzable AF intervals	AF ≤6 min		AF >6 min	
		% Analyzable AF intervals with AF ≤6 min	Sensitivity (95% CI)	% Analyzable AF intervals with AF >6 min	Sensitivity (95% CI)
Asian	854	0.5% (4/854)	0.0 (0.0–60.2) <sup>†</sup>	99.5% (850/854)	99.2 (98.3–99.7) <sup>†</sup>
Black/African American	314	45.9% (144/314)	31.9 (20.2–57.8)	54.1% (170/314)	97.8 (80.0–100.0)
White	6118	2.0% (125/6118)	16.7 (10.8–23.4)	98.0% (5993/6118)	98.8 (97.9–99.5)

AF indicates atrial fibrillation.

\*There were no evaluable reference-positive intervals within the Native Hawaiian/Pacific Islander or Other groups, so sensitivity could not be determined for these groups.

<sup>†</sup>For the Asian group, some cluster bootstrapped estimates of CIs were incalculable or resulted in estimates of [0, 1] due to most events occurring in only a few subjects. The CIs reported above are estimated based on the exact binomial.

specificity was maintained across the Fitzpatrick scales (Table 5). Patient-level sensitivity remained high across Fitzpatrick scale subgroups (Table 5). Furthermore, the time to detect AF and photoplethysmography-based estimation of AF burden was similar across different Fitzpatrick scales (Table 6).

All 111 subjects included in the analysis had some unanalyzable intervals. Only 10 subjects among the 111 subjects had >50% unanalyzable data (Figure S3). The mean percentage of unanalyzable intervals at the subject level was 23.7% (min, 0.8%; max, 91.2%; median, 19.8%; and IQR, 22.3%). There were 27 115 unanalyzable intervals (22.8% of 118 972 intervals with patch ECG as ground truth). Moderate or vigorous activity occurred during 37% of the unanalyzable intervals. The detailed breakdown of unanalyzable intervals by activity, arm hair index, and skin tone subgroups are provided in Tables S3 through S5.

## DISCUSSION

The Study Watch and its photoplethysmography-based algorithm had high interval-level sensitivity and specificity for detection of AF and provided an accurate estimation of AF burden in an at-home environment compared with continuous ECG monitors during the

14-day study period. This is the first study to report the performance of a medical grade, FDA-cleared device in different subgroups in an ambulatory environment with continuous nonperiodic photoplethysmography-based AF monitoring when added to wrist-worn ECG event-recording capability.

The Study Watch detected AF on an interval level with high sensitivity (96.1%) and specificity (98.1%). There have been several clinical studies validating the performance of photoplethysmography-based algorithms for detection of AF with the range of sensitivity and specificity between 85% and 99%.<sup>11–18</sup> These studies were limited by short photoplethysmography recording time and were conducted in controlled clinical settings unlike the ambulatory environment where these devices are most likely to be used. There have been several recent studies that reported the validity of different photoplethysmography-based algorithms for detection of AF in an ambulatory environment.<sup>6,19,20</sup> In an exploratory analysis of 1057 patients in the Fitbit Heart Study where the 1-week ECG patch monitor was used as the reference, the episode-level sensitivity was 67.6% (95% CI, 62.4%–72.6%) and specificity was 98.4% (95% CI, 97.3%–99.2%).<sup>20</sup> The Apple Heart Study reported a PPV of 71% (97.5% CI, 69%–74%); however, it did not report the episode or interval-level

