The Tailored-AF trial

Tailored vs Anatomical Ablation Strategy for Persistent **A**trial **F**ibrillation

NCT Number: NCT04702451

Protocol ID: CLIPL-01-002

Document type : Statistical Analysis Plan

Document version: 1.4

Document date: 19/FEB/2024

Volta Medical

Tailored vs. Anatomical Ablation Strategy for Persistent AF
CLIPL-01-002-E

Statistical Analysis Plan

Version: 1.4

Volta Medical
65 Avenue Jules Cantini
13006 Marseille, FRANCE

TABLE OF CONTENTS

1 PROTOCOL SUMMARY 3
2 ABBREVIATIONS 5
3 INTRODUCTION 6
4 ENDPOINT ANALYSES 6
4.1 ANALYSIS SETS 6
4.2 PRIMARY ENDPOINT 7
4.3 SECONDARY ENDPOINTS 8
4.4 EXPLORATORY ENDPOINTS 10
4.5 SAFETY ENDPOINTS 11
5 ASSESSING THE PRIMARY OUTCOME 13
6 GENERAL STATISTICAL CONSIDERATIONS 14
6.1 INTERIM ANALYSIS 14
6.2 BLANKING PERIODS 15
6.3 MISSING DATA MANAGEMENT 16
6.4 POOLING DATA ACROSS CENTERS 16
6.5 RANDOMIZATION 16
6.6 TREATMENT MASKING 16
6.7 STUDY STOPPING RULES 17
6.8 MEASURES TAKEN TO MINIMIZE BIAS 17
6.9 CORRECTION FOR MULTIPLE-TESTING 17
7 CHANGES FROM PROTOCOL 17
8 REFERENCES 18

1 Protocol Summary

Primary Objective	To demonstrate that a tailored ablation strategy targeting areas exhibiting spatio-temporal dispersion in association with PVI, is superior to an anatomical probabilistic ablation strategy targeting PVI alone for the treatment of persistent AF.
Study Device/Procedure	Tailored ablation with VX1
Control Device/Procedure	Anatomical Ablation
Study Design	Interventional, prospective, randomized, controlled, two-arm, multicenter clinical investigation
Primary Endpoint(s)	Freedom from documented AF, with or without AADs, 12 months after a single AF ablation procedure. Where freedom from documented AF is defined as no documented episodes of AF>30 seconds with conventional non-invasive monitoring between 3 and 12 months after index procedure. Freedom from AF is analyzed as a time-to-event comparison between survival distributions of the treatment and control arms.
Secondary Endpoint(s)	1. Freedom from documented AF/AT, after one or two procedures, with or without AADs, at 12 months 2. Freedom from documented AF/AT, after one procedure, with or without AADs, at 12 months 3. Composite safety endpoint of death, cerebrovascular event, or serious treatment-related adverse event at 12 months after index procedure.

Exploratory Endpoints	1. Freedom from documented AF, after one or two procedures, with or without AADs, at 12 months 2. Freedom from documented AF or AF/AT, after one procedure, or one procedure or more, without AADs, at 12 months 3. Freedom from documented AF or AF/AT, after one procedure, or one procedure or more, without AADs or with previously failed AAD at the same dose, at 12 months 4. Freedom from documented AF or AF/AT, after one procedure, or one procedure or more, with or without AADs, at 12 months, without the use of TTM 5. Freedom from documented symptomatic AF or AF/AT, after one procedure, or one procedure or more, with or without AADs, at 12 months 6. Freedom from AF or AF/AT, with or without AADs, 12 months after one procedure, or one procedure or more, for subjects with persistent AF <1 year 7. Quality of life measurements (SF-36 and AFEQT) at baseline, 3, 6 and 12 months post-index procedure 8. Incidence of peri-procedural complications including stroke, PV stenosis, cardiac perforation, esophageal injury, and death 9. Acute AF termination rate by ablation 10. Rate of conversion to sinus rhythm by ablation 11. Ablation procedure duration 12. Fluoroscopy time and dose 13. Mapping time 14. Radiofrequency (RF) time to terminate AF 15. Total RF time 16. Total delivered energy (mean power x RF duration) 17. Total delivered energy (mean power x RF duration) / biatrial surface 18. Ablation surface area to terminate AF 19. Total ablation surface area 20. Blinded baseline Voltage maps before ablation 21. TTM compliance (% of transmissions expected vs received at each week post index procedure) 22. Assessment of blind maintenance per group
Follow-Up Schedule	Follow-up assessments at 3, 6, and 12 months post index procedure. Holter and 12-lead ECGs will be performed at 3, 6, and 12 months. Portable trans telephonic event monitors will be recorded weekly for approximately 52 weeks.

