Medtronic

Alleviating Pain · Restoring Health · Extending Life

AdaptResponse

Statistical Analysis Plan

Version 3
15 NOV 2021

Project Statistician:
[redacted]

TABLE OF CONTENTS

1 VERSION HISTORY

Version	Version Date	Summary of Changes	Author(s)/Role
1.0	*30 SEP 2015*	• *Not Applicable, New Document*	[redacted], *Project Statistician*
2.0	*29 APR 2016*	• *'Version History' is added.* • *For secondary objective #2 (intervention for heart failure decompensation) and #4 (incidence of atrial fibrillation), the primary analysis in the presence of competing risk is changed into the Cox regression while the Fine-Gray regression becomes a secondary analysis. The subsequent sensitivity/subgroup/ancillary analyses are changed as well into the framework of Cox regression.*	[redacted], *Project Statistician*
3.0	*15 NOV2021*	• *Removed conditional disengagement date from calculation of follow-up time, as it is an administrative date and does not reflect patient contact* • *For frailty models used in analysis of primary objective and secondary objective #1, added example code to specify site-level random effects from a gamma distribution, rather than the default lognormal*	[redacted], *Project Statistician*

2 PURPOSE

This Statistical Analysis Plan (SAP) has been designed to document, before data are analyzed, the rationale for the study design, and the planned analyses that will be included in study reports. This SAP does not limit the analysis in reports, and additional analysis of the study data beyond this plan is expected.

This SAP is developed based on Version 1 of the AdaptResponse Clinical Investigation Plan (CIP) dated April 1, 2014.

3 RATIONALE FOR STUDY DESIGN

The study design and rationale can be found in the CIP. AdaptResponse is a prospective, randomized, controlled, interventional, single-blinded, multi-center, post-market, global Cardiac Resynchronization Therapy (CRT) in heart failure (HF) clinical study. The purpose of this clinical study is to test the hypothesis that market released CRT devices which contain the AdaptivCRT® (aCRT) algorithm have a superior outcome compared to standard CRT devices in CRT indicated patients with normal atrio-ventricular (AV) conduction and left bundle branch block (LBBB).

Following enrollment and baseline assessment, eligible subjects will be implanted with a CRT system containing the aCRT algorithm and randomized in a 1:1 fashion to either treatment (aCRT ON) or control (aCRT OFF) groups. **aCRT ON** means CRT therapy with ambulatory optimization and preferential LV-only pacing; **aCRT OFF** means standard bi-ventricular CRT therapy. The randomization will be stratified by two factors: site and NYHA class at enrollment. The primary objective is to compare the incidence of the combined endpoint of all-cause mortality and intervention for heart failure decompensation between the aCRT ON and aCRT OFF groups. Because the event rate is expected to be relatively low, the study utilizes an "event-driven" design that requires a review of accumulating endpoint events during the trial to possibly modify enrollment goal and follow-up duration in order to maintain study power. Study subjects will be followed until the required number of endpoint events is reached and/or a decision by Medtronic, Ethics Committee, or regulatory authority, whichever occurs first.

4 DESCRIPTION OF ANALYSIS

4.1 General

Medtronic employed statisticians will perform the statistical analyses described in this SAP. Interim analyses will be performed by a Medtronic statistician other than the lead statistician for the study. The lead study statistician will be blinded to all interim analyses.

All tests of treatment effects will be conducted in order to preserve an overall two-sided alpha level of 0.05 unless otherwise stated. An Intention-to-Treat analysis will be performed and will serve as the primary analysis for all objectives in this study. The Intention-to-Treat cohort will include all randomized subjects.

4.1.1 Subject Characteristics at Enrollment and Baseline

A number of subject characteristics recorded on the enrollment and baseline case report forms (CRFs) will be summarized in tables for all the study subjects. **Table** 1 lists these characteristics:

Table 1: Subject characteristics in enrollment and baseline forms

CRFs	Subject Characteristics
Enrollment	Gender PR interval QRS duration Left ventricular ejection fraction (LVEF) NYHA class
Baseline	Age Physical exam status Cardiac disease classification Current HF symptoms status Blood measurements Medical history

For continuous variables, descriptive statistics including mean, standard deviation, median, the first quartile and third quartile, minimum and maximum will be presented; for categorical variables, counts and percentages will be given. Subject characteristics will also be summarized by treatment (aCRT ON and aCRT OFF) and NYHA class at enrollment (II, III and IV).