**Table 5. Interval-Level and Patient-level Sensitivity and Specificity by Fitzpatrick Scale Score**

Fitzpatrick scale score	Unevaluable intervals/ total number of intervals (%)	Interval-level sensitivity (95% CI)	Interval-level specificity (95% CI)	Patient-level sensitivity (95% CI)	Patient-level specificity (95% CI)
I (10 subjects)	1696/11 825 (14.3%)	98.4 (96.8–99.4) <sup>†</sup>	98.0 (96.6–99.6)	100.0 (2.5% - 100.0)	55.6 (21.2–86.3)
II (17 subjects)	3600/19 175 (18.8%)	94.2 (87.5–98.3)	97.5 (95.3–100.0)	100.0 (29.2–100.0)	85.7 (57.2–98.2)
III (42 subjects)	8610/42 975 (20.0%)	97.4 (92.2–99.2)	99.5 (99.1–99.8)	87.5 (61.7–98.4)	61.5 (40.6–79.8)
IV (26 subjects)	6257/27 686 (22.6%)	95.4 (83.6–97.8)	95.9 (93.0–99.4)	87.5 (47.3–99.7)	50.0 (26.0–74.0)
V/VI* (16 subjects)	6952/17 311 (40.2%)	92.5 (56.7–99.1)	99.5 (99.3–99.8)	100.0 (39.8–100.0)	50.0 (21.1–78.9)

\*Only 5 subjects were observed with a Fitzpatrick scale score of VI. Due to the low numbers, subjects with Fitzpatrick scale scores of V and VI were combined into a single group.

<sup>†</sup>Low number of subjects in this group precluded the use of cluster bootstrap estimates of CIs for sensitivity. CIs are instead reported above based on exact binomial.

Though the Fitzpatrick Category I subgroup includes 10 total subjects, only a single subject in this subgroup had AF episodes during the study, thus contributing to the wide confidence interval for sensitivity among Fitzpatrick Category I subjects.



**Table 6. AF Time to Detect and Accuracy of Burden Estimation by Skin Tone**

Fitzpatrick skin type	Number of subjects (with AF detected/ total)	Time (days) from first AF episode to first true positive prompt for ECG (mean±SD)	Absolute error of AF burden (mean±SD)
I/II	(4/27)	0.7±1.2	0.5±1.2
III	(14/42)	1.2±1.7	0.5±1.1
IV	(6/26)	0.1±0.1	1.0±3.5
V/VI	(4/16)	1.1±1.7	1.0±2.6

AF indicates atrial fibrillation.