2 Abbreviations

AAD	Antiarrhythmic Drug
ADE	Adverse Device Effect
AE	Adverse Event
AF	Atrial Fibrillation
AFEQT	AF Effect on Quality-of-Life Questionnaire
AT	Atrial Tachycardia
CP	Conditional Power
ECG	Electrocardiogram
HR	Hazard Ratio
mITT	Modified Intention-To-Treat
PP	Per-Protocol
PV	Pulmonary Vein
PVI	Pulmonary Vein Isolation
QoL	Quality-of-Life
RF	Radio Frequency
SADE	Serious Adverse Device Effect
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SF-36	The Short Form (36) Health Survey
TTM	Trans Telephonic Monitoring

3 Introduction

This statistical analysis plan (SAP) describes the planned analyses for data collected under the Clinical Trial Protocol CLIPL-01-002 “Tailored vs. Anatomical Ablation Strategy for Persistent AF”, last revised 06 July 2022 (Revision E). Specified analyses may be used for scientific presentations and/or manuscripts, and regulatory submissions. The primary analysis will be based on the data through 12 months post-index procedure.

If there are any discrepancies between this SAP and the study protocol, the SAP shall prevail.

4 Endpoint Analyses

4.1 ANALYSIS SETS

The analysis sets making up this study will include the following:

4.1.1 The Safety Population

The Safety population will include all randomized Subjects who underwent their first ablation procedure.

4.1.2 The Modified Intent-to-Treat Population (mITT)

The modified Intention-To-Treat (mITT) set will include all randomized subjects except for:

- Subjects who are deemed ineligible after randomization or who do not have any AF ablation, and
- Subjects lost to follow-up during the 3-month blanking period.

Subjects in the Tailored group for whom AF cannot be induced and who undergo conventional ablation will be included in the mITT analysis.

Subjects who complete at least their 3-month follow-up visit but who are lost to follow-up before the 12-month follow-up visit will be included in the mITT analysis.

Subjects who undergo a 3rd ablation procedure will be included in the analyses but the data will be censored at the time of the 3rd procedure for multiple procedure endpoints.

4.1.3 The Per-Protocol Population (PP)

The Per-Protocol (PP) analysis set will include only those subjects of either given treatment arm defined as the treatment actually received at the start of the study, and will consist of subjects who met all the inclusion criteria, none of the exclusion criteria, were successfully treated with an index procedure, had no important protocol deviations, and attended the scheduled follow-up visits (did not miss more than one mandatory Holter recording, and TTM compliance >50%).

The PP analysis set will also be used to evaluate endpoints already evaluated using the mITT as the primary analysis set, thus serving as supportive analyses to that of the mITT efficacy analysis.

4.2 PRIMARY ENDPOINT

Freedom from documented AF, 12 months after a single index ablation procedure to be understood as a single AF ablation procedure (i.e. including redo procedures for non-AF indications).

4.2.1 Hypotheses

Freedom from documented AF after 12 months will be higher in the tailored group than the anatomic group.

The null and alternative hypotheses are as follows:

$$H_0: S_1(t) = S_2(t)$$

$$H_A: S_1(t) > S_2(t),$$

where S1(t) and S2(t) represent the survival functions for the Tailored and Anatomical arms, respectively, at any given time t.

