PCORNET ANALYTIC PLAN TEMPLATE

Date (mm/dd/yyyy): 05/12/2017	
Title of Project: Individualized Studies of Triggers of Paroxysmal Atrial Fibrillation (I-STOP-AF)	
NCT03323099	
Period Covered by this Report: 01/01/2017 to 04/01/2017	

0

Principal Investigator & Institution Updated Contact Information:	
PI First Name:	Gregory
PI Last Name:	Marcus
Office Phone:	415-476-3450
Institution Legal Name:	University of California, San Francisco
Address (street, city, state, zip code):	505 Parnassus Ave, M1180B, San Francisco, CA, 94143
Telephone:	415-476-5706

Study Coordinator & Institution Updated Contact Information:	
First Name:	Madelaine
Last Name:	Faulkner
Office Phone:	415-476-3393
Institution Legal Name:	University of California, San Francisco
Address (street, city, state, zip code):	550 16th Street, Second Floor
Telephone:	

Executive Summary

Atrial fibrillation (AF) is the most common arrhythmia and substantially reduces quality of life. The disease is often episodic, referred to as paroxysmal AF, and many patients report that certain exposures appear to trigger episodes. The effect of potential triggers on AF has never been rigorously investigated. A previous meeting between patients and investigators organically led to the formation of an AF interest group. That interest group, led primarily by AF patients, decided the most important research questions to patients would be to uncover whether AF triggers are in fact important and if involving AF patients in an "N-of-1" study of their own triggers would help improve their quality of life. We therefore designed the I-STOP-AF study, which will involve randomizing almost 500 paroxysmal AF patients to either AF episode tracking versus engaging in testing the relationship between patient-selected triggers and AF episodes utilizing a mobile-app based N-of-1 study design. Both groups will complete a validated survey to assess AF severity, essentially a measure of quality of life while living with AF, before and after the 10-week study period.

AIM 2: I-STOP- AF STUDY

1. Definitions and Roles

1.1. Principal Investigators: Gregory M Marcus, MD, MAS will serve as the principal investigator for the I-STOP-AF study and will be responsible for all aspects of the study.

1.2. Site Coordinating Center: The site will be coordinated using the Eureka platform based at the University of California, San Francisco under the supervision of Dr. Marcus. Madelaine Faulkner will serve as the Project Director.

1.3. Data Management Center: The Eureka platform at the University of California will manage the data. Eureka is an NIH-funded infrastructure built to facilitate mobile-health based research under the supervision of co-principal investigators Drs. Gregory M Marcus, Jeffrey Olgin, and Mark Pletcher. The data is hosted on a HIPAA-compliant AWS cloud using a multi-tenant architecture.

1.4. Study Biostatistician: Dr. Christopher Schmid at Brown University will serve as the lead biostatistician. He will supervise a graduate student/statistician who will write the automated programs to conduct the N-of-1 trials results analysis as part of the Eureka N-of-1 technology platform and will perform the statistical meta-analysis of all N-of-1 trials as well as the comparison of N-of-1 vs. data tracking.

2. Study Overview:

This is a randomized controlled trial that examines the comparative effectiveness of N-of-1 study protocols vs. symptom surveillance (data tracking alone) for reducing AF episode frequency and severity. We will recruit 478 patients with AF, equally randomized to the two study arms. Throughout the duration of the study, all study patients will use the Eureka mobile app to self-report daily mood and sleep quality, AF episode duration (in minutes) and severity, and will use the AliveCor mobile electrocardiogram (ECG) to record ECG tracings. Patients will be instructed to take ECG tracings at least once per day as well as any time they think they are having an AF episode. Patients will be able to visualize their data in real time and will receive weekly summaries of their AF frequency and severity via Eureka. The primary aims will be:

1. To compare the 10 weeks, change in AF severity (using the Atrial Fibrillation Effect on Quality of Life, or AFEQT, survey[19]) among paroxysmal AF patients randomized to AF episode tracking versus N-of-1 experiments to assess their triggers.
2. Determine from analyses of the individual N-of-1 experiments in that arm whether individual patients can determine triggers for their AF and can implement suggested behavioral changes (if any).

N-of-1 Trial arm: Patients will use the Eureka mobile application and AliveCor device to execute at least one N-of-1 trial with the goal of identifying and better controlling their AF triggers. Each N-of-1 trial will last a total of 6 weeks and will include up to 3 periods of trigger exposure and 3 periods of trigger elimination with each exposure/elimination period lasting 1 week. Patients will be randomly assigned a sequence of trigger exposure and elimination periods such that they receive each treatment in each two-week block. During each N-of-1 trial, patients will track daily AF duration and severity, daily mood and sleep quality, daily AliveCor tracings and daily trigger exposure. At the end of each trial, patients will be able to review their trial results which will include visualizations of their daily AF symptom and trigger tracking over time. After completing a trial, patients will be instructed to implement any lifestyle changes they deem appropriate based on what they learned from the results of their trial. Patients will implement these changes for a period of 4 weeks during which they will continue to track AF episode duration and severity via the app. At the end of the 4-week lifestyle change period, patients will complete the Atrial Fibrillation Effect on Quality of Life survey (AFEQT) (Appendix A) and will then have the option of testing another trigger or ending their study participation.

Symptom Surveillance arm: Patients will use the Eureka app and AliveCor device to record daily AF duration and severity, daily AliveCor readings and daily mood and sleep quality for a period of 10 weeks. Patients will be able to visualize their AF, sleep and mood data in real time and will receive a weekly summary of their data via the Eureka app. At the end of the 10-week data tracking period, patients will complete the AFEQT survey and will then have the option of either ending their study participation or crossing over to the N-of-1 trial arm to test their triggers.