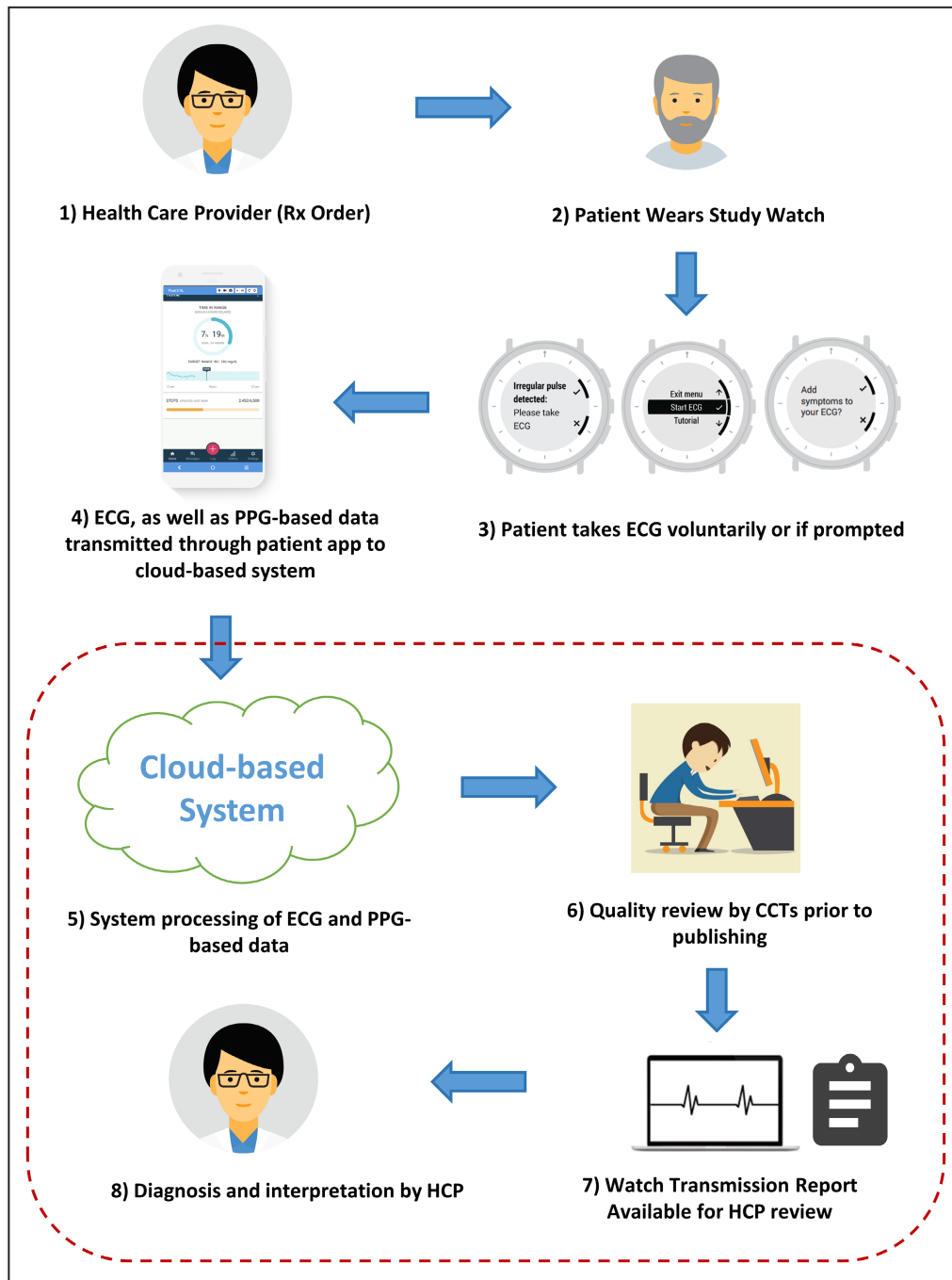
sensitivity of the photoplethysmography-based algorithm for detection of AF.<sup>6</sup> A recent study reported that the Apple Watch algorithm had an episode-level sensitivity of 60.0%.<sup>21</sup> The My Heart Lab Validation study evaluated the performance of the Samsung watch against a continuous ECG patch during a 4-week study period in an at-home environment. The study reported 5-minute interval-level sensitivity of 86.4% (95% CI, 81.8%–90.0%) and specificity of 97.4% (95% CI, 97.1%–97.7%). The sensitivity and specificity changed to 79.4% (95% CI, 68.1%–87.5%) and 99.4% (95% CI, 99.0%–99.5%) when the 2 out of 3, 5-minute intervals were used for analysis.<sup>19</sup> Our results show significantly higher 15-minute interval-level sensitivity for detection of AF in a similar population in an at-home environment. This is the first study to report subject-level sensitivity and specificity for AF detection using a wrist-worn device in an ambulatory environment. Although the Study Watch’s high subject-level sensitivity allows for detection of the majority of clinically relevant AF episodes, the subject-level specificity of 60.8% may produce a large number of FP alarms that may increase the burden on clinicians. The Study Watch is connected to an end-to-end system that includes real-time prompts for a user to record a single-lead ECG from the Study Watch when AF is suspected by the photoplethysmography algorithm and automated ECG rhythm classification followed by validation by CCTs (Figure 2). A comprehensive report is created and sent to the prescribing physician to confirm diagnosis of AF. The ability to continuously monitor for AF, coupled with ECG recording capability and CCT review, can provide timely and clinically relevant ECGs to the clinician. This is distinct from consumer grade wearables that currently only periodically assess irregular pulse and have no user prompting or ECG event-recording capabilities.

In the Study Watch ECGs sampled for further evaluation, sensitivity of the ECG classification algorithm was 100% in identification of AF episodes, indicating a low likelihood of false negatives in this population. PPV of the ECG classification algorithm was 54.9%; however,

this represents only an initial step in event-triggered ECG processing. Following automated identification of AF within the watch ECGs, the Study Watch implementation involves review and episode classification by CCTs before inclusion in a report to the prescribing physician. This multistage, end-to-end system with CCT overread may limit the burden on clinicians and avoid the deluge of ECGs and irregular pulse notifications generated by other wrist-worn devices because clinically actionable findings are found in a minority of patients.<sup>22</sup> This workflow model is analogous to what is currently used and reimbursed for a range of ambulatory ECG monitoring service types, including Holter, long-term continuous, mobile cardiac telemetry, and event monitors. Future work will evaluate the performance of this end-to-end system prospectively to assess its effectiveness.