4.2.2 Sample Size

The primary endpoint is freedom from AF, 12 months after a single ablation procedure. Assuming that 62% of the Subjects in the anatomic group will be free from AF at the end of follow-up1,2, that at least 77% of the Subjects of the tailored group will be free from AF at the end of follow-up (i.e., Hazard Ratio (HR) 0.547), then 292 Subjects (146 in each group) are required to show that time to AF is significantly different between groups (Log-Rank test for a one-sided superiority trial, significance level α=0.025, statistical power of 80%, under the assumption that the hazard rates are proportional). Assuming a drop-out rate of 22% (no index ablation performed or loss to follow up), 374 Subjects are required overall (187 in each group). Following 50% of subjects having a primary endpoint collected, the sample size will be re-estimated using the Mehta-Pocock conditional power (CP) approach.

4.2.3 Primary Analysis

The primary endpoint will be analyzed using a Cox Proportional hazards model to conduct a time-to-event comparison (of hazard rates from risk of AF recurrence at a time t) between the Tailored (treatment) and Anatomical (control) groups. The model will include an indicator for group (study arm), a dichotomous effect for AF duration (persistent AF ≥ 6 months; persistent AF < 6 months), and applying right-censoring of subjects lost to follow-up. The model will be used to estimate the hazard ratio of AF occurrence between the Tailored arm compared to that of the Anatomical arm, along with a two-sided 95% CI and p-value for the HR on behalf of the group predictor. The actual null hypothesis of the Primary Endpoint will

be tested with a one-sided Log-Rank test at a significance level of α=0.025. The null hypothesis will be rejected if the p-value is < 0.025, and the survival distribution of the Tailored arm will be considered superior relative to that of the Anatomical arm. All of the outputs described above will be generated under the same single Cox Proportional hazards model, which will be conducted in the mITT population. A Kaplan-Meir plot will be generated to visually display the survival distributions of the Tailored and Anatomical arms, respectively to accommodate the p-value from their comparison.

4.2.4 Sensitivity Analyses of Primary Endpoint

The Primary Endpoint Analysis will additionally be conducted in the PP population.

4.2.5 Other analyses on primary endpoint

Primary analyses will be repeated with the following changes in the Cox Proportional Hazards model:

Without AF duration fixed effect. With AF duration fixed effect numerically expressed as the logarithm of the AF duration (in months) instead of the categorical definition. The format of the AF duration variable (dichotomous vs. numerically expressed as the log of months) will be compared between the models with which each is fit.

4.2.6 Subgroup Analyses

The following subgroup analyses will be performed following the approach described for the primary analysis:

1. All subjects with AF duration within the following categories: < 6 months, ≥ 6 months, < 12 months, and ≥ 12 months
2. AF presentation at index procedure (spontaneous AF, induced AF)
3. With/without AADs
4. Patient gender (male and female)

4.3 SECONDARY ENDPOINTS

The secondary endpoints are as follow:

1. Freedom from AF/AT after one or two procedures at 12 months following index procedure.
2. Freedom from AF/AT after a single procedure at 12 months following index procedure.
3. The safety endpoint is a composite endpoint, the events of which it is comprised include that of death, cerebrovascular event, or serious treatment-related adverse event at the 12-month timepoint since initial index procedure (this endpoint is inclusive of the Blanking Period – marked by the 3 months immediately following the initial index procedure undergone by each enrolled subject).

4.3.1 Secondary Analysis

The secondary endpoint analyses will consist of the following:

Two tests of the primary endpoint in the mITT population, but expanded to be defined as freedom from AF/AT, with or without AADs, 12 months after one or two procedures. A superiority hypothesis will be tested using a Cox proportional hazards model with randomization group, AF duration, and AF/AT strata as fixed effects, with a one-sided alpha level of 0.025:

Both analyses under will be evaluated according to the following hypotheses:

$$H_0: S_1(t) = S_2(t)$$

$$H_A: S_1(t) > S_2(t),$$

where $S_1(t)$ and $S_2(t)$ represent the survival functions for the Tailored and Anatomical arms, respectively, at any given time *t*. If the null hypotheses for each of the tests is rejected, the following superiority statements will be considered on behalf of each endpoint:

1. Tailored ablation is superior to anatomical with respect to freedom from AF/AT, with or without AADs, 12 months after one or two procedures.

