STATISTICAL ANALYSIS PLAN

Protocol Number: 013

(099-20494)

ECG APP 2.0 CLINICAL VALIDATION STUDY

Apple Inc.
One Apple Park Way
Cupertino, CA 95014
United States

[redacted]

Initial Version Date: February 24, 2020
Current Version Number: [redacted]
Current Version Date: April 6, 2020

TABLE OF CONTENTS

Statistical Analysis Plan Approval Signature Page

The undersigned have reviewed and approve the Statistical Analysis Plan.

SIGNATURES

Sponsor Signatory	**Signature**	**Date**
Name/Role	_____	_____
[redacted]		
Name/Role	_____	_____
[redacted]		
Name/Role	_____	_____
[redacted]		
Name/Role	_____	_____
Name/Title	_____	_____

All approvals are maintained in Agile.

LIST OF ABBREVIATIONS

AE	Adverse Event
AF	Atrial Fibrillation
AV	Atrioventricular
AVNRT	Atrioventricular Nodal Reentry Tachycardia
AVRT	Atrioventricular Reentrant Tachycardia
BMI	Body Mass Index
bpm	Beats Per Minute
CAS	Classifiable Analysis Set
cm	Centimeters
CRO	Clinical Research Organization
CSR	Clinical Study Report
DEN	De Novo
ECG	Electrocardiogram
FAS	Full Analysis Set
FDA	Food and Drug Administration
HR	Heart Rate
I/E	Inclusion/Exclusion
IRB	Institutional Review Board
kg	Kilogram
LBBB	Left Branch Bundle Blockage
m	Meter
mm	Millimeters
mmHg	Millimeters of Mercury
msec	Millisecond(s)
PAF	Paroxysmal Atrial Fibrillation
QTc	Individual-based corrected QT interval
QTcB	Bazett's corrected QT interval
QTcF	Fridericia's corrected QT interval
R	R software
RBBB	Right Branch Bundle Blockage
SAE	Serious Adverse Event
SAS	Statistical Analysis Software
SR	Sinus Rhythm
Std. Dev.	Standard Deviation
SVT	Supraventricular Tachycardia
WAAS	Waveform Assessment Analysis Set
WHO	World Health Organization

1. Version History

[redacted]

2. Introduction

This statistical analysis plan describes the analysis of the ECG App Clinical Validation Study data.

2.1. Background

Atrial fibrillation (AF) is the most common cardiac arrhythmia, and when left untreated, is a leading cause of morbidity and mortality from stroke, heart failure and myocardial infarction[1,2]. Data from the Framingham Heart Study indicates that by age 40 years, lifetime risk for developing AF is 1 in 4[3]. AF is also a growing public health problem with prevalence projected to triple between 2010 and 2050, with an estimated 12.1 million diagnosed cases in 2030 in the United States (US) alone[4].

Early detection and treatment of patients with AF minimizes the risk of sequelae of thromboembolism including >60% reduced risk of stroke[2,5]. However, atrial fibrillation is commonly underdiagnosed; many people with AF are asymptomatic or experience mild nonspecific symptoms for which they do not seek medical attention or screening[2]. As a result, asymptomatic patients are 3 times as likely to have sustained an ischemic stroke prior to diagnosis than those with symptoms[1,6]. These findings raise concerns and have prompted several variations of screening programs to detect patients with asymptomatic AF to prevent an embolic event[1,2]. While systematic and opportunistic screening programs have demonstrated increased rates of detection when compared to detection during routine clinical practice, such screening programs are not yet widely implemented[2]. Additionally, AF may be paroxysmal (PAF or intermittent AF) and therefore missed by recording a single in-clinic electrocardiogram (ECG). This is especially true for those patients with intermittent symptoms. Holter devices are commonly used for ambulatory 24-hour ECG monitoring in at-risk patients, but have limited sensitivity for the detection of new AF[7]. Given that mobile ECG devices permit an on-demand assessment by a user, our hypothesis is that use of the ECG app on Apple Watch can facilitate identification of AF.

2.2. Device Description

The ECG app software comprises a pair of mobile medical apps—one on Apple Watch and the other on the iPhone.

The ECG Watch app analyzes data collected by the integrated electrical sensors on a compatible Apple Watch to generate an ECG waveform similar to a Lead I, calculate average heart rate, and provide a rhythm classification to the user for a given 30 second session. When a user opens the ECG Watch app while wearing the Watch on one wrist and

places the finger of the opposite hand on the digital crown, they are completing the circuit across the heart which begins a recording session.

Once the recording session is complete, the ECG Watch app performs signal processing, feature extraction and rhythm classification to generate a session result.

2.3. Study Rationale

This clinical study is being conducted to support investigational ECG app 2.0 algorithms (test device) for the Apple ECG app which expand the classification heart range, introduce new classification results, and introduce minor, non-user-facing algorithm updates, and evaluate the performance of the test device.

3. Study Objective

The objective of this study is to evaluate the performance of the Test Device.

4. Study Design Overview

This is a prospective, non-significant risk study. The study protocol will undergo review and approval by an Institutional Review Board (IRB) prior to enrolling study subjects.

Written, informed consent will be obtained from all subjects before any protocol-directed procedures are performed. Potential subjects will participate in a screening visit, and if eligibility is confirmed, subjects may begin study participation at the same visit.

The investigational ECG app 2.0 algorithms (test device) will be run post-hoc on the sensor dataset to generate ECG app waveforms and classifications. The performance of the test device will be determined based upon adjudication of concurrent data from the reference device.