When assessing the performance of wrist-worn wearable devices in an ambulatory setting, it is important to monitor patients during all activities of daily living. The algorithm used in the Apple Heart Study recorded photoplethysmography signals for 1 minute every 2 hours when the patient was completely stationary and therefore did not provide continuous photoplethysmography monitoring during the study period.<sup>6</sup> Similarly, the algorithm in the Fitbit Heart Study did not continuously monitor the photoplethysmography for AF but only when patients were stationary to reduce the FP rates and increase specificity for detection of AF.<sup>20</sup> The photoplethysmography-based algorithm used in the My Heart Validation Study included intervals during various levels of activity and only excluded them if the photoplethysmography signal was unanalyzable. However, they did not report the performance of the algorithm during different levels of activity and the device has yet to receive FDA clearance.<sup>19</sup> Our study is the first to report the performance of a photoplethysmography-based algorithm during different levels of activity as determined by the device accelerometer. Sensitivity and specificity for detection of AF were comparable during low and moderate levels of activity. Reduced interval-level sensitivity during high levels of activity for AF detection was observed; however, this represents a small portion of the overall monitoring period (1.7% of total intervals in our study). Therefore, the Study Watch photoplethysmography algorithm will be able to detect the majority of clinically relevant episodes of AF (ie, AF episode duration >6 minutes).

The Study Watch photoplethysmography-based algorithm correctly identified most patients with clinically relevant AF episodes (AF episode duration >6 minutes) during the 14-day monitoring period (29/32, 90.6%). AF episode durations >6 minutes have been shown to be a clinically relevant threshold when considering initiating anticoagulation in patients with a new diagnosis of AF during long-term monitoring, and it has been used



**Figure 2.** The Study Watch is a prescription device connected to an end-to-end system that includes real-time prompts for a user to record a single-lead ECG when AF is suspected by the on-watch PPG algorithm.

The acquired ECGs are transmitted to a cloud server where automated rhythm classification is reviewed by human experts. Only validated findings are then shared with the prescribing physician through reports. CCTs indicates certified cardiographic technicians; HCP, health care provider; and PPG, photoplethysmography.

as a clinical end point in several large clinical trials.<sup>4,23</sup> In a large study of individuals with paroxysmal atrial fibrillation, the majority of patients (96.2%) experienced at least 1 long-duration AF episode within a 2-week monitoring period. Patients with short AF episodes

tend to have many episodes, up to >10 events per day for AF episodes ≤6 minutes.<sup>24</sup> Thus, even though sensitivity for detecting an individual photoplethysmography interval with short-duration AF <6 minutes was lower, the effective sensitivity increases to 90%

with 6 intervals of short-duration AF, and 97.8% with 10 intervals of short-duration AF. Therefore, interval-level performance is maintained when focusing on clinically relevant AF episodes.

Our subgroup analyses also revealed a lower interval-level sensitivity in the Black/African American subgroup. The sensitivity for intervals with short-duration AF ( $\leq 6$  minutes) was lower compared with long-duration AF ( $> 6$  minutes) across all racial groups. Because a greater percentage of intervals with short-duration AF were detected in Black/African American patients, this may account for the overall reduction in interval-level sensitivity in this subgroup. Although prior research has demonstrated a higher incidence of AF among White individuals as compared with Black/African American individuals, there is no prior literature showing differences in AF episode durations among different racial subgroups.<sup>25</sup> The observed differences in our study may also be related to the low number of Black/African American patients with AF events during the monitoring period in our study. In our study, independent of the duration of AF, at the patient level, the device correctly identified 100% of Black/African American patients with AF events during the monitoring period due to enhanced opportunity to detect AF with continuous monitoring. Furthermore, the time to detect AF was similar across the racial subgroups. However, the current dataset limits our ability to evaluate the relative importance of these effects and other potential confounders due to a small number of patients with Fitzpatrick Type V and VI.

