# Statistical Analysis Plan

## ASSURE WCD Clinical Evaluation – Detection and Safety Study

## (ACE-DETECT)

**Protocol: 3324582_B (dated 17 Aug 2018)**
**NCT03887052**

**Sponsor:**
Kestra Medical Technologies.
3933 Lake Washington Blvd Suite 200
Kirkland, WA 98033.
(800) 957-0028

**Version: #1**
19 MAR 2019

## Confidential Information



**Primary Author:**
Kaisa Kivilaid, MS
Principal Biostatistician
Regulatory and Clinical Research Institute, Inc.
Minneapolis MN, 55416
Ph: 952.746.8030

Signature

19-Mar-2019
Date

**Primary Sponsor Contact:**
Laura Gustavson
VP Clinical Marketing
Kestra Medical Technologies, Inc.
Kirkland, WA 98033
Ph: 425-633-0651

Laura Gustavson
Signature

19-MAR-2019
Date

# Table of Contents

# 1. Purpose of the Statistical Analysis Plan

The purpose of this Statistical Analysis Plan (SAP) is to summarize the proposed statistical methodology used to determine the required sample size and describe the statistical analyses that will be employed to evaluate safety and detection performance of a new wearable automatic cardioverter defibrillator device in adult cardiac subjects at risk for sudden cardiac arrest as part of the overall system validation.

# 2. Study Design

## 2.1 Study Design/Overview

The ASSURE WCD Clinical Evaluation – Detection and Safety Study (ACE-DETECT) is a multicenter single arm open label evaluation. All subjects will be fitted with the WCD and will be trained and provided written instructions for use, including safety precautions and adverse event reporting. Demographic data and limited medical history data will be collected. The subjects will be asked to complete a BORG Comfort Scale at Visit 1 (baseline). Subjects will wear the device for approximately 30 days during normal daily activities including sleep. Research Coordinators will conduct weekly phone interviews with the subjects during the participation period to address subject questions, review potential adverse events, and review usage. Subjects will return for a final visit, Visit 2, at the end of the wear period. At Visit 2 subjects will be interviewed regarding any adverse events they may have experienced since the most recent phone follow-up and the status of any adverse events that were recorded during the participation period. They will also be asked to complete a device usability survey and the BORG Comfort Scale after the device is removed. All episodes recorded by the ASSURE WCD, the subject's ICD, or other devices when available, will be reviewed by independent clinical experts.

## 2.2 Investigational Device

The WCD used for this study will be a configuration controlled by the Sponsor from managed inventory designated for clinical evaluations only. The devices will be production-equivalent systems with shock alarms disabled (Shock Alarm Event Markers will be recorded by the device for analysis purposes), defibrillation therapy programmed OFF, and detection parameters at default settings. Detection parameter default settings include Ventricular Tachycardia (VT) rate threshold at 170 bpm and Ventricular Fibrillation (VF) rate threshold at 200 bpm.

The ASSURE™ Cardiac Recovery System is comprised of several components:

- A WCD including a Garment with embedded ECG sensors that is worn on the body next to the skin, a Monitor which houses the battery and the graphic user interface, Cable Assembly which connects the Garment to the Monitor, and a Carry Pack
- A Charger that allows recharging of the monitor batteries
- A Tablet that allows the WCD to be programmed via wireless connection to the Monitor
- CareStation™ website which allows users to view WCD data associated with study subjects
- An ASSISTANT™ mobile device that allows data uploads from the WCD to CareStation website

# 3. Study Objectives

Primary: To evaluate ambulatory detection performance of the ASSURE Wearable Cardioverter Defibrillator (WCD)

Secondary: To evaluate arrhythmia detection performance and safety of the ASSURE WCD

# 4. Endpoints and Hypotheses

The endpoint analyses will be performed after all subjects have either completed the final visit or withdrawn early from the study, the study database has been cleaned, verified and locked, and adjudication/reconciliation of all episodes has been completed.

## 4.1 Primary Endpoint

The primary endpoint is the WCD False Positive Alarm (FPA) Rate calculated as False Positive Alarms per subject-day where days per subject will be calculated by taking the total wear time in minutes as recorded in the eCRF by the Research Coordinator from the CareStation remote monitoring system divided by 1,440 minutes. The study will be considered successful if the WCD FPA Rate is less than 0.29 FPAs per subject-day.