4.1.2 Special Considerations

4.1.2.1 Analysis Blinding

It is planned that Medtronic employed statisticians will perform all statistical analyses. Interim analyses will be performed by the unblinded Medtronic statistician of the study. The lead study statistician will be blinded to all interim analyses. The unblinded statistician will keep results strictly confidential per the DMC charter during the study.

4.1.2.2 ECG Core Lab vs. Site-Determined ECG Data

Per CIP section 1.3, the ECG Core Lab will review all baseline ECGs for presence of LBBB and normal AV conduction, and feedback on accuracy rates will be presented to the Steering Committee and to the sites. However, analyses that involve PR interval, QRS duration and LBBB diagnosis at enrollment will be based on the site-determined values, except for the sensitivity analyses of the primary objective where presence of LBBB needs to be confirmed by the ECG Core Lab.

4.1.2.3 Intention-to-treat and Sensitivity Analyses

An intention-to-treat (ITT) analysis will be performed and will serve as the primary analysis for all the objectives in this study. Based on the ITT principle, this analysis will include all randomized patients in the treatment groups to which they were randomly assigned, regardless of their adherence with the entry criteria, regardless of the treatment they actually received, and regardless of subsequent withdrawal from treatment or deviation from the protocol[1]. With this being said, patients with noncompliance of randomization including permanent crossover to the other study arm will stay in the treatment arm they were originally randomized to in the primary ITT analysis. A permanent crossover means the programming will not be corrected to the original randomization assignment (CIP section 5.5).

In addition, two sensitivity analyses will be performed for the primary objective:

1. ITT analysis for randomized patients whose presence of LBBB is confirmed by the ECG Core Lab. In this sensitivity analysis, patients with noncompliance of randomization including permanent crossovers will stay in the treatment arm they were originally randomized to per the ITT principle.
2. As-treated analysis for randomized patients whose presence of LBBB is confirmed by the ECG Core Lab. In this sensitivity analysis, patients' actual aCRT programming will be a time-varying covariate that has 3 categories: ACRT ON, ACRT OFF, and Neither, counting from the time randomization programming is implemented. Patient with major eligibility violations discovered after randomization will be excluded. Patients with major protocol violations after randomization, such as being programmed permanently to a mode that deviated from original randomization assignment due to a system modification after randomization, will be censored at the time of protocol violations.

4.1.2.4 Subgroup Analysis/Poolability

Poolability

Site will be included as a random effect to account for the site-by-site variability in the primary analysis of the primary endpoint, and baseline characteristics will be adjusted for in an ancillary analysis. Descriptive statistics for the primary endpoint will be reported study-wide and by investigational sites. In the circumstance where the statistical model of the primary analysis cannot converge due to low enrollments of some sites, sites with less than 5 randomized subjects will be combined to "small sites" first within city; if it doesn't work, then within state (if applicable), country, geographic region, and continent in order. Once a combined site has at least 5 randomized subjects, a new combined site will be started at the same geographic level. If there is only one site with less than 5 randomized subjects left at one geographic level, this site will be combined to the nearest site at the same geographic level.

Subgroup analysis

Subgroup analysis of the primary efficacy endpoint will be performed for a limited number of pre-specified subgroups by testing the effect of treatment by subgroup interaction on primary endpoint. The subgroups include:

- Geography: US/Canada vs. International
- Age
- Gender: Males vs. Females
- QRS duration: ≤150ms vs. >150ms
- AV conduction time: median as cutoff
- LBBB: Core Lab confirmed vs. Core Lab not confirmed
- NYHA class (in different model where NYHA class is a covariate not a stratification factor)
- HF etiology
- Renal dysfunction
- Diabetes

The treatment by subgroup interaction effects for the subgroup described above will be tested at α=0.05. Adjustment for multiple comparisons will be done. The interaction effects, confidence intervals, and unadjusted p-values for multiple comparisons will be reported using a Forest plot.