Photoplethysmography signals are sensitive to external factors such as motion artifacts and variations in peripheral circulation, which can lead to unanalyzable data. This has been reported in multiple studies published in the literature with the number of unanalyzable photoplethysmography intervals ranging between 32.2% and 93.4%.<sup>19,26–29</sup> Our device demonstrates one of the lowest percentages of unanalyzable photoplethysmography intervals in the published literature, resulting in identification of all but 3 subjects with at least 1 reference-positive interval during the study period.

Accurate and reliable estimation of AF burden requires long-term monitoring and is critical for appropriate management of patients with AF. Previous studies have demonstrated a dose–response relationship between AF burden, symptoms, heart failure, and stroke.<sup>30–33</sup> The median wear time of 18.3h/d in this study shows that it is feasible to continuously monitor heart rhythm using a wrist-worn wearable device. Our photoplethysmography-based algorithm allows for continuous monitoring during the day compared with other FDA-cleared consumer devices that provide non-continuous monitoring during the day and include only periods of time where the patient is stationary.<sup>6,20</sup> Our results show that one can accurately estimate AF burden over long-term monitoring periods in an ambulatory,

at-home setting via photoplethysmography-based wrist-worn devices.

This device offers an opportunity to continuously and noninvasively monitor patients outside of the clinical setting for weeks, months, or years. Longer monitoring periods provide multiple opportunities to detect AF episodes and improve the ability to detect clinically relevant AF in patients with or suspected of having AF. It provides a noninvasive option for long-term monitoring of patients with infrequent symptoms or history of cryptogenic stroke suspected of having AF. It is not intended to replace interval ECG monitoring where continuous ECG is needed but rather aims to bridge long-term continuous ECG monitoring and the consumer space with a practical, easy solution for assessment of AF burden that also allows the ability to transmit ECGs based on AF notifications or for other patient-activated reasons. As such, we anticipate that this device will complement the existing devices by providing an opportunity to monitor for AF during longer periods that are not currently possible with the existing noninvasive technologies such as patch ECG. In clinical practice, most individuals who use wearables do not use it for screening or diagnosis but rather for management of AF.<sup>5</sup> The Study Watch can be implemented in clinical practice to provide both ECG event-recording capability and continuous AF monitoring to improve management of patients with established diagnosis of AF through assessment of burden, characterization of AF episodes, optimizing heart rate control and medications. There is a need for prospective, randomized clinical trials focused on wearable use for management of patients with established diagnosis of AF to fully understand the association of wearables with health outcomes.

## Limitations

The study did not include a large number of patients with Fitzpatrick scale V/VI who had episodes of AF during the study period. Although substantial differences in performance of our algorithm for detection of clinically significant AF in this population were not observed, future studies with a larger number of patients with AF and Fitzpatrick V/VI are needed to further evaluate algorithm performance in this population. Although there were  $> 90\,000$  intervals with simultaneous analyzable reference and test data and this was powered at the episode level, the AF episodes came from a smaller number of participants who had AF. Future clinical studies including a larger cohort of patients with AF are needed to further validate findings in a clinical setting. The wrist-worn device relies on photoplethysmography signals, which can be influenced by motion artifacts and peripheral circulation, potentially leading to unanalyzable periods. Exclusion of unanalyzable periods may affect the overall performance of the wrist-worn device. The study

population did not include patients with ventricular pacing; therefore, performance results of our algorithm cannot be generalized to this patient population. Finally, this study did not evaluate the impact of end-to-end systems including irregular pulse notifications, on-demand ECG acquisition via the Study Watch with CCT over-reads on rates of false positive notifications, overall health care utilization, and cost.

## CONCLUSIONS

The results have demonstrated that the Study Watch, a wrist-worn device with a lead-I ECG and continuous photoplethysmography capability and algorithm, can accurately detect AF in a free-living environment. Our findings also suggest that the Study Watch, in combination with cloud-based assessment, is effective in providing estimation of AF burden. Furthermore, the algorithm for classification of symptom- or event-triggered ECGs obtained from the wrist-worn device and confirmatory CCT readings may provide critical information around the AF episodes, and provide the context necessary for health care professionals to effectively manage patients with AF.

## ARTICLE INFORMATION

Received May 31, 2023; accepted August 4, 2023.