The hypotheses to be tested are as follows:

$H_0$: $p_1 \geq 0.29$[1]
$H_1$: $p_1 < 0.29$[1]

where
$p_1$ = ASSURE False Positive Alarm Rate

## 4.2 Sample Size Estimation

The following assumptions were used to determine the sample size for the primary endpoint of WCD False Positive Alarm Rate:

- The primary analysis is a one-sided comparison of the WCD FPA rate per subject-day against the value of 0.29
- The WCD FPA rate is expected to be 50% lower than 0.29
- At least 90% power is considered sufficient
- The statistical test will be performed with a 2.5% significance level
- Average wear time per subject, per day, is 14 hours
- The FPA rate has a Poisson distribution with variance to mean ratio, ρ, = 14 and an intra-site correlation =0.026.[2]
- For large total number of expected events, the Poisson (λ) distribution is approximately normally distributed as N ($\mu=\lambda$, $\sigma^2=\rho*\lambda$).

To maintain at least 90% power to detect a FPA rate 50% lower than 0.29 a sample size of approximately 105 subjects is necessary. However, the study will enroll at least 130 subjects to increase the likelihood that a subject will experience a true ventricular arrhythmic event.

Sample size calculations were performed using the R statistical package.

## 4.3 Secondary Endpoints

- True Positive Detection:

[1] Equivalent to one false alarm every 3.43 subject days (2.0 per subject week of use) as cited in LifeVest Model 4000 Operator's Manual, Rev C; 7-5.)
[2] Estimates based upon sponsor study 'Electrocardiogram (ECG) Recordings from Cardiac Patients during Ambulatory Use with Assure™ Prototypes', protocol 3330766, September 2017.

The incidence of WCD True Positive Detections will be summarized for each adjudicated episode (see Appendix A for format of episode summary). Additional details are found in the Section 10. Definitions.

- Missed Events (False Negative Detections):
  The incidence of WCD Missed Events (False Negative Detections) for each adjudicated episode will be summarized (see Appendix A for format of episode summary). Additional details are found in the Section 10. Definitions.

- Estimated Inappropriate Shock Rate:
  An estimate of the Inappropriate Shock Rate, per subject-day, will be calculated as the product of the WCD FPA rate and the Missed Shock Alarm Rate where the WCD FPA rate, per subject-day is estimated from the primary endpoint of this study, and the Missed Shock Alarm Rate is estimated as the upper limit of the 95% confidence interval of 23.1% based on data from two previous studies (GDR 3333597 ASSURE UISim Sleep Study and GDR 3334121 ASSURE UISim 7-day study). See the protocol for additional information.

- Adverse Events Determined to Be at Least Possibly Related to the Device:
  Summary of adverse events that are determined to be at least possibly related to the device.

Other measures, for consideration of future investigation, include, but are not limited to, average daily wear time, which will be calculated for each subject, overall wear time, with the first and last day of the study not included, BORG Comfort Scale, and a Usability Survey.

## 4.4 Subgroup Analyses

Subgroup analyses based on, but not limited to, gender, age, BMI and baseline QRS width will be provided on the primary effectiveness endpoint. These analyses will be regarded as exploratory, and have been selected for the following reasons:

- Gender – The ASSURE WCD is available in two styles (male and female). Garment fit may influence false alarm performance.
- Age – Higher levels of activity may be associated with higher levels of noise, and activity levels may be influenced by age. Subgroups will be based on a cutoff point of 75 years (<75/≥75).
- BMI – Data presented previously on the LifeVest suggest that obese subjects experience a significantly higher percentage of noise, which may influence false alarm performance. (Wan C et al. ESC 2013. The impact of body mass index on wearable cardioverter defibrillator efficacy. Abstract P4926 Poster.) Subgroups will be based on a cutoff point of 30 (</≥ 30).
- QRS width – Variable used for rhythm classification in the ASSURE WCD detection algorithm. Subgroups will be based on a cutoff of 120 ms (</≥ 120).
- ICD Device Type - Subjects are eligible for enrollment with either a single or dual chamber ICD. Identification of Episode Rhythm type may be facilitated by the presence of a dual chamber ICD.

## 4.5 Adverse Events

All adverse events reported over the course of the study will be summarized and tabulated by category, seriousness, and relationship to the study device (possibly, probably or definitely). No formal statistical testing is planned.

# 5. Analysis Methods

## 5.1 Primary Endpoint

A random effects Poisson regression model will be fit with the number of FPAs for each patient as the outcome variable, the logarithm of days of wear as an offset term, and a random site effect. The fitted model will be used to calculate a two-sided 95% confidence interval for the FPA rate of the study device. The null hypothesis will be rejected if the upper bound of the 95% confidence interval is less than the rate reported for the comparator device (0.29). Note that a two-sided 95% confidence interval corresponds to a two-sided test with significance level 0.05 or a 1-sided test with significance level 0.025.

If the null hypothesis is rejected, the conclusion will be that the FPA rate for the WCD study device is significantly lower than the rate reported for the comparator device.