Similar subgroup analysis may also apply to the two secondary endpoints that are contributed to the primary endpoint: all-cause mortality, and intervention for heart failure decompensation.

4.1.2.5 Ancillary Analysis for Covariates Adjustment

Ancillary analysis will be performed for the primary objective to adjust for subject characteristics collected at enrollment and baseline. Such analysis may also be considered for the first two secondary objectives that are components of the primary endpoint. Modeling selection and determination will be described in the corresponding sections for those objectives.

4.1.2.6 Missing Data

All the data collected from the study will be reported and unauthentic data (if any) will not be included in any analysis dataset. Details on how to handle missing data are provided in the following sections for all the study objectives.

4.1.3 Reports for which this Statistical Analysis Plan applies

This analysis plan shall apply to the final report and other study analyses. Analyses that are conducted for Data Monitoring Committee (DMC) meetings will use the methods from this analysis plan as guidance when applicable. Statistical analysis for study-related publications will not be limited to this plan.

This study includes DMC reports not associated with an interim analysis, DMC reports based on 3 interim analyses, and a final report. A set of analyses including primary ITT analysis, sensitivity analyses per LBBB and aCRT programming, subgroup analysis for interaction with treatment, and ancillary analysis for covariates adjustment have been described in section 4.1.2. The analyses to be conducted at each stage of the study may vary by report type and study objectives that are considered. Table **2** indicates the analyses for each study objective that we will be conducted for the different study reports. Contents of this table are subject to change, depending on the needs of the DMC or the study itself. Note that, in order to control for family-wise α level, confidence intervals and p values will not be provided in any non-interim DMC report.

Table 2: Analysis plan matrix – when, what and for which

	DMC Reports Non-Interim	**DMC Report Interim**	**Final Report**
Primary Objective	Primary Analysis Sensitivity Analysis	Primary Analysis Sensitivity Analysis Subgroup Analysis	Primary Analysis Sensitivity Analysis Subgroup Analysis Ancillary Analysis
Secondary Objective #1: all-cause mortality	Primary Analysis	Primary Analysis Sensitivity Analysis	Primary Analysis Sensitivity Analysis Subgroup Analysis Ancillary Analysis
Secondary Objective #2: HF events	Primary Analysis	Primary Analysis Sensitivity Analysis Recurrent Events	Primary Analysis Sensitivity Analysis Subgroup Analysis Recurrent Events

			Ancillary Analysis
Secondary Objective #3 - #7	None	Primary Analysis	Primary Analysis
Ancillary Objectives	None*	None*	Primary Analysis

* Except for Ancillary Objective #3 (cardiovascular adverse events) that will be summarized in these reports,