### Affiliations

Verily Life Sciences, South San Francisco, CA (M.-Z.P., L.-F.C., J.L., A.P., R.S.P., H.G.); iRhythm Technologies, San Francisco, CA (A.J.B., J.T.E., L.A.C., Y.T., M.P.T.); Colorado Heart and Vascular, Lakewood, CO (G.S.V.); San Diego Cardiac Center, San Diego, CA (C.A.A.); University of Texas Southwestern Medical Center, Dallas, TX (N.S.P.); University of Pittsburgh Medical Center Harrisburg, Harrisburg, PA (C.P.P.); Ascension Providence Hospital, Southfield, MI (C.E.M.); Stanford University Medical Center, Palo Alto, CA (R.S.P.); and University of Michigan, Ann Arbor, MI (H.G.).

### Acknowledgments

The authors would like to thank Leera Choi for her role in managing the study documentation and for her assistance in reviewing the manuscript.

### Sources of Funding

The study was funded and sponsored by Verily Life Sciences.

### Disclosures

AJB, JTE, LAC, YT, and MPT are paid employees of iRhythm Technologies. HG is a paid consultant for Verily Life Sciences, Boston Scientific, Johnson and Johnson and Huxley Medical Inc. L-FC, JL, AP, and RSP are paid employees of Verily Life Sciences. M-ZP is a paid employee of Google. The remaining authors have no disclosures to report.