2. Tailored ablation is superior to anatomical with respect to freedom from AF/AT, with or without AADs, 12 months after one procedure.

Similarly to the primary endpoint analyses, Cox Proportional hazards regression models will be fit and assessed for goodness-of-fit. Then to test each null hypothesis, one-sided Log-Rank Tests will be. Just like under the primary analysis, the p-value from the one-sided Log-Rank Test run under each Cox Proportional hazards model will ultimately be what is used to either reject ($p < 0.025$) or fail to reject ($p \geq 0.025$) the null hypotheses above. Both Cox models will be conducted in the mITT population.

Kaplan-Meier plots will be generated for each endpoint, to visually display the differences among the survival distributions of the Tailored and Anatomical arms, respectively, with each plot accommodated by the p-value generated from their respective one-sided Log-Rank test.

3. The composite adverse event (AE) rate defined as the number of AEs per month occurring in each treatment arm from time of index procedure through to 12-month follow-up, where composite AEs are defined by either death, a cerebrovascular event, or any serious procedure-related adverse event, will be conducted. in the Safety population (see Section 4.5, Safety Endpoints below).

4.3.2 Sensitivity Analyses of Secondary Survival Endpoints

The first two Secondary Endpoint Analyses will additionally be conducted in the PP population.

4.4 EXPLORATORY ENDPOINTS

The following endpoints compared in anatomical vs. tailored groups will be considered exploratory:

1. Freedom from documented AF, after one or two procedures, with or without AADs, at 12 months
2. Freedom from documented AF or AF/AT, after one procedure, or one procedure or more, without AADs, at 12 months
3. Freedom from documented AF or AF/AT, after one procedure, or one procedure or more, without AADs or with previously failed AAD at the same dose, at 12 months
4. Freedom from documented AF or AF/AT, after one procedure, or one procedure or more, with or without AADs, at 12 months, without the use of TTM
5. Freedom from documented symptomatic AF or AF/AT, after one procedure, or one procedure or more, with or without AADs, at 12 months
6. Freedom from AF or AF/AT, with or without AADs, 12 months after one procedure, or one procedure or more, for subjects with persistent AF <1 year.
7. Freedom from AF or AF/AT, with or without AADs, 12 months after one procedure, or one procedure or more, for subjects with AF ≥ 6 months (including long-standing persistent AF)
8. Estimated AF or AF/AT burden, with or without AADs, after one procedure, or one procedure or more
9. Quality of life measurements (SF-36 and AFEQT) at baseline, 3, 6 and 12 months post-index procedure
10. Incidence of peri-procedural complications including stroke, PV stenosis, cardiac perforation, esophageal injury, and death
11. Acute AF termination rate by ablation
12. Rate of conversion to sinus rhythm by ablation
13. Ablation procedure duration
14. Fluoroscopy time and dose
15. Mapping time
16. Radiofrequency (RF) time to terminate AF
17. Total RF time
18. Total delivered energy (mean power x RF duration)
19. Total delivered energy (mean power x RF duration) / biatrial surface
20. Ablation surface area to terminate AF
21. Total ablation surface area
22. Blinded baseline Voltage maps before ablation
23. TTM compliance (% of transmissions expected vs received at each week post index procedure)

4.4.1 Exploratory Endpoint Analysis

Endpoints 1 – 8 will be tested using a Cox proportional hazards model with randomization group and AF strata as fixed effects. Hypotheses will be tested in the following order, and each at a less stringent one-sided alpha level of 0.05, given the exploratory nature of these analyses)

Endpoint 9 will be tested by comparing the QoL score at each visit between tailored and anatomical ablation using a t-test for difference in means. A Shapiro-Wilks test will be used to assess whether the data meets normality assumptions for parametric. Should the normality assumption be violated, a non-parametric method such as the Wilcoxon test will be used instead. The effect of time will also be assessed using a generalized linear mixed model for repeated measures.