5. Study Procedures

5.1. Screening

Informed consent will be obtained before any study protocol-directed procedures are performed. After the signing of informed consent, the subject will be evaluated for eligibility according to the study inclusion/exclusion criteria.

During screening, the following should be obtained on each subject:

- Age
- Gender
- Race and ethnicity

- Diagnosed medical conditions such as: past surgical history, illness, previous diagnoses
- Prior and Concomitant Medications
- Allergies
- Tobacco and Nicotine history
- Alcohol use (self-reported)
- Recreational drug use (self-reported)
- Physical examination
- Height and weight
- Body mass index (BMI)
- Vital signs will be collected after sitting for at least 5 minutes
- Urine pregnancy testing for women of childbearing potential*
- Caffeine intake
- Exercise habits
- Dominant hand
- Wrist circumference of wrist wearing the Watch
- Skin fold thickness of wrist wearing the Watch
- Wrist hairiness of wrist wearing the Watch
- Assessment of tattoos, moles, scars on wrist of wrist wearing the Watch
- 12-lead ECG will be collected supine after lying for at least 5 minutes

* For women of child-bearing potential only. A woman will be considered not of child-bearing potential if they are surgically sterile and have provided documentation, or they are ≥55 and have not had a cycle for ≥2 years.

All subjects will undergo a 12-lead ECG. The 12-lead ECG will be reviewed and interpreted by the investigator. If subjects were not previously aware of any abnormal findings on ECG, the investigator will inform the subject of the findings; any further medical care will be at the discretion of the subject's non-study medical provider. Non-AF subjects must be in normal sinus rhythm at the time of the 12-lead ECG screening in order to be assigned to Cohort 1. AF subjects must be in AF at the time of the 12-lead ECG at screening in order to be assigned to Cohort 2. AF subjects that do not present AF at time of screening may come back another day to determine their eligibility for Cohort 2 at a maximum of one repeat time.

If any of the study staff believe the vital signs or screening 12-lead ECG are of poor quality, they may repeat the measurements.

5.2. Study Day 1

Subjects will report to the clinic for the Day 1 visit at their scheduled date and time. No special preparation for this visit is required. If eligibility is confirmed after completion of all screening procedures, subjects may begin study participation following the screening.

Subjects will be asked to remove jewelry and any underwire bras during 12-lead ECG data collection. NOTE: Non-underwire bras are acceptable during data collection.

5.2.1. Study Procedures

5.2.1.1. Data Collection Study Equipment Set up

The following procedures will be performed for each study subject at study participation:

Assign an Apple Watch and the paired iPhone to each subject and document the Apple Watch and paired iPhone numbers and the version of Study App. The Watch configurations were selected at random from the total set of Watch configuration. The random assignment of watches to subjects is based on a pre-determined randomization schedule stratified by site, age group, and cohort.

1. Instruct the subject to choose his/her preferred wrist to wear the Apple Watch. The Apple Watch can be worn on whichever side the subject prefers unless there is a skin condition on one wrist as noted above in the exclusion criteria, in which case, the Apple Watch must be worn on the non-affected wrist. Record the chosen wrist. If the subject does not have a side preference or skin condition on one wrist, the device will be placed on the left wrist.

2. The subject will be asked to put the Apple Watch on his/her wrist (as indicated in 1). The Apple Watch should be fitted tightly enough that it does not move when the hand/wrist is shaken.

3. Check band tightness for a snug fit and adjust as needed.

4. Instruct the subject to read the provided instructions on how to Take an ECG. Instructions will come from the "How to Take an ECG" onboarding screen of the ECG app (DEN180044) on iPhone.

5. Ask the subject to take a practice ECG. After the first practice run, site study staff will guide subject to appropriate study posture and grip (if not achieved independently during the first test run).

6. Allow subject up to three additional practice runs if more practice runs are needed to ensure study posture and grip instruction understanding (fewer than 3 additional practice runs are acceptable if study posture and grip are correct and patient is comfortable with the positioning).

5.2.1.2. Data Collection – Rest ECG

1. Instruct subject to remain seated for a 5-minute resting period prior to collecting a 12-lead ECG and set the reference system to 40 Hz filter. Use standard placement of all 12-lead electrodes (with limb lead for Lead 1).

2. Collect a 12-lead ECG (review to ensure there is no artifact). The recording will be approximately 30 seconds in duration. Simultaneously, instruct the subject to record a separate complete single-lead ECG using the Apple Watch. Both data streams will be synced together by the Study Staff through a flash sync method.

    a) Monitor the 12-lead ECG to ensure a good quality recording is obtained.

    b) A total of three trials consisting of simultaneously recorded 12-lead ECG and single-lead ECG using the Apple Watch will be performed for each subject.

    c) The first trial will be used for adjudication and analysis.

        - [redacted]

    d) The baseline resting heart rate for exercise (5.2.1.3) will be determined from the 1st trial.