4.2 Interim Analysis Plan

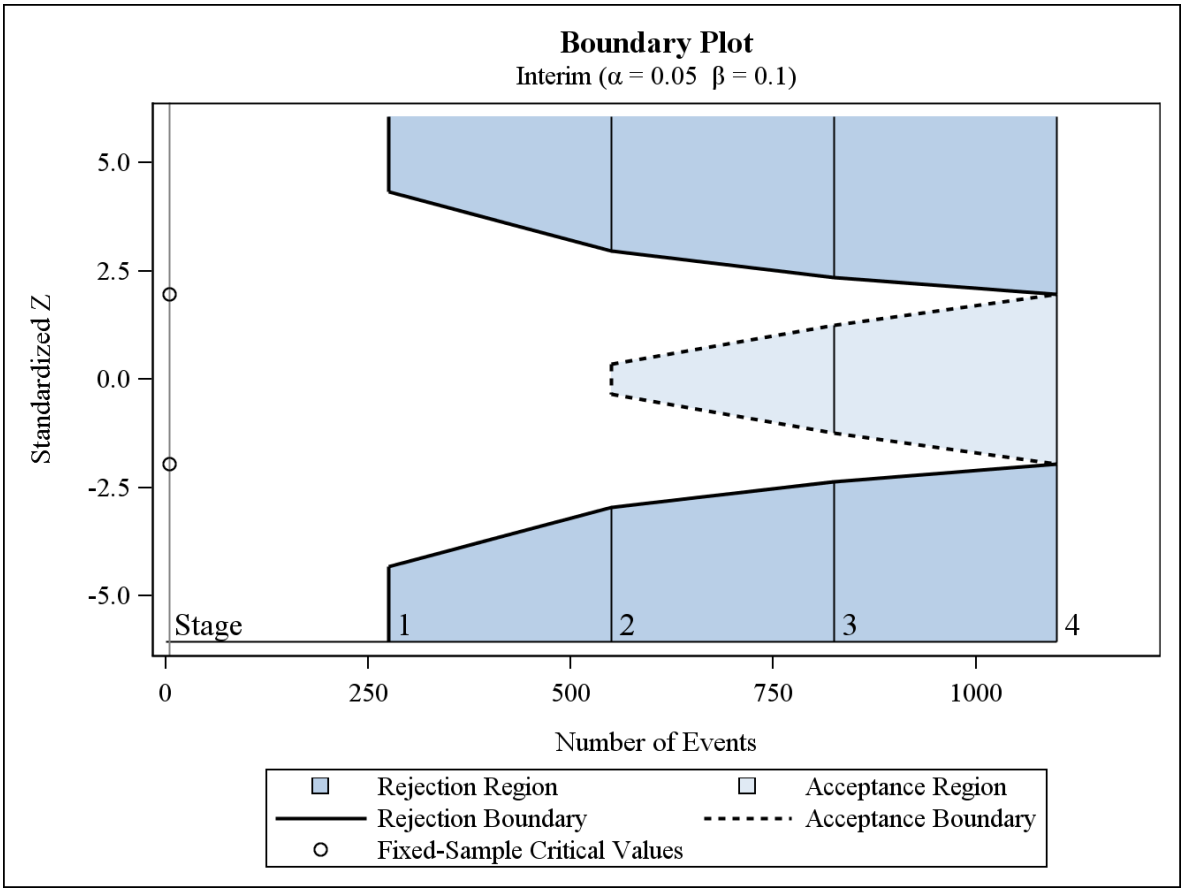
Four analyses including three interim analyses and a final analysis are planned for the primary objective. The 3 interim analyses will follow a symmetric group sequential design using the alpha-spending methodology of Lan and DeMets[2] with O'Brien-Fleming[3] type boundaries. Specifically, the cumulative alpha-spending function at the k^{th} look or analysis:

$$\alpha(t_k) = 2 - 2\Phi(\frac{Z_{\alpha/2}}{\sqrt{t_k}})$$

where k=1, 2,…, K and t_k is the information fraction at the k^{th} look, In the group sequential design for time-to-event endpoint, the information fraction at the k^{th} look can be defined as the cumulative count of events at the kth look divided by the projected number of events at the final look.

The O'Brien-Fleming function was chosen because it conservatively preserves most of the type I error for the final analysis, and the conservation of type I error is desired due to the multiple endpoints and objectives of interest in this trial. The statistical stopping rules are illustrated in the figure below.

Figure 1: Statistical stopping boundaries



The horizontal axis shows the number of events, indicating the timing of interim analyses after 275, 550, and 825 events, and the final analysis when 1100 events are accrued. If at any interim analysis the standardized Z statistic for the primary objective is in the darker blue area with solid boundary line, the DMC may advise to stop the study claiming the null hypothesis is rejected. This corresponds to a cumulative α-level of 0.00002, 0.00304, 0.0193 and 0.05 spent at the subsequent stages (see section 5.1 “Error Spending Information” table). If the Z Statistic is in the lighter blue area with dashed boundary line, the DMC may advise to stop the study for futility.

In case the study is stopped for efficacy after an interim analysis, the secondary objectives will be analyzed using a family-wise type I error rate derived from a Pocock-type alpha-spending function. This corresponds to a cumulative α-level of 0.018, 0.031, 0.041 and 0.05 spent at the subsequent stages[4]. Hommel procedure will be used to correct for multiplicity.[5] This will be implemented by the SAS MULTTEST procedure.