Primary Study Aim:
To test the comparative effectiveness of N-of-1 trials vs. data tracking alone to identify and eliminate individual level triggers and reduce AF frequency and severity as measured by the atrial fibrillation quality of life scale (AFEQT)

Secondary Aims:
- Using individual N-of-1 trials, identify modifiable exposures that increase the risk for an AF episode on an individual-level.
- By combining data from different N-of-1 trials, identify modifiable exposures that increase the risk for an AF episode on a population-level.

Exploratory Aims:
1. To compare exposures with respect to how much they increase risk of AF episodes
2. To discover factors that determine patients who respond well to N-of-1 studies
3. To discover modifiers of the effect of different types of AF triggers

3. Study population:

3.1 Recruitment and Enrollment:

Patients will be recruited from UCSF Cardiology and General Medicine clinics, by ancillary study teams based elsewhere, and also (primarily) on the internet. At UCSF, patients are approached by a study coordinator (or the treating MD) while they are at clinic. We also post IRB approved materials, such as brochures and posters, inviting persons in the waiting room to sign up for the study, on UCSF shuttles, and other locations. We will also use the Health eHeart Study, a national, internet-based study, as a recruitment tool.[20-23] Potential subjects are referred to the website by friends, family and colleagues who have heard about the study, through collaborators such as StopAfib.org, a patient-led organization, or go to the mobile application because of our materials or press. The process of obtaining informed consent is especially important for patients signing up without an in-person coordinator to explain the study. We have carefully engineered this process, as described below.

Step 1: *Provide Information About the Study* - The majority of study patients will be approached via email. We will send a patient information about the study based on their subscription to partner organizations (stopafib.org) or the Health eHeart Study. Persons approached in the waiting room or in clinic are asked if they would be interested in participating in a research study, and given basic information about the study verbally by the study coordinator; Persons who find the mobile application read about the goals of the study, potential benefits from being in the study, who the investigators are, and which institutions are supporting the study.

Step 2: *Registration* – Individuals enter basic identifying information, including name, email address, phone number and date of birth, confirming no duplicates are in the study, and creation of a new study profile. Persons are then logged into the study website and able to provide electronic informed consent.

Step 3: *Review Study FAQ's*--On this mobile screen the patients can review an overview of the study and frequently asked questions. This is to help set expectations for participation without the legalese required by the consent form.

Step 4: *General Informed Consent (Main Consent and Surveys)* - The general informed consent form is shown on the mobile application with a reminder to scroll through to the end. The Privacy and Data Security Policy receives special emphasized treatment. After reading the form, assuming they would like to continue, they indicate that they accept and would like to join the study.

Step 5: Modular Electronic Informed Consent – Patients are invited to participate in different study activity "modules" for which they are consented separately. These modules are designed and labeled clearly as to their purpose and the incremental risks and benefits. Patients for this study will be consented using a unique modular consent. Patients will be asked to consent to participation in the Aim 2 activities (described in this proposal) and will be "sheltered" from all additional HeH surveys except those integral to Aim 2 (demographics, past medical history, AF severity).

3.2 Inclusion criteria:

- Symptomatic paroxysmal AF with observed triggers
- Smartphone/tablet with data plan

3.3 Exclusion criteria:

- children (age <18 years)
- patients with plans to substantially change AF management (e.g., ablation or change in antiarrhythmic drugs) in the 6-months after enrollment; and
- non-English speakers
- Unwillingness to test AF triggers.
- Patients who have had an AV node or AV Junction ablation

3.4 Vulnerable populations:

Importantly, this research will not involve patients performing any activities or undergoing any exposures that would not normally be subjected to during the course of their normal lives. Instead, this research will primarily seek to organize those exposures in a systematic fashion. Therefore, for example, pregnant women are not necessarily excluded. However, prisoners will likely be automatically excluded as all federal and state prisons do not allow personal smartphones. As above, children will be excluded-however, atrial fibrillation in children is extremely rare.[24]

4. Data Considerations:

4.1 **Measures**

Data that will be collected during the study are detailed in Table 2:

Table 2

Construct	Measure(s)	Categorization	Timing	Source
Primary Outcome (comparing N-of-1 group to data tracking group)				
Atrial fibrillation quality of life	AFEQT score	Integer (20-140). Survey includes 20 questions 1-7 score for each	Baseline and 10 weeks	Eureka app
Primary Outcomes (for individual-level N-of-1 trials within the N-of-1 group)				
Atrial fibrillation per AliveCor device	ECG flagged by FDA-approved algorithm	1=Yes 0=No	Daily and whenever a symptomatic episode occurs	AliveCor app
Atrial fibrillation per patient report	Patient report	1=Yes 0=No	Daily	Eureka app
Atrial fibrillation duration	Patient report (hrs and minutes)	Integer 0-60 hrs, 0-60 min	Daily (queried only if self-reported atrial fibrillation=Yes on that day)	Eureka app
Atrial fibrillation severity	Patient report	Integer 1-7, ordinal scale 1=least severe	Daily (queried only if self-reported atrial fibrillation=Yes on that day)	Eureka app
SECONDARY ENDPOINTS				
AliveCor confirmed atrial fibrillation event	Patient reported atrial fibrillation event confirmed	1=Yes 0=No	Daily and whenever a symptomatic episode occurs	AliveCor app and Eureka app