5.2.1.3. Exercise

For the exercise sessions, a target HR will be defined as 85% of predicted max HR (PMHR = 220 – age) which is not to exceed 150 bpm; thus the maximum target HR in any subject is 150 bpm. Subjects who have been cleared and deemed fit by the Investigator to exercise using the treadmill or stationary bike will complete training on how to use the exercise equipment and perform a practice session to determine adequate resistance for increasing the HR to > target HR. There will be up to 3 total exercise trials. Each trial consists of up to 5 minutes of exercise, depending on how long it takes for the target BPM of >target HR to be achieved, followed by an approximately 1 minute data capture session where the subject is stationary to complete a data capture session. Subject should be at target HR for at least 5 seconds prior to starting data collection. At the discretion of the investigator, blood pressure reading may be taken for safety purposes. During the exercise, the subject will be encouraged by the Study Staff to increase effort if the subject's initially chosen exercise intensity is too low to raise their HR by the requisite amount. For treadmill exercise, the subject is recommended to move off the treadmill and into a seated position, then supporting their arms and elbows on their lap while completing data collection with the Apple Watch. For stationary bikes, the subject can remain on the bike and place their hands on the bike handles in a still position for data collection with the Apple Watch.

The subject may rest in between each trial for approximately 3 minutes. The subject will be monitored throughout the session. The reference ECG will be continuously monitored during the exercise session by the Investigator and/or Study Staff.

The first trial will be used for adjudication and analysis. [redacted]

5.2.2. Apple Watch Data Processing

Apple Watches will be collected, data will be transferred from the Watches to the paired iPhones by CRO staff, and data will be transferred to study sponsor.

Procedures for ECG waveform generation and ECG app classification based on the collected Apple Watch sensor data are described in Section 9.1.

5.2.3. Adjudication of ECG Data

5.2.3.1. 12-lead ECG

Thirty second 12-lead ECGs will be reviewed by 2 independent US Board Certified Cardiologist adjudicators for heart rate and rhythm. In the event of discrepancy, ECGs will be reviewed by a third adjudicator. All adjudicators will be blinded to cohort, stage (resting or exercise) and ECG app classification. Adjudicators will be instructed to review ECGs independently and separately of each other and not to confer about diagnoses. The 12-lead ECGs generated from the first trial at rest and the first trial after exercise will be adjudicated for each subject.

For purposes of the primary endpoint analysis, the rhythm classification and the heart rate will be adjudicated separately. The final adjudicated rhythm classification result will be used if there is a rhythm classification discrepancy between the first two cardiologists.

5.2.3.1.1. Heart Rate

Heart rate will be calculated for each 12-lead ECG. Record the heart rate and select the heart rate diagnostic code that corresponds to the heart rate observed on the reference ECG. The final adjudicated heart rate result will be used if there is a heart rate discrepancy between the first two cardiologists.

5.2.3.1.2. Rhythm Diagnoses

The following rhythm diagnoses will be adjudicated to the 12-lead ECG data for each 12-lead ECG:

1. Sinus Rhythm (will include sinus bradycardia, normal sinus rhythm, and sinus tachycardia)
2. Atrial Fibrillation
3. Supraventricular tachycardia (SVT) with regular beat-to-beat intervals (e.g., AVNRT, AVRT, Atrial Tachycardia with HR over 100 beats per minute)
4. Other Abnormal Rhythm
    a. Frequent Premature Atrial Contractions (>3 in 30 seconds)
    b. Frequent Premature Ventricular Contractions (>3 in 30 seconds)
    c. Atrial flutter
    d. Ventricular tachycardia
    e. Ventricular fibrillation
    f. Second degree AV block, Type I
    g. Second degree AV block, Type II
    h. Third degree AV block
    i. Other
5. Uninterpretable

5.2.3.1.3. Combining Heart Rate and Rhythm Diagnoses on 12-lead ECG

Table 5.1: Heart Rate and Rhythm Diagnoses on 12-Lead ECG

Heart Rate	**Rhythm on 12-lead ECG**				
	Sinus Rhythm	Atrial Fibrillation	SVT with regular intervals	Other	Uninterpretable
<50	HR<50	HR<50	N/A	HR<50	Uninterpretable
50-99	SR 50-99	AF 50-99	N/A	Other 50-99	Uninterpretable
100-150	SR 100-150	AF 100-150	SVT 100-150	Other 100-150	Uninterpretable
>150	HR>150	HR>150	HR>150	HR>150	Uninterpretable

5.2.3.2. ECG Strip Overlay

1. Strip preparation

    a. The Sponsor will prepare one paired set of at rest ECG strips from approximately 100 randomly selected subjects (50 AF; 50 SR) comprised of the Apple Watch ECG and lead 1 ECG from the 12-lead.

2. Cardiac Core Lab review:

    a. Three blinded, independent cardiac technicians will review each paired set of strips on 25mm/s, 10 mm/mV standard.

    b. One reviewer will identify the first six consecutive distinct readable PQRST complexes without artifact that match between the subject device strip and reference device strip for evaluation. For example, if the reference strip has an artifact in beats 2 and 5 but all other beats are good and the strip from the test device has artifact in beats 1 and 6 but all other beats are good, the six consecutive beats to be used will begin at beat 7 of both strips. The strips will be excluded if six consecutive beats cannot be found. The six PQRST complexes identified by the initial reviewer will be used by the 2 other reviewers.

    c. The strips will be rendered semi-transparent through the use of back lighting and overlaid to visually assess for morphological similarity.

    d. Each reviewer will assign a pass/fail to the strips by visually assessing all 6 PQRST complexes. A "pass" is given when the morphology of the PQRST complexes appears to overlay to the unaided eye.

    e. Measure R amplitude from the isoelectric baseline of the PR segment at the onset of the QRS complex to the highest vertical deflection of the R-wave. Measure to the nearest 0.5 millimeter for the first QRS complex of the six consecutive distinct readable PQRST complexes in both the reference strip and test device strip. The reviewers will be blinded to the identity of the reference strip and test device strip.