Endpoints 10 – 21 and 23 will be tested by comparing tailored and anatomical ablation groups. using a t-test for differences in means. Should the normality assumption be violated, a non-parametric method such as the Wilcoxon test will be used instead.

Endpoint 22 will not undergo any statistical analysis, but the data collected (blinded baseline voltage maps prior to ablation) will be provided to the sponsor in a dataset / spreadsheet format to enable any post-hoc analysis that may be desired.

4.5 SAFETY ENDPOINTS

The safety endpoint is a composite endpoint at 12 months of the following: death, cerebrovascular events, or serious treatment-related adverse event. Because the rates of this endpoint which are observed under each of the study arms may or may not prove to be relatively rare (< 5% within a given arm), two methods of analysis are provided, the one which will ultimately be used to analyze the safety endpoint will depend on its rate of occurrence, as observed across the two study groups. The safety endpoint analysis, regardless of the method employed, will be conducted in the Safety population. The same analysis will also be conducted in the mITT and PP populations, since if rare occurrence of the safety endpoint is observed and the few occurrences observed happen to all occur within the mITT/PP populations, then the smaller subject sample sizes will serve to increase the rates/proportions observed in each study arm.

METHOD 1 (safety endpoint occurrence > 5% in each study arm):

The safety analysis will be to compare the AE composite rate per month (defined as the 4th secondary endpoint in Section 1 Protocol Summary) under the Tailored (treatment) arm against the Anatomical (control) arm AE composite rate per month from the time of index procedure until the 12-month follow-up. A two-sided 95% CI will be generated for the rate ratio (λ_t/λ_c) comparing the AE composite rate for the two study arms.

Let λ_c and λ_t denote the rate of composite adverse events for the Anatomical and the Tailored treatment arms, respectively. Given a non-inferiority margin δ, the null and alternative hypotheses are:

$$H_0: \frac{\lambda_t}{\lambda_c} \geq 1 + \delta$$

$$H_1: \frac{\lambda_t}{\lambda_c} < 1 + \delta$$

$\lambda_t, \lambda_c,$ and $\frac{\lambda_t}{\lambda_c}$ will be estimated with a linear fixed effects model with a log link function and a Poisson distribution for the number of adverse events (referring only to those defined under the composite AE endpoint). The outcome of the model is the number of composite AEs for each subject enrolled and randomized to one of the study arms. The model will

include an indicator variable for Study Arm, an offset for the log of the number of weeks each subject remained enrolled in the study will be added (the date of each subject's index procedure will be regarded as their first day on the study).In the event the model fails to converge, it will be refit with a Negative Binomial distribution.

H_0 is rejected if the upper limit of the 95% confidence interval for $\frac{\lambda_t}{\lambda_c}$ is less than $1 + \delta$. If H_0 is rejected, non-inferiority is demonstrated. Based upon the number of months and the expected rate of adverse events, the current study design is under-powered to detect a clinically meaningful (less than 5%) non-inferiority margin for the Tailored treatment arm. Because the null hypothesis test must be conducted under a specified non-inferiority margin, the analysis will use $\delta = 0.05$.

METHOD 2 (safety endpoint occurrence (≤ 5% in each study arm):

If both model approaches fail to converge, most likely due to the incidence of the composite AE rates being relatively rare (less than or equal to 5%) under both study arms, then the composite AE endpoint will be treated as a frequency instead of a rate, such that P_1 will be used to denote the proportion defined by the number of composite AEs out of total subjects in the Tailored (treatment) arm and P_2 will be used to denote the proportion defined by the number of composite AEs out of total subjects in the Anatomical (control) arm. The difference in proportions ($P_1 - P_2$) will be reported along with a two-sided exact 95% CI using an exact test statistic with continuity correction if not already defaulted (a Fisher's Exact test may be used if CI can be generated from its statistic for the difference of proportions estimate). Given a non-inferiority margin δ, the null and alternative hypotheses are as follow:

$$H_0: p_1 - p_2 \geq \delta$$
$$H_1: p_1 - p_2 < \delta$$

If the upper bound of the 90% CI for ($P_1 - P_2$) is less than $\delta = 0.025$, the null hypothesis is rejected, and non-inferiority is demonstrated (2.5% has been allocated as a clinically meaningful non-inferiority margin in the absence of literature). In the event that the composite AEs have a zero occurrence in both study arms, a confidence interval for the difference of proportions cannot be computed, in which case 95% exact CIs for P_1 and P_2 will be reported, and non-inferiority will be considered demonstrated.

4.5.1 Other Adverse Events

All adverse events will be summarized through 12 months. The proportions of subjects with events will be reported. No imputation for missing data is planned.

Tabular summaries by treatment group of all adverse events, adverse device effects, serious adverse events, and serious adverse device effects will be presented.

An adverse event (AE) is defined as any untoward medical occurrence, unintended disease or injury or untoward clinical signs (including an abnormal laboratory finding) in subjects,

users or other persons, whether or not related to the investigational medical device. An Adverse Device Effect (ADE) is defined as an adverse event related to the use of an investigational medical device. A serious adverse event (SAE) is an adverse event that:

a) led to a death

b) led to a serious deterioration in the health of the subject, that either resulted in: - a life-threatening illness or injury, or - a permanent impairment of a body structure or a body function, or - in-patient or prolonged hospitalization, or -medical or surgical intervention to prevent life threatening illness or injury or permanent impairment to a body structure or a body function.

c) led to fetal distress, fetal death or a congenital abnormality or birth defect. A serious adverse device effect (SADE) is an adverse device effect that has resulted in any of the consequences characteristic of a serious adverse event.

The investigator will assess each adverse event for its relationship to the study device, and whether it was anticipated in the protocol. The determination of the relationship between the adverse event and the device is made according to the following definitions:

- Definitely related (causal relationship)
- Probably related
- Possibly related
- Not related
- Unknown.

Acute peri-procedural complications (related to the study procedure and occurring within one week of the ablation procedure), will be specifically analyzed as part of the exploratory endpoints.

5 Assessing the Primary Outcome

Freedom from documented AF between 3 and 12 months following the index ablation procedure will be defined as no documented episodes of AF > 30 seconds with conventional non-invasive monitoring.

The following scenarios result in primary endpoint failure:

1) A repeat procedure where AF is confirmed during the procedure (Patient is in spontaneous AF at the beginning of the procedure)

2) A 30-second or greater episode of AF is documented on a 24-hour Holter

3) A 30-second of AF on a weekly TTM transmission confirmed by 30-second of AF on same (but different day) or next week TTM transmission. Two AF transmissions on the same day do not count as a confirmed recurrence.

4) Since 12-Lead ECGs are 10 seconds in duration, an AF episode of >30 seconds will be documented via 12-lead EGC under only certain conditions (AF symptoms, multiple ECGs recorded at a single visit showing AF, confirmation of single Kardia in AF in same, previous or next week).

TTM will be measured weekly by the patient via a portable ECG device (KardiaMobile, AliveCor). The recordings will be automatically uploaded for blind review by the ECG core lab.

24-hour Holter monitor assessments will be performed prior to or subsequent to each in-office visit at 3, 6, and 12 months post-index procedure.

Diagnostic 12-lead ECGs will be performed at each in-office visit at 3, 6, and 12 months following the index procedure.