within the N-of-1 group and compared to data tracking in last month of tracking	by report from AliveCor device			
Count of AF Frequency	The number of patients reported AF episodes during the final month of enrollment	Integer 0-7—count	Daily	Eureka app
Total duration of AF episodes	Total duration of patient reported AF episodes daily during the final month of enrollment	Continuous 0-60 min, 0-24 hours 0= no answer	Daily (only if self-reported atrial fibrillation=Yes on that day)	Eureka app
Average severity of AF episodes	average severity of patient reported AF episodes daily during final month of enrollment	Continuous (range 1-7 using, Ordinal Scale 1=least severe 0= no answer	Daily (only if self-reported atrial fibrillation=Yes on that day)	Eureka app
Safety Measures				
Emergency room visit	Patient report	1=Yes 2=No 3=Don't know *If yes, date and hospital location*	Monthly: Reported at 4 & 8 weeks; actual emergency room visits will be categorized by day	Eureka app
Emergency room visit related to atrial fibrillation	Patient report	1=Yes 2=No 3=Don't know	Monthly (queried only if emergency room visit=Yes for that month) Reported at 4 &8 weeks.	Eureka app
Hospitalization	Patient report	1=Yes 2=No 3=Don't know *If yes, date and hospital location*	Monthly; actual hospitalizations will be categorized by day reported at 4 and 8 weeks.	Eureka app
Hospitalization related to atrial fibrillation	Patient report	1=Yes 2=No 3=Don't know	Monthly (queried only if hospitalization=Yes for that month)	Eureka app

			Reported at 4 &8 weeks.	
Cerebrovascular accident	Patient report	1=Yes 2=No 3=Don't know *If yes, date and location diagnosis made*	Monthly; reported at 4 &8 weeks.	Eureka app
Myocardial infarction	Patient report	1=Yes 2=No 3=Don't know *If yes, date and location diagnosis made*	Monthly; reported at 4 &8 weeks.	Eureka app
Heart failure	Patient report	1=Yes 2=No 3=Don't know *If yes, date and location diagnosis made*	Monthly; reported at 4 &8 weeks.	Eureka app
MODIFIERS				
Age	Patient report	Continuous	Baseline only	Eureka app
Gender Identity	Patient report	1=Woman 2=Man 3=Transgender Woman 4=Transgender Man 5=Genderqueer 6=Another gender identity 7=Decline to state	Baseline only	Eureka app
Sex	Patient report	1=Male 2=Female 3=Prefer not to state	Baseline only	Eureka app
Race	Patient report	1=Black 2=White 3=Asian 4=Native Hawaiian or Pacific Islander 5=American Indian or Alaska Native 6=Some other race 7=Don't know	Baseline only	Eureka app

		8=Prefer not to state		
Hispanic	Patient report	1=Not Hispanic 2=Yes: Mexican, Mexican American or Chicano 3=Yes: Puerto Rican 4=Yes: Cuban 5=Yes: Mixed Hispanic, Latino or Spanish Origin 6=Yes: Other Hispanic, Latino, or Spanish Origin 7=Don't know 8=Prefer not to state	Baseline Only	Eureka app
Education level	Patient report	1=No formal schooling 2=Some school, but did not graduate high school 3=High school diploma or equivalent 4=Associate degree 5=Some college 6=Bachelor's degree 7=Master's degree 8=Doctorate 9=Professional doctorate 10=Other 11=Don't know 12=Prefer not to state	Baseline Only	Eureka app
Household income	Patient report	1= Under $10,000 2= $10,000 to under $20,000 3= $20,000 to under $30,000	Baseline Only	Eureka App

		4=$30,000 to under $40,000 5= $40,000 to under $50,000 6= $50,000 to under $75,000 7= $75,000 to under $100,000 8= $100,000 to under $150,000 9= More than $150,000 10= Don't know 11= Prefer not to state		
Household size	Patient report	Integer 0-20	Baseline Only	Eureka app
Height	Patient report	Integer 1-10 ft 0-11 inches	Baseline Only	Eureka app
Weight (pounds)	Patient report	Integer 1-999	Baseline Only	Eureka app
Hypertension	Patient report	1=Yes 2=No 3=Don’t Know	Baseline Only	Eureka app
Diabetes	Patient report	1=Yes 2=No 3=Don’t know	Baseline Only	Eureka app
Coronary artery disease	Patient report	1=Yes 2=No 3=Don’t know	Baseline Only	Eureka app
Heart Attack	Patient report	1=Yes 2=No 3=Don’t know	Baseline Only	Eureka app
Peripheral Artery Disease	Patient report	1=Yes 2=No 3=Don’t know	Baseline Only	Eureka app
Congestive Heart Failure	Patient report	1=Yes 2=No 3=Don’t know	Baseline Only	Eureka app
Stroke	Patient report	1=Yes 2=No 3=Don’t know	Baseline Only	Eureka app
Heart murmur	Patient report	1=Yes 2=No 3=Don’t know	Baseline Only	Eureka app

Sleep Apnea	Patient report	1=Yes 2=No 3=Don't know	Baseline Only	Eureka app
First degree family member with atrial fibrillation	Patient report	1=Yes 2=No 3=Don't know	Baseline Only	Eureka app
Currently on antiarrhythmic drugs	Patient report	1=Flecainide 2=Propafenone 3=Aminodarone 4=Dronedarone 5=Sotalol 6=Dofetilide 7=None of the above	Baseline Only	Eureka app
Currently take antiplatelets and anticoagulants	Patient report	1= Wafarin 2=Aspirin 3=Dabigatran 4=Rivaroxaban 5=Apixaban 6 =Edoxaban 7=Clopidogrel 8=Prasugrel 9=None of the above	Baseline Only	Eureka app
Sleep	Patient report	1=Amazing 2=Good 3=Average 4=Bad 5=Horrible	Daily	Eureka app
Mood	Patient report	1=Neutral 2=Happy 3=Depressed 4=Angry 5=Anxious 6=Stressed	Daily	Eureka app