6. Study Endpoints

6.1. Primary Study Endpoints

Co-Primary Endpoints:

1a. Correct classification of subjects with normal sinus rhythm (HR 50-150) on simultaneous 12-lead ECG as "Sinus Rhythm" or "High Heart rate" on a readable and classifiable ECG app strip. [Specificity]

1b. Correct classification of subjects with AF (HR 50-150) on simultaneous 12-lead ECG as "AF" or AF (high heart rate) on a readable and classifiable ECG app strip. [Sensitivity]

6.2. Secondary Endpoints

2a. Correct classification of subjects with normal sinus rhythm (HR 50-99, sinus rhythm (SR) on simultaneous 12-lead ECG) as "Sinus Rhythm" on a readable and classifiable ECG app strip

2b. Correct classification of subjects with AF (HR 50-99, AF on simultaneous 12-lead ECG) as "AF" on a readable and classifiable ECG app strip

2c. Correct classification of subjects with sinus tachycardia (HR 100-150, SR on simultaneous 12-lead ECG) as "High Heart Rate" on a readable and classifiable ECG app strip

2d. Correct classification of subjects with AF with high heart rate (HR 100-150, AF on simultaneous 12-lead ECG) as "AF (high heart rate)" on a readable and classifiable ECG app strip

2e. Equivalence of the ECG app waveform to Lead I from a 12-lead ECG as measured by acceptable morphology of PQRST complexes

2f. Equivalence of the ECG app waveform to Lead I from a 12-lead ECG as measured by R-Wave amplitude agreement

6.3. Study Hypotheses

[redacted]

[redacted]

7. Statistical Considerations

7.1. Sample Size

[redacted] The sample sizes to achieve 80% power with a one-sided type I error of 0.025 using an exact binomial test are [redacted] with no known diagnosis of AF and [redacted] with a known diagnosis of AF.

Table 7.1 below presents the expected percent correct associated with the four secondary endpoints 2a, 2b, 2c, and 2d and are summarized below along with the proposed performance goals.

Table 7.1: Sample size determination for four secondary endpoints

[redacted]

[redacted] To account for obtaining readable/classifiable waveforms and to ensure enough data is collected from the exercise stage, 168 subjects with no known diagnosis of AF and 400 subjects with a known diagnosis of AF will be enrolled.

[redacted] To account for obtaining readable waveforms and to assess waveform equivalence across a

range of heart rates and rhythms, approximately 100 subjects (50 SR; 50 AF) will be randomly selected according to the sampling scheme presented in Section 7.2.

7.2. Randomization and Blinding

Subjects in this study will not be randomized to any treatment regimens.

Approximately 100 subjects (50 SR; 50 AF) will be randomly selected to assess waveform equivalence of the ECG app to Lead 1 from a 12-Lead ECG. Subjects will be randomly selected using a systematic sampling approach. A number between 1 and 6 will be randomly selected for the AF cohort and then every 6th consecutively enrolled AF subject will be selected at random within each site. Likewise, a number between 1 and 5 will be randomly selected for the SR cohort and then every 5th consecutively enrolled SR subject will be selected at random within each site.

US board-certified cardiologists will be blinded to the subjects' past medical history, stage (resting or exercise) and ECG app classifications during the assessment and adjudication of ECG data.

Cardiac technicians will be blinded to the subjects' past medical history, ECG app classifications, stage (resting or exercise) and origin (ECG app or 12-lead ECG) of the waveform being reviewed during the assessment and adjudication of ECG data.

Adjudicators will be instructed to review ECGs independently and separately of each other and not to confer about diagnoses.

7.3. Significance Level

The primary hypothesis, the four secondary endpoint hypothesis tests of percent correct, and the two secondary endpoint hypotheses of waveform assessment will use a one-sided significance level of 0.025.

7.4. Missing Data

Rigorous efforts will be made to ensure all subjects are compliant with the protocol. However, some subjects may drop out prematurely or some planned measurements may not be analyzable due to missing data or uninterpretable results from either the ECG adjudication process or the ECG app test device. The data analyses will be conducted on all analyzable data.

7.5. Subgroup Analyses

In accordance with the FDA Guidance[8] on the reporting of age-, race-, and ethnicity-specific data, the primary endpoints will be reported by the following subgroups (data permitting). Some subgroups may be combined depending on data availability.

- Age group (<55; ≥55 to <65; ≥65 years)

- Sex (Male; Female)
- Race (White; Black or African American; Asian; American Indian or Alaska Native; Native Hawaiian or other Pacific Islander)
- Ethnicity (Hispanic or Latino; Not Hispanic or Latino)
- Watch series/material types (Series 4, Series 5: Aluminum, ceramic, titanium, stainless steel)

7.6. Interim Analyses

There are no interim analyses planned in this study.

8. Analysis Sets

The following analysis sets are defined for this study.

Full Analysis Set (FAS): All subjects who sign informed consent and are enrolled into the study. This analysis set will be used to summarize subject 12-Lead, and ECG app device accountability in addition to demographic and baseline characteristics and safety data. The data for this analysis set will be presented overall and separately for the enrollment cohorts.

Classifiable Analysis Set (CAS): All subjects who have readable paired ECG app and 12-Lead ECG adjudicated results. This analysis set will be used for analyzing the co-primary endpoints (1a, 1b), the four secondary endpoints (2a, 2b, 2c, and 2d) and the additional analyses.