### Supplemental Material

Data S1  
Tables S1–S5  
Figures S1–S3

## REFERENCES

- January CT, Wann LS, Alpert JS, Calkins H, Cigarroa JE, Cleveland JC, Conti JB, Ellinor PT, Ezekowitz MD, Field ME, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *Circulation*. 2014;130:2071–2104. doi: 10.1161/CIR.0000000000000040
- Kaura A, Sztrihla L, Chan FK, Aeron-Thomas J, Gall N, Piechowski-Jozwiak B, Teo JT. Early prolonged ambulatory cardiac monitoring in stroke (EPACS): an open-label randomised controlled trial. *Eur J Med Res*. 2019;24:25. doi: 10.1186/s40001-019-0383-8
- Rosenberg MA, Samuel M, Thosani A, Zimetbaum PJ. Use of a noninvasive continuous monitoring device in the management of atrial fibrillation: a pilot study. *Pacing Clin Electrophysiol*. 2013;36:328–333. doi: 10.1111/pace.12053
- Reiffel JA, Verma A, Kowey PR, Halperin JL, Gersh BJ, Wachter R, Pouliot E, Ziegler PD; REAL AF Investigators. Incidence of previously undiagnosed atrial fibrillation using Insertable cardiac monitors in a high-risk population: the REVEAL AF study. *JAMA Cardiol*. 2017;2:1120–1127. doi: 10.1001/jamacardio.2017.3180
- Wang L, Nielsen K, Goldberg J, Brown JR, Rumsfeld JS, Steinberg BA, Zhang Y, Matheny ME, Shah RU. Association of wearable device use with pulse rate and health care use in adults with atrial fibrillation. *JAMA Netw Open*. 2021;4:e215821. doi: 10.1001/jamanetworkopen.2021.5821
- Perez MV, Mahaffey KW, Hedlin H, Rumsfeld JS, Garcia A, Ferris T, Balasubramanian V, Russo AM, Rajmane A, Cheung L, et al. Large-scale assessment of a smartwatch to identify atrial fibrillation. *N Engl J Med*. 2019;381:1909–1917. doi: 10.1056/NEJMoa1901183
- Arges K, Assimes T, Bajaj V, Balu S, Bashir MR, Beskow L, Blanco R, Calif R, Campbell P, Carin L, et al. The project baseline health study: a step towards a broader mission to map human health. *NPJ Digit Med*. 2020;3:1–10. doi: 10.1038/s41746-020-0290-y
- Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol*. 1988;124:869–871. doi: 10.1001/archderm.1988.01670060015008
- von Schuckmann LA, Hughes MC, Green AC, van der Pols JC. Forearm hair density and risk of keratinocyte cancers in Australian adults. *Arch Dermatol Res*. 2016;308:617–624. doi: 10.1007/s00403-016-1680-5
- Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med*. 2019;25:65–69. doi: 10.1038/s41591-018-0268-3
- Brasier N, Raichle CJ, Dörr M, Becke A, Nohturfft V, Weber S, Bulacher F, Salomon L, Noah T, Birkemeyer R, et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). *Europace*. 2019;21:41–47. doi: 10.1093/europace/euy176
- McManus DD, Lee J, Maitas O, Esa N, Pidikiti R, Carlucci A, Harrington J, Mick E, Chon KH. A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. *Heart Rhythm*. 2013;10:315–319. doi: 10.1016/j.hrthm.2012.12.001
- Chan P, Wong C, Poh YC, Pun L, Leung WW, Wong Y, Wong MM, Poh M, Chu DW, Siu C. Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting. *J Am Heart Assoc*. 2016;5:e003428.
- Krivoshei L, Weber S, Burkard T, Maseli A, Brasier N, Kühne M, Conen D, Huebner T, Seeck A, Eckstein J. Smart detection of atrial fibrillation. *Europace*. 2017;19:753–757.
- Koenig N, Seeck A, Eckstein J, Mainka A, Huebner T, Voss A, Weber S. Validation of a new heart rate measurement algorithm for fingertip recording of video signals with smartphones. *Telemed J E Health*. 2016;22:631–636. doi: 10.1089/tmj.2015.0212
- Dörr M, Nohturfft V, Brasier N, Bosshard E, Djurdjevic A, Gross S, Raichle CJ, Rhinisperger M, Stöckli R, Eckstein J. The WATCH AF trial: SmartWATCHes for detection of atrial fibrillation. *JACC: Clin Electrophysiol*. 2019;5:199–208. doi: 10.1016/j.jacep.2018.10.006
- Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, Lindsay BD, Wazni OM, Tarakji KG. Smartwatch algorithm for automated detection of atrial fibrillation. *J Am Coll Cardiol*. 2018;71:2381–2388. doi: 10.1016/j.jacc.2018.03.003
- Nemati S, Ghassemi MM, Ambai V, Isakadze N, Levantsevych O, Shah A, Clifford GD. Monitoring and detecting atrial fibrillation using wearable technology. *Annu Int Conf IEEE Eng Med Biol Soc*. 2016;2016:3394–3397. doi: 10.1109/EMBC.2016.7591456
- Avram R, Ramsis M, Cristal AD, Nathan V, Zhu L, Kim J, Kuang J, Gao A, Vittinghoff E, Rohdin-Bibby L, et al. Validation of an algorithm for continuous monitoring of atrial fibrillation using a consumer smartwatch. *Heart Rhythm*. 2021;18:1482–1490. doi: 10.1016/j.hrthm.2021.03.044