4.2. Primary Endpoints

1. **Atrial fibrillation quality of life** will be measured by the AFEQT survey monthly and will be used in comparing of the N-of-1 arm to the tracking arm. The primary outcome will be change from baseline (randomization) to 10 weeks.
2. **Self-reported atrial fibrillation event** Each day patients will report if they have had an atrial fibrillation event.
3. **Atrial fibrillation duration** Each day patients will report the total amount of time during which they had atrial fibrillation adding up time over all events over the day. For days on which no event occurred, the time will be zero.

4. **Atrial fibrillation severity** Each day patients will record the severity of their atrial fibrillation on a scale of 1-7 with a score of 1 given if no event occurred.

4.3. Secondary Endpoints

1. **AliveCor confirmed atrial fibrillation event** Daily self-reported atrial fibrillation event confirmed by report from AliveCor device.
2. **Average frequency of AF episodes per week** For each patient, the average number of AF episodes per week during the final month of each patient's enrollment as reported by the patient will be calculated.
3. **Average duration of AF episodes per week** For each patient, the average duration among all AF episodes per week during the final month of each patient's enrollment as reported by the patient will be calculated.
4. **Average severity of AF episodes per week** For each patient, the average severity among all AF episodes per week during the final month of each patient's enrollment as reported by the patient will be calculated.

4.4. Safety Endpoints

1. **Hospitalization** Days on which any hospitalization occurs together with date and hospital location. This will be recorded from the Eureka app and will be reported monthly by the patient.
2. **Hospitalization related to atrial fibrillation** Days on which a hospitalization related to atrial fibrillation occurs together with date and hospital location. This will be recorded from the Eureka app and will be reported monthly by the patient.
3. **Emergency department visit** Days on which any emergency department visit occurs together with date and hospital location. This will be recorded from the Eureka app and will be reported monthly by the patient.
4. **Cerebrovascular accident** Days on which any cerebrovascular accident occurs together with date and hospital location. This will be recorded from the Eureka app and will be reported monthly by the patient.
5. **Myocardial infarction** Days on which a myocardial infarction occurs together with date and hospital location. This will be recorded from the Eureka app and will be reported monthly by the patient.
6. **Heart failure** Days on which heart failure occurs together with date and hospital location. This will be recorded from the Eureka app and will be reported monthly by the patient using language from previous HeH surveys.

4.5 **Potential Treatment Effect Modifiers**

Within the N-of-1 experiments and in the comparison of N-of-1 to data tracking, it is recognized that external factors could influence willingness to adhere to a treatment assignment, could interact with the exposure and AF, and could affect AF episodes. Per suggestions from AF patients, we will therefore perform daily assessments of mood, sleep (the night previous) and stress. These will be obtained via push notifications and will therefore rely on patient self-report. Relationships between these factors and both treatments (N-of-1 and usual care) and outcomes will be investigated; given evidence of statistically significant relationships, results will be adjusted for these covariates.

1. **Medical History:** Characteristics as outlined in Table 2 collected at baseline by self-report include hypertension, diabetes, heart failure, history of cerebrovascular accident, first degree

family member with atrial fibrillation and history of atrial fibrillation and history of no disease other than atrial fibrillation.

2. **Demographics:** Basic demographic data as outlined in Table 2 including age, gender, race/ethnicity, height, weight, and education level will be collected both to describe the population of patients and to assess for heterogeneity of treatment effects at the population level.
3. **Current medications:** Current medications being taken including antiarrhythmic drugs, aspirin, Plavix, warfarin or others as noted in Table 2 will be recorded.
4. **Mood and sleep:** Patient self-report on mood and sleep will be collected via short survey on the Eureka app on a daily basis.

4.6 **Data Sources**

4.6.1. **Case Report Forms**: Because the study is conducted remotely and via a mobile app, it will not employ formal case report forms. Instead, surveys will be delivered via the mobile app in two forms: 1. During initial on-boarding, patients will answer a series of surveys delivered in consecutive order. These surveys will remain on the app (and will be the first thing a patient sees when they interact with the app) until they are completed. 2. Short surveys will be delivered via push notifications from the app.

4.6.2 Eureka Technology Platform: Individuals will enter data via the Eureka mobile app. As with the great majority of research in the Health eHeart Study that is obtained remotely and from patient self-entry, the study will rely on a framework of patient reported data. Patients who fail to complete surveys will be sent reminders via app-related push notifications or SMS text messages.

4.7 **Data Access and Reports** During the study, the Eureka data management team at the University of California, San Francisco will have access to the raw study data from all data sources. Standardized reports will be generated to allow the study team to monitor recruitment and retention, progression through the phases of the N-of-1 trial and/ or data tracking arm, completion of surveys, and responses to study activity notifications. These outputs will be visualized in real-time using the Eureka administrative portal and reports can be printed or provided as a PDF as needed. In coordination with the study statistician, the Eureka data management group will also prepare reports for the data safety and monitoring board (DSMB). These reports will include data on: accrual, data quality and completeness, study withdrawals, rates of completion of N-of-1 trials, and hospitalizations and emergency room visits.