Waveform Assessment Analysis Set (WAAS): Randomly selected subjects with readable paired strips from the ECG app and Lead 1 of the 12-Lead ECG. If 6 consecutive beats for analysis cannot be found in these strips, they will be excluded from further analysis. This analysis set will be used for assessing the clinical waveform accuracy associated with secondary endpoints 2e and 2f.

9. Analysis Approach

9.1. Dataset Generation

A single sensor dataset to generate ECG app waveforms and classifications will be used for this study to which the test device algorithm will be applied in a post-hoc fashion.

9.2. Subject Accountability

A summary of the number of enrolled subjects, eligible subjects, withdrawn subjects, and completers will be presented by cohort assigned at baseline (SR or AF), overall, and by investigational site. Reasons for withdrawal will be summarized according to the following categories:

- Adverse Event

- Death
- Protocol Deviation
- Withdrawal of Consent
- Investigator Discretion
- Termination of the Study
- Device malfunction
- Lost to Follow-up
- Other

9.3. ECG App and 12-Lead ECG Measurement Accountability

A summary table will be reported which presents the accountability of ECG app and 12-Lead ECG test measurement results overall and by enrollment cohort.

9.4. Demographic Characteristics

Descriptive statistics (e.g., N, Mean, Std. Dev., Min, Max) for continuous data types and frequencies for categorical data types will be displayed for the demographic characteristics by cohort assigned at baseline (SR or AF), overall, and by investigational site. The demographic characteristics and data types are listed below:

Table 9.1: Demographic Characteristics

Characteristic	Data Type
Age at Enrollment (years)	Continuous
Age Group at Enrollment	Categorical
Sex	Categorical
Ethnicity	Categorical
Race	Categorical

The following age group categories will be used: <55, ≥55 to <65, and ≥65 years. Subjects may choose more than one Race category.

9.5. Baseline Characteristics

The following baseline characteristics will be summarized using descriptive statistics by enrollment cohort assigned at baseline (SR or AF) and overall.

9.5.1. History of Heart Rhythm Abnormalities

Heart rhythm abnormalities will be summarized with frequencies and percentages according to the following categories:

- None
- AF (permanent)
- AF (persistent)
- AF (paroxysmal)
- AF (other)

- Atrial flutter
- Ventricular Tachycardia
- Atrial Tachycardia
- 1st degree AV block
- 2nd degree AV block (Type I)
- 2nd degree AV block (Type II)
- 2nd degree AV block (other)
- Left Branch Bundle Blockage (LBBB)
- Right Branch Bundle Blockage (RBBB)
- Other

9.5.2. Baseline Electrocardiogram Results

The baseline ECG results at screening will be summarized descriptively as follows:

Table 9.2: ECG Results at Screening

Characteristic	Data Type
Ventricular rate (bpm)	Continuous
PR interval (msec)	Continuous
QRS duration (msec)	Continuous
QT interval (msec)	Continuous
RR interval (msec)	Continuous
QTc (msec)	Continuous
QTcB (msec)	Continuous
QTcF (msec)	Continuous
Rhythm	Categorical
Results interpretation (normal/abnormal)	Categorical

Abnormal ECG findings will be further summarized by clinical significance (yes/no).

9.5.3. Lifestyle Characteristics

Responses to the lifestyle questions will be summarized descriptively as follows:

Table 9.3: Lifestyle Characteristics

Characteristic	Data Type
Tobacco/nicotine use	Categorical
Alcohol use	Categorical
Recreational drug use	Categorical
Caffeine use	Categorical
Exercise	Categorical

9.5.4. Physical Exam Results

Results of the physical examination will be summarized as normal/abnormal for the following anatomical categories:

- General appearance
- Head, eyes, ears, and nose
- Chest
- Heart
- Abdomen
- Musculoskeletal
- Neurologic
- Dermatologic

Tremors will also be summarized as present/absent along with descriptive statistics of tremor severity on a scale of 0-10 for those with tremors present. The number of subjects who have tattoos on sensor locations will also be summarized. Pregnancy test results will also be summarized.

9.5.5. Vital Signs

Results of the vital signs assessments will be summarized descriptively for the following parameters:

- Systolic blood pressure (mmHg)
- Diastolic blood pressure (mmHg)
- Heart rate (bpm)
- Body temperature (ºC)

9.5.6. Medical History

Medical and allergy history information will be coded using MedDRA v22.1 and summarized by preferred term.

9.5.7. Body and Device Wear Measurements

Body and device wear measurement results at baseline will be summarized descriptively as follows:

Table 9.4: Body and Device Wear Measurements

Characteristic	Data Type
Height (cm)	Continuous
Weight (kg)	Continuous
BMI (kg/ m^2)	Continuous
Dominant hand	Categorical
Preferred wrist to wear a band on	Categorical
Device wearing wrist circumference (mm)	Continuous
Device wearing wrist skinfold thickness (mm)	Continuous
Device wearing wrist hairiness	Categorical
Wrist device size	Categorical
Band size	Categorical
Device wearing wrist	Categorical

Characteristic	Data Type
Wrist device location	Categorical
Wrist device orientation	Categorical
Wrist device interaction	Categorical

9.6. Primary Endpoint Analyses

The following truth table will be generated to cross-classify data collected from enrolled subjects according to the ECG app device and adjudicated ECG data from the 12-lead ECG for the primary endpoint analyses as presented below. The rhythm classification and the heart rate will be adjudicated separately. The final adjudicated rhythm classification result will be used if there is a rhythm classification discrepancy (including the specify type of Other Abnormal Rhythm) between the first two cardiologists. The final adjudicated heart rate result will be used if there is a heart rate discrepancy between the first two cardiologists.