6 General Statistical Considerations

6.1 INTERIM ANALYSIS

This trial will employ an adaptive design for the blinded re-estimation of the sample size. To ensure the study is adequately powered for the superiority comparison for the primary outcome measure, the assumptions regarding the rates of AF in both study arms will be verified after 50% of the originally planned total patient sample have complete primary endpoint information. A conditional power (CP) approach for sample-size re-estimation using the Mehta-Pocock approach will be utilized.[3] Following this interim assessment, the sample size may only remain the same or be increased. A third-party, independent statistician will perform this analysis to minimize operational bias. This analysis will occur after approximately 50% of the originally planned subjects a completed/analyzable primary endpoint. The statistician will assess the conditional study power assuming that the hazard ratio remains the value estimated at this analysis. The conditional power will be categorized into one of the three following zones:

Favorable: CP ≥ 80%
Promising: 36% ≤ CP < 80%
Unfavorable: CP < 36%

The lower bound on the promising results is selected from Mehta and Pocock as a function of the target power, timing of the sample size re-estimation, and maximum sample size. Selection of this boundary in particular does not inflate the overall Type I error rate.
After the analysis, the statistician will provide a recommendation to the Steering Committee and Volta Medical clinical team:

- In the event that the conditional power is found to be in the Unfavorable zone, the study will continue to the originally planned sample. The independent statistician will report ONLY that the trial will continue as originally planned.
- In the event that the conditional power assessment is Promising, the sample size will be increased to 100 total subjects (an additional 50 maximum per treatment arm, predetermined by Volta). The independent statistician will report that the sample size should be increased and by how many subjects (in regard to the 292 subjects originally planned).
- In the event that the conditional power is found to be in the Favorable zone, the study will continue to the originally planned sample. The independent statistician will report ONLY that the trial will continue as originally planned.

Also, the following information will be shared with senior, non-operational management staff that have a strict business need to review and act on the results:

- The recommendation (sample size increase or continue as originally planned)

- The rate of freedom from AF for each arm for the overall interim analysis population

- The rate of freedom from AF for each arm for two subgroups depending on the AF type (persistent AF and long-standing persistent AF)

None of the clinical operational staff or study investigators will have access to the interim data on which the analysis is based to further minimize operational bias. The scientific soundness of the study is not likely to be impacted as the only potential change to the study as a result of this analysis is an increase in sample size.

6.2 BLANKING PERIODS

After the index procedure, a 3-month blanking period + tolerance of 1 week (98 days) will be applied for all efficacy endpoints, during which recurrences of AF/AT (excluding repeat procedure) will not be counted.

If a repeat procedure is performed during the 3-month blanking period (and patient is in spontaneous AF), the recurrence and the repeat procedure will be counted as occurring at 3 months.

If a repeat procedure is performed after the end of the 9th month, the procedure will not be considered for multiple procedures endpoints.

After a repeat procedure, a 2-week blanking period (14 days) will be applied, during which AF/AT recurrences will not be counted. Any re-repeat procedures (repeat procedure #2) will result in the data being censored at the time of the re-repeat procedure.

Note: for the purpose of counting months to establish blanking period boundaries in the study, one month will be considered equal to 30.4 days.

6.3 MISSING DATA MANAGEMENT

Subjects who miss a study visit should be contacted immediately by the Investigational site to determine the reason for the missed visit and to reschedule the visit as soon as possible to meet the study visit window. If the subject cannot be located after 3 attempts by phone, email and/or letter, then the subject will be considered lost to follow-up. All attempts to contact the subjects will be documented and retained in the subject study record. Regarding the main judgement criteria, data will be censored if the patient is lost to follow-up. As the burden of study participation does not differ between the tailored and the anatomical arm, we expect all study losses to be non-differential with respect to treatment assignment. Thus, our results will not be biased because of these losses. Additional supportive analyses will be performed on the primary endpoint to assess the effect of missing data. For missing demographic data, the data will not be replaced because it has no weight on the main judgment criterion. A rigorous monitoring is expected; any missing data will be filled in during these visits.