At the conclusion of the study, the data coordinating center will work with the study biostatistician to prepare the analytic data files to be used in the final analysis. The study statistician will only be obtaining and using de-identified data for this project. Please see attached letter from Brown IRB regarding the non-human subjects waiver.

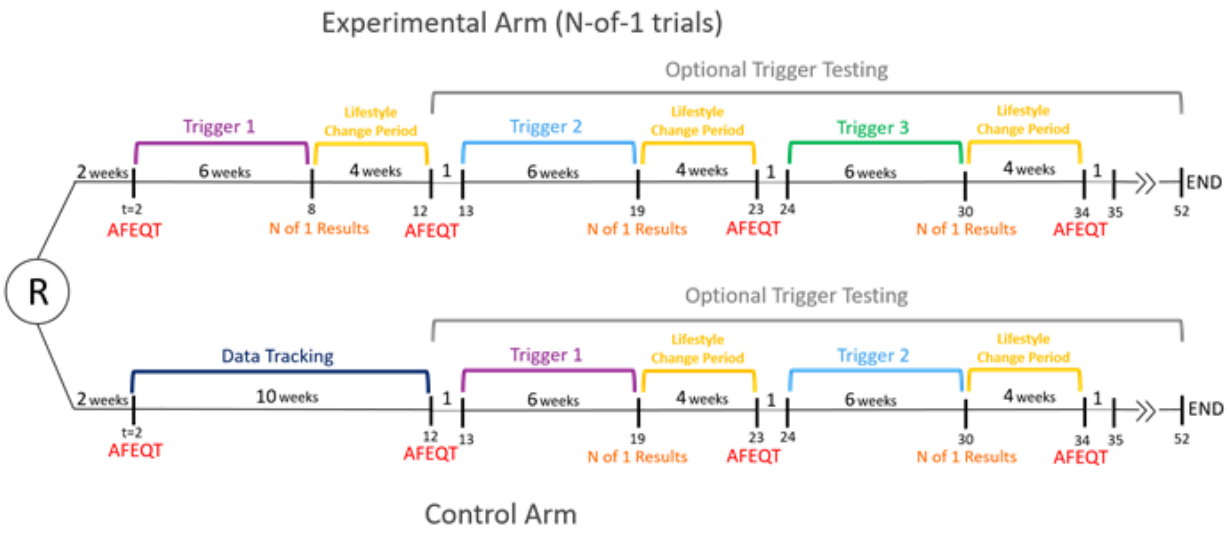
5. Analytic Plan

5.1 Overview

Figure 2 details the study design for Aim 2 which is designed as an RCT to compare use of N-of-1 studies vs. symptom surveillance (data tracking alone) using comparable mHealth technology in both study arms. Working with the HeH patient stakeholders, we will design a menu of triggers that people with AF may select to test in their own personal N-of-1 study. This menu will include approaches for modifying triggers including alcohol, caffeine, exercise (frequency and intensity), sleep, liquid intake/dehydration,

meal size, sleep position (e.g., lying on left side), beverage and food temperature, or choices at the option of the patient. These N-of-1 protocols will involve 6 randomly assigned treatment periods (each 1-week in duration) of exposure (trigger "on") and elimination (trigger "off").

Figure 2



5.2 Sample Size for Number of Study Patients

As those with at least moderate AF severity have an AFEQT score of 58 ± 19[19] and the most modest of therapies typically improves this score by a mean of 5-10, we have powered this study conservatively to detect an improvement of 5 points in the AFEQT. We estimate 239 patients in each group would provide 80% power to detect this small, yet still clinically meaningful difference.

5.3 Randomization to Treatments

Completion of the electronic consent via the Eureka portal will trigger 1:1 randomization to the N-of-1 or the symptom surveillance conditions. The randomization schema will be conducted within the Eureka platform and communicated directly to patients via the mobile app.

5.4 Analysis

5.4.1 Comparison of N-of-1 to data tracking

The experimental design for Aim 2 is a two-arm randomized trial comparing N-of-1 vs. symptom surveillance. The primary endpoint is AF severity via the AFEQT from baseline to 10 weeks. Differences between treatment groups will be compared using analysis of covariance with the 10 week measurement as the outcome and the baseline measurement as covariate. We will also test for an interaction between treatment and baseline AFEQT in order to measure possible treatment effect heterogeneity. The average frequency, duration and severity of AF episodes per week during the final month of each patient's enrollment will be secondary outcomes. Their group means will be compared with a t-test. Additional secondary outcomes will include safety events: hospitalizations, emergency room visits, stroke, myocardial infarctions and heart failure. The main analysis will treat each of these as a binary outcome indicating the incidence of any event. We will also investigate models for multiple

events per person if these are common. We will check for balance in covariates across the randomized groups and will adjust for any imbalances by including those covariates in regression analyses and reporting the adjusted treatment effect as a secondary analysis. The null hypothesis of no difference will be tested at a significance level of 0.05.

5.4.1.1 Heterogeneity of treatment effects (HTE)

N-of-1 trials may be more effective for some types of patients than for others. In addition, to checking interaction with baseline AFEQT, we will examine interactions of treatment with the potential between-patient modifiers listed in Table 2 in regression models as exploratory analyses. These will include medical history, demographics, medications, mood and sleep.

5.4.1.2 Missing Data

We will attempt to obtain a 10-week AFEQT value when a patient drops out of the study. When the primary 10-week endpoint is missing for a patient, we will use different approaches to estimate sensitivity of conclusions to the missing data assumption: 1) disregard the missing outcome and discard the patient from analysis; 2) multiple imputation; 3) assume no change from baseline to 10 weeks. The third analysis is likely to be conservative, but makes a fairly strong assumption about the missing not at random assumption (MNAR).