Table 9.5: Cross-classification truth table for primary endpoint analyses

Algorithm Classification	**Reference Strip Final Adjudicated Result (via 3 US Board Certified Cardiologists)**			
	SR ($50 \leq$ HR ≤ 150)	AF ($50 \leq$ HR ≤ 150)	Other (SVT or Other with HR 50-150; HR<50; HR>150)	Uninterpretable
SR (SR (50-100) + High HR no AF (101-150))	**n11**	**n12**	n13	n14
AF ($50 \leq$ HR ≤ 150)	**n21**	**n22**	n23	n24
Inconclusive (including HR< 50; HR>150)	n31	n32	n33	n34
Poor Recording	n41	n42	n43	n44

Note: Bold indicates data values used for the primary endpoint analyses.

The following primary endpoint hypothesis will be tested using a one-sided type I error of 0.025:

[redacted]

[redacted]

Using the cross-classified data from Table 9.5, the sensitivity and specificity for the primary endpoint hypothesis will be estimated as follows:

Sensitivity estimate = n22/(n12 + n22)

Specificity estimate = n11/(n11 + n21)

Each subject may contribute up to 2 paired ECG app and Adjudicated 12-Lead ECG results for analysis; one at rest (first trial run) and the other after exercise (first trial run).

Because data will be collected from the same subjects at rest and after exercise, a bootstrap approach will be implemented to obtain two-sided 95% confidence intervals for the sensitivity and specificity to account for potential within-subject correlation. Subjects with at least one adjudicated result of AF (for sensitivity), SR (for specificity) and with a classifiable algorithm result (i.e., SR or AF) will be selected at random with replacement and the 2.5th and 97.5th percentiles of the distribution of bootstrap estimates will represent the two-sided 95% confidence bounds. If the lower confidence bounds for both sensitivity and specificity exceed the pre-established performance goal associated with these performance metrics, the null hypothesis will be rejected in favor of the alternative hypothesis. Sensitivity and specificity estimates will also be presented separately for at rest and after exercise measurements along with their corresponding Fisher's Exact 95% confidence intervals.

As an additional analysis associated with the primary endpoint, the sensitivity and specificity will also be estimated and reported along with their corresponding bootstrap two-sided 95% confidence intervals by including the Inconclusive algorithm classification category in the calculation. More specifically, the sensitivity and specificity for the additional analysis will be estimated as follows:

Sensitivity estimate (additional analysis) = n22/(n12 + n22+n32)

Specificity estimate (additional analysis) = n11/(n11 + n21+n31)

9.7. Secondary Endpoint Analyses

The following truth table will be generated to cross-classify data collected from enrolled subjects according to the ECG app device and adjudicated ECG data from the 12-lead ECG for secondary endpoint analyses as presented below:

Table 9.6: Cross-classification truth table for secondary endpoint analyses

	Ground Truth via Adjudicated 12-lead ECG									
ECG app Device	HR	*<50*	**50-99**	**50-99**	**100-150**	**100-150**	*>150*	50-150		

Output / Rhythm	Any	**SR**	**AF**	**SR**	**AF**	Any	SVT or Other	Uninterpretable	Total
Low HR (<50)	n11	n12	n13	n14	n15	n16	n17	n18	N1D
SR (50-99)	n21	**n22**	**n23**	**n24**	**n25**	n26	n27	n28	N2D
AF (50-99)	n31	**n32**	**n33**	**n34**	**n35**	n36	n37	n38	N3D
High HR (100-150) with no AF	n41	**n42**	**n43**	**n44**	**n45**	n46	n47	n48	N4D
AF (100-150)	n51	**n52**	**n53**	**n54**	**n55**	n56	n57	n58	N5D
HR >150	n61	n62	n63	n64	n65	n66	n67	n68	N6D
Inconclusive	n71	n72	n73	n74	n75	n76	n77	n78	N7D
Poor Recording	n81	n82	n83	n84	n85	n86	n87	n88	N8D
Total	N1	N2	N3	N4	N5	N6	N7	N8	N

Note: Bold indicates data values used for the secondary endpoint analyses.

The following four secondary endpoint hypotheses will be tested:

[redacted]

[redacted]

Using the cross-classified data from Table 9.6, the percent correct for each of the four secondary endpoint hypotheses will be estimated as follows:

Hypothesis 2a: Percent Correct = n22/(n22+n32+n42+n52)

Hypothesis 2b: Percent Correct = n33/(n23+n33+n43+n53)

Hypothesis 2c: Percent Correct = n44/(n24+n34+n44+n54)

Hypothesis 2d: Percent Correct = n55/(n25+n35+n45+n55)

Each subject may contribute up to 2 paired ECG app and Adjudicated 12-Lead ECG results for analysis; one will be at rest (first trial) and the other after exercise (first trial).

Because data will be collected from the same subjects at rest and after exercise, a bootstrap approach will be implemented to obtain two-sided 95% confidence intervals for the percent correct associated with each of the four secondary endpoint hypotheses to account for potential within-subject correlation. Subjects with at least one adjudicated result of SR (HR 50-99) [hypothesis 2a], AF (HR 50-99) [hypothesis 2b], SR (HR 100-150) [hypothesis 2c], AF (HR 100-150) [hypothesis 2d] and with a classifiable algorithm result will be selected at random with replacement and the 2.5th and 97.5th percentiles of the distribution of bootstrap estimates will represent the two-sided 95% confidence bounds. If the lower confidence bound of the percent correct equals or exceeds the pre-established performance goal associated with the associated secondary endpoint hypotheses, the null hypotheses will be rejected in favor of the alternative hypotheses.