6.4 POOLING DATA ACROSS CENTERS

The analyses will be presented using data pooled across study sites. Summary statistics for the rate of freedom from AF will be calculated by site. A formal assessment of site-to-site heterogeneity will be assessed using a linear mixed-effects binomial regression model with a logit link function for the event of AF recurrence (corresponding to a difference of proportions analysis). The outcome of the model will be the binary status of AF recurrence per subject (0=Subject did not experience AF recurrence, 1=AF recurrence documented on behalf of subject). The model will include an indicator variable for study arm, AF duration as a fixed effect, site as a fixed effect, and a study arm-by-site interaction term. Within the model results, if the study arm-by-site interaction coefficient has a *p*-value <0.15, it will be regarded as there being heterogeneity detected between sites. Lastly, poolability by geography (US vs. OUS) will be performed based on linear mixed-effects binomial regression as described above. As US subjects will only be patients with AF duration of 3 months to < 12 months, only OUS patients with AF <12 months will be included in the poolability by country assessment.

6.5 RANDOMIZATION

After successful enrollment, the patient will be automatically randomized in a 1:1 ratio to one of two arms through the eCRF system: “Anatomical” arm (PVI) or “Tailored” arm (Dispersion + smart-PVI). Randomization will be stratified in order to balance the treatment arms in terms of subjects with persistent AF (< 12 months) and long-standing persistent AF (≥ 12 months). Subjects will not be informed of their randomization assignment.

6.6 TREATMENT MASKING

Special care will be taken to ensure that the patient will be blinded to his/her randomization assignment throughout the study. Patient unblinding will be documented as a non-important deviation. Moreover, a blind rhythm monitoring is planned: all ECG recordings for long-term effectiveness endpoints (Holter, TTM and 12-lead ECG) will be analyzed and adjudicated blindly by an independent ECG core lab. Subjects will be asked at the 12-month visit to guess which arm of the study they are in.

6.7 STUDY STOPPING RULES

Premature termination of the study may occur motivated by a general decision (Ethics Committees, Institutional Review Boards, regulatory Competent Authorities, Data Safety Monitoring Board, Steering Committee) for safety reasons related to the use of the device or study procedures.

6.8 MEASURES TAKEN TO MINIMIZE BIAS

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

1) This is a multi-center trial to help ensure that investigator or site or subject enrollment bias is minimized. Selection of subjects will be made from the Investigator's usual subject load. Consecutively eligible subjects should be enrolled into the study.
2) This document specifies appropriate statistical methods to ensure that bias is minimized.
3) Standardized and validated case report forms will be used to collect data during the study.
4) Steps to ensure blinding to treatment received have been taken (see Section 6.6)

6.9 CORRECTION FOR MULTIPLE-TESTING

For the two secondary efficacy endpoints, the same variable will be observed after up to two procedures (one or two procedures), or after a single procedure. Given the potential number of patients receiving a second ablation procedure (between 22% and 33%[1]), the variable for both secondary efficacy endpoints will be the same for most patients, meaning that the two variables will be highly correlated. The third secondary endpoint relates to safety, and is therefore of a different nature. Consequently, no correction for multiple testing will be applied for secondary endpoints.

7 Changes from protocol

Primary endpoint hypothesis in the protocol was wrongly based on the effect size observed in a previous study and used in the sample size estimation (delta 15%, HR = 0.547). The study is powered to detect such an effect size. The tested hypotheses were corrected to H0: S1(t) = S2(t) and HA: S1(t) > S2(t).

The gatekeeping approach for secondary endpoints was abandoned. There are 3 secondary endpoints of different nature (2 highly-correlated efficacy, 1 safety) and therefore no multiple-testing correction will be applied.

8 References

1. Verma, A. et al. “Approaches to catheter ablation for persistent atrial fibrillation.” *New England Journal of Medicine* 372 (2015): 1812–1822.

2. Fink, T. et al. “Stand-alone pulmonary vein isolation versus pulmonary vein isolation with additional substrate modification as index ablation procedures in patients with persistent and long-standing persistent atrial fibrillation: the randomized alster-lost-AF Trial (ablation at St. Georg hospital for long-standing persistent atrial fibrillation).” *Circ Arrhythm Electrophysiol* 10, (2017).

3. Mehta, Cyrus R., and Stuart J. Pocock. "Adaptive increase in sample size when interim results are promising: a practical guide with examples." *Statistics in medicine* 30.28 (2011): 3267-3284.