Other endpoints such as average frequency, severity, duration and number of safety events will be calculated based on the data reported. We will assume that days where no safety events or AF events are reported are event-free and so there will be no missing values for the other outcomes.

5.4.2 Analysis of Individual N-of-1 Trials

We will analyze each patient's trial of a specific trigger using the same methods as in Aim 1 applied to the outcomes of frequency, severity and duration of events reported daily. Frequency will be analyzed as a binary event on each day. Severity and duration will be analyzed as a mixture model between non-event days and event days with the mixture probability corresponding to the probability of an event.

5.4.3 Meta-Analysis of N-of-1 Trials

We will conduct separate meta-analyses of each set of N-of-1 trials that test a particular trigger. We expect this will likely be most relevant to the triggers offered in the menu which patients are more likely to try and which were also the most commonly identified triggers in a survey recently sent to 1,000 AF patients. The meta-analyses will use the same Bayesian multilevel methods as those outlined in Aim 1 and will be applied to the outcomes of AF events, duration and severity as well as the safety endpoints listed in Table 2. We will use distribution and link functions appropriate to the scale of each outcome variable. We will calculate the average treatment effect in the group of patients as well as the individual effects adjusted for the other patients.

Patients with AF are particularly interested in whether certain triggers are more likely than others to induce AF. Because patients in this study will be choosing to study different triggers and often will be conducting more than N-of-1 trial, the set of trigger comparisons will form a network of triggers (plus no trigger) that may be analytically compared using methods of network meta-analysis.[25] Triggers may then be compared indirectly through the common control using techniques from network meta-analysis. The complicating factor here is that trials done by the same patient will be correlated and it will be necessary to adjust for this correlation in the network meta-analysis model. This can be done in two different ways. The first would be to treat each patient's trials as separate and introduce a common patient effect into a

mixed model. The second would be to combine all the trials for a patient into one trial with several arms, one for each different trigger and one for no trigger that would combine all the results from the no trigger arms. If a patient changed the control (no trigger) arm in a new trial (e.g., by testing two triggers vs. one trigger), then the control arm would be analyzed according to the trigger used. This approach would create multi-arm trials for patients with multiple trials. These would then be analyzed by methods for multi-arm studies in a network meta-analysis.

5.4.3.2 Heterogeneity of Treatment Effects

While N-of-1 trials are specifically constructed to evaluate individual treatment effects, it is also of interest whether particular triggers are more or less likely to have an effect on individuals with particular characteristics, particularly if the differential effects of triggers are moderated by modifiable factors or by individual characteristics that may indicate which triggers are most important. We will examine the set of modifiers listed in Table 2 that include medical history, demographics, medications, mood, stress and sleep by including these between-patient factors in the multilevel models.

5.4.3.3 Missing Data

Missing data in individual N-of-1 trials are likely going to be most pertinent to the daily measures of AF and duration and severity. We will treat days on which the patient makes no report as having no AF event. As with the analyses in Aim 1 with cases of missing outcome data, we will both ignore the missing values (assuming missing at random) and treat the missing values as latent parameters and estimate them as part of the Bayesian model assuming them to be missing at random. If we suspect that data are missing not at random (MNAR), we will construct MNAR models based on the appropriate missing data patterns to determine sensitivity of inferences to this assumption. If any data on effect modifiers are missing, we may also attempt multiple imputation, particularly if the number is large and the number of latent parameters to be estimated becomes too large.

6.Potential Impact:

6.1 Potential Risks & Protection Against Risks

1. *Loss of Confidentiality:* There is a minor risk related to loss of confidentiality of responses to questions asking for an assessment of the individual's physical condition. There is minimal risk of a loss of privacy if another person sees responses through the mobile platform. The primary reasons this would occur would be because a patient lost their device or if they provided another person access to their mobile device.
    - Patients will submit their data via SMS text messages, via mobile applications or via secure websites. SMS messages and mobile notifications are transmitted through 3rd party services (e.g., Twilio and Urban Airship). Simple, non-PHI carrying messages are sent by the system, and typically unidentifiable numeric responses are returned. The only PHI involved in the transaction is the phone number or unique ID used to identify the patient's phone. Patients can opt out of these mechanisms and respond entirely through a secure website if they choose.
    - The platform will also be able to communicate with commercial devices and other data sources managed by the patient (e.g., AliveCor). When logging into the platform, the patient can actively elect to allow the platform to communicate with these external data sources and devices. The use of these devices is elective and therefore not necessary in order to participate in the study. The patient will be responsible for the

security of these data when housed with the external vendor. Once data from these sites/devices are imported in the platform, data security procedures as previously described will apply.