Additional secondary endpoint analyses will be performed similarly by including the HR<50, HR>150, and Inconclusive algorithm results in the computations of percent correct.

Two hypotheses associated with secondary objective to demonstrate that the ECG app produces a waveform that provides clinically equivalent information to Lead 1 of the 12-Lead ECG will be tested.

[redacted] The majority result from the pass/fail readings of the 3 independent, certified cardiac technicians will be used for analysis. The hypothesis for this objective is stated as follows:

[redacted]

[redacted]

The second hypothesis will be based on the results of the quantitative analysis of the R-wave amplitudes as measured by each of 3 independent, certified cardiac technicians on the paired ECG app and Lead 1 from the 12-Lead ECG reference strips. The majority result

from the three technicians (i.e., if at least two technicians measured the difference in amplitudes between the ECG app and reference strips to be ≤2 mm then the paired strips for a given subject were deemed ≤2 mm) was used [redacted]

[redacted]

[redacted]

For each subject, the QRS complex from the 12-lead ECG with the largest R amplitude (out of the two QRS complexes with R-wave amplitude measurements) will be used for purposes of determining the majority result.

Due to the multiple observations per subject while at rest and during exercise, a bootstrap approach will be used to construct two-sided 95% confidence intervals for the Morphology Pass Rating Proportion and the R-Wave Amplitude Agreement Proportion. If the lower 2.5th percentile of the bootstrap distribution for the Morphology Pass Rating Proportion [redacted], the null hypothesis H_{0_2e} will be rejected. [redacted] to the evaluation of the R-Wave Amplitude Agreement Proportion.

9.8. Additional Analyses

The following additional analyses will be performed.

1) Percent correct associated with the Other abnormal rhythm truth category will be estimated as n77/(N7 - n87).

2) Percent correct associated with the HR under 50 (HR<50 on simultaneous 12-lead ECG) truth category will be estimated as n11/(N1-n81).

3) Percent correct associated with the HR over 150 (HR >150 on simultaneous 12-lead ECG) truth category will be estimated as n66/(N6-n86).

4) Percent correct associated with the AF rhythm classification for adjudicated heart rate values between 50-120. The numerator of this percent correct calculation will be the number of test device classifications of AF and test device heart rate between 50-120 with an adjudicated AF rhythm classification and adjudicated heart rate value between 50-120. The denominator will include all classifiable test device results with an adjudicated AF rhythm classification and adjudicated heart rate between 50-120.

Two-sided 95% bootstrap confidence intervals will be constructed during the same bootstrap sampling procedure outlined for the primary and secondary endpoints for these four additional analyses. If the number of samples within any of these categories is ≤20, only summary statistics will be presented.

5) The number and percentage of the ECG app Inconclusive and the ECG app Poor Recording trials within each category of ground truth will be reported as well as overall as a percentage of the total number of trials.

6) A cross-classification table of the ECG app results with each of the types of Other abnormal rhythm truth categories (outlined in Section 5.2.3.1.2) will be presented.

7) Descriptive summary statistics (N, mean, std. dev, min, max) of the paired heart rate differences (bpm) between the ECG app and the 12-Lead ECG will be reported by cohort and overall.

9.9. Safety Analyses

All adverse events will be recorded throughout the entire study period, whether they are considered to be related to the study procedures or not. Signs and symptoms of each AE will be described in detail: date of event, description of event, severity, relationship to study procedures, action taken and outcome. Adverse events will be collected as spontaneously reported by the subjects.

Adverse events will be coded using MedDRA v22.1. The number of any adverse events and the number and percentage of subjects reporting each type of adverse event will be presented by Preferred Term. Multiple occurrences of the same event reported by the same subject will be counted only once.

Adverse event summaries (number of events and incidence) will be presented for:
- All adverse events
- Serious adverse events (SAEs)
- Study procedure related adverse events
- Severe adverse events

Study procedure related events will include those events classified as possibly or definitely related to the study procedure.

The intensity of an AE will be categorized as follows:
- Mild: Events that are easily tolerated with no disruption of normal daily activity
- Moderate: Events that cause sufficient discomfort to interfere with daily activity
- Severe: Events that incapacitate and prevent usual activity

Causal relationship assessment to study procedures is required for purposes of reporting AEs. To promote consistency, the following guidelines should be taken into consideration along with good clinical and scientific judgment when determining the relationship of study procedures to an AE:

- Definitely Related: A clinical event, including laboratory test abnormality, occurring in a plausible temporal relationship to the study procedures, and which cannot be explained by concurrent disease or other drugs or chemicals.

- Possibly Related: A clinical event, including laboratory test abnormality, with a reasonable temporal relationship to the study procedures, but which could also be explained by concurrent disease or other drugs or chemicals.
- Unlikely Related: A clinical event, including laboratory test abnormality, with little or no temporal relationship to the study procedures, and which could be explained by concurrent disease or other drugs or chemicals.
- Not Related: A clinical event, including laboratory test abnormality, that has no temporal relationship to the study procedures

Action taken will be defined as:
- None;
- Study procedures interrupted;
- Study procedures stopped

Outcome will be defined as:
- Resolved;
- Ongoing or stabilized and followed by private MD;
- Lost to follow up

9.10. Medication Usage

Prior and concomitant medication usage will be coded using the WHO Drug Dictionary Global (version 01SEP2019) and presented separately in the listings. A prior medication is defined as a medication end date that is more than 30 days prior to the informed consent date. All other medications will be considered concomitant medications.