Assumptions regarding the Poisson regression model will be checked by assessing for example the deviance and goodness of fit statistics and dispersion. If substantive violations of model assumptions are suspected, a sensitivity analysis will be performed using the bootstrap method with resampling of cases. Ten thousand bootstrap samples will be created, and confidence intervals will be constructed by the percentile method. In the event there are significant violations of model assumptions for the Poisson regression model, the bootstrap confidence interval will be utilized to test the hypotheses on the primary endpoint.

An additional random effects Poisson regression model will be fit including patient characteristics (including age, gender, BMI, baseline QRS width, and ICD type) as covariates in an exploratory analysis to determine whether any of the anticipated covariates influences the estimate of FPA rate.

The primary endpoint will also be summarized as the median FPA rate, and the percentage of subjects with at least one FPA will be reported. If subjects have more than one FPA, the distribution of number of FPAs per subject may also be reported.

## 5.2 Secondary Endpoints

Secondary endpoints are exploratory and do not have specific performance criteria requirements.

## 5.3 General Methodology

Unless otherwise stated, continuous variables will be summarized by number of observations, mean, standard deviation, median, minimum and maximum values, and categorical/dichotomous variables will be summarized using frequency tables (frequencies, percentages, confidence intervals).

Hypothesis tests of continuous variables will be based on two sample t-tests. Non-parametric Wilcoxon signed rank test may be used when underlying data are found to be non-normal.

Hypothesis tests of categorical variables will be based on a Chi-square (or Fisher exact) test.

Level of significance (α), for purposes of statistical comparisons/tests, will be 0.05.

# 6. Safety Evaluation

## 6.1 Safety Endpoints

All adverse events that occur during the study will be recorded per the protocol definitions. Per the study protocol, adverse event (AE) is defined as any untoward medical occurrence in a subject during the study that in the opinion of the investigator is at least possibly-related to use of the ASSURE WCD System. For each AE, the Investigator will assess the causality/relationship to the study device (ASSURE™ WCD) and categorize the AE as mild, moderate, or severe, depending on the event's impact on the subject's daily activity level.

Adverse events occurring during the study will be tabulated by type, seriousness, severity, and relationship to the study device. The Medical Monitor assigned to this study will review the adverse events to determine whether any are reportable as UADEs. The Medical Monitor will also review serious adverse events to verify relatedness and the seriousness classification. The Medical Monitor's assessment of seriousness and relatedness, rather than the investigator's assessment, will be used for summarizing and analyzing safety data.

Any deaths or serious device associated AEs will be summarized. AEs leading to subject withdrawal from the study will be listed separately. In cases where the same event is reported more than once per subject, the event will only be counted once in the incidence tables. A listing of all adverse events will be provided.

# 7. Measurement Scales and Subject Reported Outcomes

## 7.1 BORG Comfort Scale

The BORG Comfort Scale consists of eight questions that ask the subject to check the number that best describes the level of discomfort the subject feels in each area of the body: upper right chest, upper left chest, lower right chest, lower left chest, upper left back, upper right back, lower left back and lower right back. The response options range from 0 (None at all) to 10 (Very, Very Severe). For the purpose of summarizing the BORG Comfort Scale, the data will be considered continuous and will be summarized at relevant visit intervals, including change from baseline.

## 7.2 Device Usability Survey

The subject will be asked to complete the Device Usability Survey at Visit 2 after the Garment is removed. The survey consists of ten questions and ask the subject to indicate the level of agreement or disagreement with each of the questions, based on five response options: Strongly Disagree, Disagree, Neutral, Agree and Strongly Agree. The responses to the ten questions are summed for each survey and then multiplied by 2.5 to convert the original scores of 0-40 to 0-100. The usability scores will then be analyzed as a continuous measure.

# 8. General Statistical Methods

## 8.1 Missing Data

All reasonable efforts will be made to ensure complete and accurate data are collected on all enrolled subjects as set forth in the protocol and CRFs. Subject WCD wear-time, and the associated heart monitoring during wear-time, however, do not lend themselves to imputation in the case where data are found to be missing, for any reason, and no efforts will be made to do so. Any extended periods of absence of WCD wear-time and/or and heart monitoring will be investigated for cause.

## 8.2 Analysis Sets

- The Intent-to-Treat (ITT) set of subjects will include the data from all enrolled subjects. No subjects will be excluded from this data set for any reason (e.g., protocol violation, study device not used, etc.).
- The Per Protocol (PP) set of subjects will include all subjects enrolled who met the eligibility criteria and were appropriately fitted with a WCD (correct Garment size based on the sizing chart in the ASSURE System Training Manual). The determination of subjects excluded from the PP set will be made prior to locking the final database.
- The Safety (SAF) set of subjects will include all subjects who used the device for at least two hours. No subjects (or data) will be excluded from this dataset because of protocol deviations that occur during the study.