2. *Physical Risk:* All patients will be instructed to follow their physician's advice regarding treatment of their AF. The number of symptomatic AF episodes and severity of these episodes (the main measures proposed in Aim 2 of this study and the target of behavioral modification strategies), while extremely troublesome for many patients are not a safety concern.[26] It is important to emphasize that AF is not an imminently dangerous rhythm and that the primary issue regarding safety is stroke and thromboembolic prophylaxis. Stroke and thromboembolic risk is not thought to be related to the number of AF episodes or potentially even AF episodes themselves. Specifically, per the AF guidelines, anticoagulation needs to be dictated by stroke risk score and not by whether AF is paroxysmal or persistent or based on the frequency of disease.[26] Recent trials with continuous monitoring of atrial activity have demonstrated that there is no relationship between the timing or duration of AF episodes and the timing of strokes that occur.[27] This is also supported by evidence from multiple randomized trials that have clearly demonstrated that a rhythm-control strategy (i.e., a strategy involving obtaining and maintaining normal sinus rhythm) does not protect against stroke.[28,29] Therefore, changing potential triggers to reduce episodes primarily serves quality of life purposes and should not influence the overall safety of patients. The behavioral modification that patients will undertake as part of an N-of-1 trial is analogous to current recommendations that view strategies to obtain and maintain a normal rhythm (whether it be with drugs or invasive ablation) as a means to attain improved quality of life.
3. *Psychological Risk:* Patients and parents will be assured during the informed consent discussions that participation is voluntary. There may be a risk of psychological discomfort for patients associated with completion of the surveys, questionnaires, or other information required to conduct the trial, including feelings of embarrassment or distress when reviewing data with the doctor. They may also feel anxious or confused by using a device or web application that they would not have otherwise used if they had not been participating in the study. Participation in the study is completely voluntary, and patients are free to leave the study at any time.

6.2 Protection of Human Subjects

The investigators will take specific protections against these risks. Patients will be advised to contact the appropriate study Principal Investigator and/or IRB with any questions, concerns or complaints associated with their participation. Study coordinators will be available by email and phone to discuss any issues or uncomfortable questions that arise or if patient have problems associated with the use of Orchestra or another third-party device or application.

6.3 Potential Benefits of the Proposed Research to Human Subjects and Others

By using an N-of-1 approach to answer patient-identified research questions, we will generate individualized evidence about the role of triggers in paroxysmal AF. This methodology provides a direct benefit to patients by helping provide them a greater certainty about their treatment choices including the relative benefits vs. burdens of avoiding potential triggers (such as alcohol, caffeine, or chocolate). This has tremendous potential to improve their individual health in a way that truly personalized and patient centered. The risks of participating are small. Also, the risks and benefits of participating in the study are easy for study patients to understand, and can be weighed on an individual basis. Participation may help AF patients identify triggers for their disease and subsequent modification of

behavior or exposures may reduce their AF burden. The risk to benefit ratio favors conducting this study.

6.4 Importance of the Knowledge to be Gained

Using N-of-1 methods to help AF patients identify the best treatment will shift the clinical paradigm towards a more personalized and patient-centric approach care and provide patients with individualized data about the role of trigger modification in controlling symptoms. We hope that study participation will lead to a better understanding of the triggers for AF and optimal means to manipulate those triggers to enhance the health of AF patients.

6.5 Data and Safety Monitoring Plan

Although randomization to N-of-1 trials versus symptom surveillance (data tracking) is low risk, we will form a DSMB to ensure the safety of patients. This board will be comprised of a patient, a cardiac electrophysiologist, and a statistician from outside the research groups or institutions of the study personnel otherwise involved in the project. They will meet approximately 12 weeks after the implementation of the study to review initial trigger testing results and adverse event reports (if applicable.) An analyst will prepare summary statistics of results collected prior to the meetings. The DSMB will review study progress and any adverse events that have been reported by patients. The DSMB will have the authority to stop the study for safety reasons or to alter it if needed. The DSMB will not perform any interim analyses for efficacy or consider stopping the study early for futility. I

References

19. Spertus J, Dorian P, Bubien R, et al. Development and validation of the Atrial Fibrillation Effect on QualiTy-of-Life (AFEQT) Questionnaire in patients with atrial fibrillation. *Circ Arrhythm Electrophysiol.* 2011;4(1):15-25.
20. Brooks GC, Vittinghoff E, Iyer S, et al. Accuracy and Usability of a Self-Administered 6-Minute Walk Test Smartphone Application. *Circ Heart Fail.* 2015;8(5):905-913.
21. Christensen MA, Bettencourt L, Kaye L, et al. Direct Measurements of Smartphone Screen-Time: Relationships with Demographics and Sleep. *PLoS One.* 2016;11(11):e0165331.
22. Dixit S, Pletcher MJ, Vittinghoff E, et al. Secondhand smoke and atrial fibrillation: Data from the Health eHeart Study. *Heart Rhythm.* 2016;13(1):3-9.
23. Whitman IR, Agarwal V, Nah G, et al. Alcohol Abuse and Cardiac Disease. *J Am Coll Cardiol.* 2017;69(1):13-24.
24. Roger VL, Go AS, Lloyd-Jones DM, et al. Heart Disease and Stroke Statistics--2012 Update: A Report From the American Heart Association. *Circulation.* 2012;125(1):e2-e220.
25. Efthimiou O, Debray TP, van Valkenhoef G, et al. GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods.* 2016;7(3):236-263.
26. January CT, Wann LS, Alpert JS, 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(23):e199-267.
27. Daoud EG, Glotzer TV, Wyse DG, et al. Temporal relationship of atrial tachyarrhythmias, cerebrovascular events, and systemic emboli based on stored device data: a subgroup analysis of TRENDS. *Heart Rhythm.* 2011;8(9):1416-1423.
28. Van Gelder IC, Hagens VE, Bosker HA, et al. A comparison of rate control and rhythm control in patients with recurrent persistent atrial fibrillation. *N Engl J Med.* 2002;347(23):1834-1840.

29. Wyse DG, Waldo AL, DiMarco JP, et al. A comparison of rate control and rhythm control in patients with atrial fibrillation. *N Engl J Med.* 2002;347(23):1825-1833.