9.11. Protocol Deviations

Protocol deviations will be presented in the subject listings.

10. Statistical Software

All analyses will be performed with SAS (v9.4 or later) and R (v3.6.0 or later for bootstrap sampling analyses).

11. Changes to Planned Analysis

Due to the COVID-19 pandemic, study enrollment was stopped on March 11, 2020 prior to completing the planned targeted enrollment. All statistical analyses will be performed as planned using all available data. Although the sample size is lower than targeted, the number of subjects enrolled exceeds the required number of subjects to achieve 80% power for both the primary and secondary endpoint hypotheses. For purposes of the subgroup analysis of the primary endpoint by race, the following race categories will be reported: White, Black or African American, Other.

Also, the subject selection sampling scheme for waveform assessment in Section 7.2 has been updated to ensure a roughly equal balance of SR and AF cohort subjects based on the number of subjects enrolled at the time study enrollment was stopped.

12. References

[1] Ben Freedman S, Lowres N. Asymptomatic Atrial Fibrillation: The Case for Screening to Prevent Stroke. JAMA. 2015 Nov 10;314(18):1911-2. doi: 10.1001/jama.2015.9846.

[2] Moran PS1, Teljeur C, Ryan M, Smith SM. Systematic screening for the detection of atrial fibrillation. Cochrane Database Syst Rev. 2016 Jun 3;(6):CD009586. doi: 10.1002/14651858.CD009586.pub3.

[3] Lloyd-Jones DM1, Wang TJ, Leip EP, Larson MG, Levy D, Vasan RS, D'Agostino RB, Massaro JM, Beiser A, Wolf PA, Benjamin EJ. Lifetime risk for development of atrial fibrillation: the Framingham Heart Study. Circulation. 2004 Aug 31;110(9):1042-6. Epub 2004 Aug 16.

[4] Colilla S1, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. Am J Cardiol. 2013 Oct 15;112(8):1142-7. doi: 10.1016/j.amjcard.2013.05.063. Epub 2013 Jul 4.

[5] Omboni S1, Verberk WJ2. Opportunistic screening of atrial fibrillation by automatic blood pressure measurement in the community. BMJ Open. 2016 Apr 12;6(4):e010745. doi: 10.1136/bmjopen-2015-010745.

[6] O'Neal WT1, Efird JT2, Judd SE3, McClure LA4, Howard VJ5, Howard G3, Soliman EZ6,7. Impact of Awareness and Patterns of Nonhospitalized Atrial Fibrillation on the Risk of Mortality: The Reasons for Geographic And Racial Differences in Stroke (REGARDS) Study. Clin Cardiol. 2016 Feb;39(2):103-10. doi: 10.1002/clc.22501. Epub 2016 Feb 16.

[7] Brachmann J1, Morillo CA2, Sanna T2, Di Lazzaro V2, Diener HC2, Bernstein RA2, Rymer M2, Ziegler PD2, Liu S2, Passman RS2. Uncovering Atrial Fibrillation Beyond Short-Term Monitoring in Cryptogenic Stroke Patients: Three-Year Results From the Cryptogenic Stroke and Underlying Atrial Fibrillation Trial. Circ Arrhythm Electrophysiol. 2016 Jan;9(1):e003333. doi: 10.1161/CIRCEP.115.003333. presented as both frequency counts and as a percentage.

[8] Evaluation and Reporting of Age-, Race-, and Ethnicity-Specific Data in Medical Device Clinical Studies. FDA Guidance Document. September 12, 2017.

13. Appendix: Schedule of Events

Schedule of Events Matrix

Procedure	Screening Before Enrollment	Study Day*** In-Lab
Informed Consent	X	
I/E Criteria Assessment	X	
Medical History including allergies	X	
Demographic information	X	
12-lead ECG	X	X**
Vital Signs	X	
Urine Pregnancy Test*	X	
Wrist/Hand Assessment (Dominant hand, skinfold thickness, wrist circumference, wrist hairiness, tattoos)	X	
Physical Examination (will include at a minimum general appearance, lungs, cardiovascular, head and neck, abdomen, musculoskeletal/extremities, lymph nodes, skin and neurological assessment)	X	
Social History (smoking status, alcohol use, recreational drug use, caffeine intake)	X	
Body Measurements and Assessments (anthropometrics) (height, weight, BMI)	X	
Concomitant Medications	X	X**
Dispense Study Equipment, test set-up, record device ID assigned to subject		X
Record watch series, size, watch material (aluminum, steel, etc.)		X
Record subject choice of wrist for device. Default is left wrist		X
Take baseline resting heart rate		X
Practice taking ECG reading		X
Collect ECG reading with both device and 12-lead ECG - 3 trials		X
Practice exercise trial to determine resistance/speed		X

Collect ECG reading after each exercise trial with both device and 12-lead ECG - 3 trials		**X**
Detach 12-lead ECG		**X**
Remove wrist device		**X**
Transfer data from wrist device to paired smartphones		**X**
Record Adverse and/or Serious Adverse Events		**X**
Disposition Status Assessment		**X**

*For women of child-bearing potential only. A woman will be considered not of child-bearing potential if they are surgically sterile and have provided documentation, or they are ≥55 and have not had a cycle for ≥2 years.
**Only needed if study day is different from screening day
*** Study Day can be on the same day or different day than the Screening Day.