20. Lubitz SA, Faranesh AZ, Selvaggi C, Atlas SJ, McManus DD, Singer DE, Pagoto S, McConnell MV, Pantelopoulos A, Foulkes AS. Detection of atrial fibrillation in a large population using wearable devices: the Fitbit heart study. *Circulation*. 2022;146:1415–1424. doi: [10.1161/CIRCULATIONAHA.122.060291](https://doi.org/10.1161/CIRCULATIONAHA.122.060291)
21. Wasserlauf J, Vogel K, Helm RH, Steinhaus D, Passman RS. Accuracy of the Apple watch for detection of atrial fibrillation: a multi-center experience. *Cardiovasc Digital Health J*. 2022;3:S23. doi: [10.1016/j.cvdhj.2022.07.053](https://doi.org/10.1016/j.cvdhj.2022.07.053)
22. 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;27:1359–1363. doi: [10.1093/jamia/ocaa137](https://doi.org/10.1093/jamia/ocaa137)
23. Svendsen JH, Diederichsen SZ, Højberg S, Krieger DW, Graff C, Kronborg C, Olesen MS, Nielsen JB, Holst AG, Brandes A, et al. Implantable loop recorder detection of atrial fibrillation to prevent stroke (the LOOP study): a randomised controlled trial. *Lancet*. 2021;398:1507–1516. doi: [10.1016/S0140-6736\(21\)01698-6](https://doi.org/10.1016/S0140-6736(21)01698-6)
24. Wineinger NE, Barrett PM, Zhang Y, Irfanullah I, Muse ED, Steinhubl SR, Topol EJ. Identification of paroxysmal atrial fibrillation subtypes in over 13,000 individuals. *Heart Rhythm*. 2019;16:26–30. doi: [10.1016/j.hrthm.2018.08.012](https://doi.org/10.1016/j.hrthm.2018.08.012)
25. Marcus GM, Olgin JE, Whooley M, Vittinghoff E, Stone KL, Mehra R, Hulley SB, Schiller NB. Racial differences in atrial fibrillation prevalence and left atrial size. *Am J Med*. 2010;123:375.e1–375.e7. doi: [10.1016/j.amjmed.2009.05.019](https://doi.org/10.1016/j.amjmed.2009.05.019)
26. Galarnyk M, Quer G, McLaughlin K, Ariniello L, Steinhubl SR. Usability of a wrist-worn smartwatch in a direct-to-participant randomized pragmatic clinical trial. *Digit Biomark*. 2019;3:176–184. doi: [10.1159/000504838](https://doi.org/10.1159/000504838)
27. Chang P-C, Wen M-S, Chou C-C, Wang C-C, Hung K-C. Atrial fibrillation detection using ambulatory smartwatch photoplethysmography and validation with simultaneous holter recording. *Am Heart J*. 2022;247:55–62. doi: [10.1016/j.ahj.2022.02.002](https://doi.org/10.1016/j.ahj.2022.02.002)
28. Bonomi AG, Schipper F, Eerikäinen LM, Margarito J, van Dinther R, Muesch G, de Morree HM, Aarts RM, Babaeizadeh S, McManus DD, et al. Atrial fibrillation detection using a novel cardiac ambulatory monitor based on photo-plethysmography at the wrist. *J Am Heart Assoc*. 2018;7:e009351. doi: [10.1161/JAHA.118.009351](https://doi.org/10.1161/JAHA.118.009351)
29. Faranesh AZ, Pantelopoulos A, Milesco A, Heneghan C. Abstract 14160: assessment of atrial fibrillation burden by smartwatch photoplethysmography. *Circulation*. 2019;140:A14160.
30. Samuel M, Khairy P, Champagne J, Deyell MW, Macle L, Leong-Sit P, Novak P, Badra-Verdu M, Sapp J, Tardif J-C, et al. Association of atrial fibrillation burden with health-related quality of life after atrial fibrillation ablation: substudy of the cryoballoon vs contact-force atrial fibrillation ablation (CIRCA-DOSE) randomized clinical trial. *JAMA Cardiol*. 2021;6:1324–1328. doi: [10.1001/jamacardio.2021.3063](https://doi.org/10.1001/jamacardio.2021.3063)
31. Mantovan R, Macle L, De Martino G, Chen J, Morillo CA, Novak P, Calzolari V, Khaykin Y, Guerra PG, Nair G, et al. Relationship of quality of life with procedural success of atrial fibrillation (AF) ablation and postablation AF burden: substudy of the STAR AF randomized trial. *Can J Cardiol*. 2013;29:1211–1217. doi: [10.1016/j.cjca.2013.06.006](https://doi.org/10.1016/j.cjca.2013.06.006)
32. Marrouche NF, Brachmann J, Andresen D, Siebels J, Boersma L, Jordaens L, Merkely B, Pokushalov E, Sanders P, Proff J, et al. Catheter ablation for atrial fibrillation with heart failure. *N Engl J Med*. 2018;378:417–427. doi: [10.1056/NEJMoa1707855](https://doi.org/10.1056/NEJMoa1707855)
33. Go AS, Reynolds K, Yang J, Gupta N, Lenane J, Sung SH, Harrison TN, Liu TI, Solomon MD. Association of Burden of atrial fibrillation with risk of ischemic stroke in adults with paroxysmal atrial fibrillation: the KP-RHYTHM study. *JAMA Cardiol*. 2018;3:601–608. doi: [10.1001/jamacardio.2018.1176](https://doi.org/10.1001/jamacardio.2018.1176)