Appendix: **Detailed derivation of sample sizes for Aim 1**

Assuming a standard crossover design in which

$$Y_{ijklt} = E\left(Y_{ijklt}\right) + \beta_{ij} + \varepsilon_{ijklt}$$

where

i indexes patient within each crossover sequence, i = 1, …, I

j indexes crossover sequence, j = 1,…,J

k indexes crossover, k = 1,…, K

l indexes observation within period l = 1,…,L

and t indexes treatment, t = 0,1

we have

$E(Y_{ijklt}) = \mu$ if t = 0

$E(Y_{ijklt}) = \mu + \tau_{ij}$ if t = 1

assuming no period effects and no carryover; $\tau_{ij} \sim N(\tau, \sigma_t^2)$ is the random treatment effect varying across patients; $\beta_{ij} \sim N(0, \sigma_b^2)$ is the random between-person effect and $\varepsilon_{ijklt} \sim N(0, \sigma_w^2)$ is the random within-person effect. The crossover treatment effect can be estimated by subtracting the average response among all control periods from the average response among all treatment periods. This may be written

$$\begin{aligned} CROS &= \bar{Y}_{ijkl1} - \bar{Y}_{ijkl0} \\ &= \sum_{j=1}^{J}\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{l=1}^{L}\left(Y_{ijkl1} - Y_{ijkl0}\right) / IJKL \\ &= \frac{1}{IJKL}\sum_{j=1}^{J}\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{l=1}^{L}\left\{\left(\mu + \tau_{ij} + \beta_{ij} + \varepsilon_{ijkl1}\right) - \left(\mu + \beta_{ij} + \varepsilon_{ijkl0}\right)\right\} \\ &= \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}\tau_{ij} + \frac{1}{IJKL}\sum_{j=1}^{J}\sum_{i=1}^{I}\sum_{k=1}^{K}\sum_{l=1}^{L}\left\{\varepsilon_{ijkl1} - \varepsilon_{ijkl0}\right\} \end{aligned}$$

This has expectation τ and variance consisting of the sum of the variance of the hetoergeneous treatment effects and the variances and correlations among the 2KL observations on each individual. Assuming the treatment effects are independent of the random errors, and that the errors have the same correlation, ρ, the variance is

$$
\begin{aligned}
Var(CROS) &= \frac{1}{(IJ)^2} Var\left(\sum_{j=1}^{J} \sum_{i=1}^{I} \tau_{ij} \right) + \frac{1}{(IJKL)^2} Var\left(\sum_{j=1}^{J} \sum_{i=1}^{I} \sum_{k=1}^{K} \sum_{l=1}^{L} \left\{ \varepsilon_{ijkl1} - \varepsilon_{ijkl0} \right\} \right) \\
&= \frac{\sigma_\tau^2}{IJ} + \frac{1}{(IJKL)^2} \left\{ 2IJKL\sigma_w^2 + 2IJ\sigma_w^2 \rho \begin{pmatrix} 2KL \\ 2 \end{pmatrix} \right\} \\
&= \frac{\sigma_\tau^2}{IJ} + \frac{2\sigma_w^2}{IJKL} \left\{ 1 + \frac{\rho(2KL)(2KL-1)}{2KL} \right\} \\
&= \frac{\sigma_\tau^2}{IJ} + \frac{2\sigma_w^2}{IJKL} \left\{ 1 + \rho(2KL-1) \right\}
\end{aligned}
$$

In particular, if the observations are independent, then the variance is $\frac{\sigma_\tau^2}{IJ} + 2\sigma_w^2 / IJKL$. Letting P = IJ be the total number of patients and M = KL be the number of measurements taken for each patient on treatment and control, we have variance of $\left(\sigma_\tau^2 + 2\sigma_w^2 / M \right) / P$ which shows that the sample size decreases by a factor proportional to the number of patients and partly to the number of measurements (assuming the measurements are independent conditional on person).

Since the factor by which the variance is inflated assuming constant correlation is 1+(2KL-1) and KL = 14, we have an inflation factor of 1+27. This is large, but unrealistic because it is unlikely that observations taken one week apart for 32 weeks would have a constant correlation. It might be more reasonable to assume first-order autocorrelation. But recall that to remove carryover the first observation of each treatment period is not evaluated so not all observations are completely equally spaced. Thus, assuming autocorrelation will overstate the total variance. But if we do assume first-order autocorrelation, it can be shown that

$$
\begin{aligned}
Var(CROS) &= \frac{\sigma_\tau^2}{IJ} + \frac{1}{(IJKL)^2} \left\{ 2IJKL\sigma_w^2 + 2IJ\sigma_w^2 \left\{ \rho\left(2KL-1\right) + \rho^2(2KL-2) + \ldots + \rho^{2KL-1} \right\} \right\} \\
&= \frac{\sigma_\tau^2}{IJ} + \frac{2\sigma_w^2}{IJKL} \left\{ 1 + \frac{2}{2KL} \left\{ \rho\left(2KL-1\right) + \rho^2(2KL-2) + \ldots + \rho^{2KL-1} \right\} \right\} \\
&\le \frac{\sigma_\tau^2}{IJ} + \frac{2\sigma_w^2}{IJKL} \left\{ 1 + 2 \sum_{k=1}^{2KL-1} \rho^k \right\} \\
&\doteq \frac{\sigma_\tau^2}{IJ} + \frac{2\sigma_w^2}{IJKL} \left\{ 1 + 2\rho / (1-\rho) \right\}
\end{aligned}
$$

so the inflation factor is $1 + 2\rho/(1 - \rho)$ which is much smaller.