The primary endpoint analysis will be based on the ITT dataset, although a secondary analysis may also be performed based upon the PP dataset (if there are differences), to assess the sensitivity of the analysis to the choice of analysis dataset. The analyses of all other endpoints will be based upon the ITT dataset only. All safety analyses will be based upon the Safety dataset.

## 8.3 Subject Disposition

Subject disposition including the number of subjects enrolled and number of subjects completing the study (and completing each study visit) will be tabulated. The percentage of subjects completing the study will be based on the total number enrolled.

Protocol deviations will also be summarized. Subject withdrawals and the reasons for withdrawal will also be summarized.

All data collected on the withdrawn subjects prior to withdrawal will be included in all analyses. However, if a protocol deviation is reported for a withdrawn subject, the subject may be excluded from the Per Protocol analysis set if the deviation is determined to potentially affect study outcomes.

## 8.4 Demographic and Other Baseline Characteristics and Concomitant Medications

Demographic data, medical history and concomitant medications will be summarized by means of descriptive statistics based upon the ITT dataset.

## 8.5 Control of Systematic Bias

Several measures are incorporated into the study design to help minimize bias as follows:

- This is a multi-center trial to help ensure that investigator or site or subject enrollment bias is minimized. Investigational sites are limited to enrolling no more than 18 subjects.
- Consecutively eligible subjects should be enrolled into the study to minimize selection bias.
- A Medical Monitor will review all assessments of seriousness and relatedness for adverse events, and the Medical Monitor's determinations will supersede the site-reported assessments in summaries of the data.

## 8.6 Pooling Data Across Centers

Approximately ten (10) investigational sites in the US will be participating in this study. The analyses will be presented using data pooled across study centers. Comparability between study sites may be examined using summary statistics calculated by site. A formal assessment of poolability of subjects across investigational sites is not planned.

# 9. Changes to Planned Analyses

Any changes to the planned statistical analyses made prior to analysis of the primary endpoint will be documented in an amended SAP approved prior to performing said analysis. Any unanticipated changes from the planned statistical methods after analyzing the primary endpoint data will be documented in the clinical study report along with the reason for deviation.

# 10. Definitions

### Alarm Definitions (for the Primary Endpoint)

When initial arrhythmia detection criteria are met, the ASSURE WCD opens an Episode and begins storage of ECG signals. If the rhythm is sustained for the confirmation period, a Shock Alarm Event Marker is recorded. Each Shock Alarm Event Marker will be reviewed and annotated per the definitions and instructions in the *Endpoint Review Committee (ERC) Charter for ASSURE WCD Clinical Evaluation - Detection and Safety Study (ACE-DETECT)* (3334131). Each Shock Alarm Event Marker will be classified for the primary endpoint as follows:

- **True Positive Alarm**: The presence of a WCD Shock Alarm Event Marker within a recorded Episode where the Shock Alarm has been annotated as:
    - Rhythm Type **VT/VF** with HR ≥ 153 bpm (170 bpm - 10%)
- **False Positive Alarm**: The presence of a WCD Shock Alarm Event Marker within a recorded Episode where the Shock Alarm has been annotated as:
    - Rhythm Type **Other**

      OR
    - Rhythm Type **Uncertain (presumed non-shockable)**

      OR
    - Rhythm Type **VT/VF** with HR < 153 bpm (170 bpm - 10%)

### Episode Detection (for the secondary Endpoint)

All WCD Episodes, and ICD Episodes for which there is not a corresponding WCD Episode, will be reviewed and annotated per the definitions and instructions in the *Endpoint Review Committee (ERC) Charter for ASSURE WCD Clinical Evaluation - Detection and Safety Study (ACE-DETECT)* (3334131). These annotation results will be used to classify the detections according to the following definitions:

- **True Positive Detection**: A WCD recorded Episode (with or without a Shock Alarm Event Marker) annotated as Rhythm Type **VT/VF** with **HR** ≥ 153 bpm (170 - 10%)
- **Missed Event (False Negative Detection)**: An Episode recorded by the ICD but not by the WCD when it is confirmed that the WCD was being worn, and where the ICD Episode is annotated as Rhythm Type **VT/VF** with **HR** ≥ 187 bpm (170 + 10%) and Duration ≥ 20 seconds.

# 11. Revision History

| Revision | Description of Changes | Revision Date |
|---|---|---|
| 1 | Initial description of analyses | 19 MAR 2019 |

# APPENDIX A: Sample Episode Summary

| | | WCD | | | | | | | ICD | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Subject ID | Date | ID | Open Marker time | Shock Alarm Marker time | Rhythm Type | Heart Rate | Annotation | Annotation BPM if Uncertain | ID | Time | Rhythm Type | Heart Rate (BPM) | Duration | VT/VF > 187 BPM for > 20 sec? | Duration | Therapy |
| | | | | | | | | | | | | | | | | |