Planned interim analyses, and any unplanned interim analyses, will be conducted under the auspices of the DMC assigned to this study. Medtronic will use the FDA recommended guidance entitled Establishment and Operation of Clinical Trial Data Monitoring Committees[6] to guide interim statistical analyses and DMC operations. Interim analyses reports will be prepared by a Medtronic statistician and SAS programmers who are

members of the study team and are unblinded to randomization . A Medtronic statistician otherwise not involved in the study will perform validation. The lead statistician and all other Medtronic personnel will be blinded to interim analyses. The DMC is authorized to review unblinded interim analyses. The DMC will disseminate interim results only if absolutely necessary. Study centers will not receive information about interim results unless they need to know for the safety of their subjects.

Data review by the DMC that occurs between interim/final analyses will not spend type I error rate on the primary objective because there is a mutual agreement between the study team and DMC that no p values will be provided for any non-interim/final analysis and DMC will not make any decision on trial continuation or termination based on the results of such analysis. Early DMC meetings will monitor adverse events for unexpected risks to the study subjects and the overall conduct of the study. For each meeting, the DMC will be provided with information about enrollment rates, eligibility, baseline characteristics, and adverse events. The DMC will also be provided with information about the primary endpoint and its components. The planned subgroup analyses will be reported so that the DMC may examine futility in these subgroups. In order to stop the study earlier than planned, priority will be given to ensuring that there is sufficient information to assess cost-effectiveness. In addition, this assessment will balance the evidence at the interim analysis about the primary objective and secondary objectives #1 and #2. According to the DMC Charter, “Should the DMC recommend terminating the study early or recommend significant changes to the protocol, the SC and Medtronic may then be unblinded to the data if deemed necessary by the DMC. Any such dissemination of unblinded data will be documented and described in the final study report […] The DMC will need to consider the level of completeness and adjudication status of the study at the moment of advice. Especially, in case of a recommendation to stop the study, the DMC will propose an implementation plan that accounts for the risk that completion of data cleaning and adjudication may alter results.”

Interim Database Freezes

For each analysis prepared for a DMC meeting and each interim analysis, a visit cutoff date will be defined. For regularly scheduled DMC meetings, the visit cutoff date will be defined to allow time for data cleaning and analysis prior to the DMC meeting. A received cutoff date will also be declared prior to each DMC-related database freeze. Only Case Report Forms dated before the determined visit cutoff date and saved complete in the study database prior to the received cutoff date will be included in each DMC analysis.

Events that would trigger an interim analysis should be adjudicated. This will be subject to a time lag associated with the activities leading up to adjudication, including the site entering the information into the study database and saving as complete, the information about the event being sent to the Endpoint Adjudication Committee (EAC) (which will typically be sent within a batch) and the event being adjudicated and its adjudication results saved into the study database. EAC meetings occur on a regular basis in order to ensure that the percentage of adjudicated events is high.

When the number of patients with any adjudicated events comes close to the target of an interim analysis (275, 550 and 825 events for the 1st, 2nd and 3rd interim analysis, respectively), a DMC meeting for the interim analysis will be scheduled. The visit cutoff date will be about 6 weeks before the DMC meeting. It will be strived for to have all events that occurred prior to the visit cutoff date adjudicated. The received cutoff date will be right after the completion of adjudication for all the events, or 2 weeks before the meeting at the

latest. The database freeze date should be no later than 2 weeks after the received cutoff date. Note that the number of events before the database freeze might not be exactly the same as the target for an interim analysis. In such a case, an updated α-value for the interim analysis will be calculated using the SAS SEQTEST procedure accordingly. To foresee what is coming based on the best available information of the trial, the same analysis will be applied to all the primary events per investigators' evaluation in the interim analysis database. This will be done regardless of the adjudication status of those events.

4.3 Primary Objective: All-cause mortality and intervention for heart failure decompensation

The primary objective is to test the hypothesis that aCRT reduces the incidence of the combined endpoint of all-cause mortality and intervention for heart failure decompensation, compared to standard CRT therapy, in patients with a CRT indication, LBBB and normal AV conduction.

4.3.1 Hypothesis

The null hypothesis is that the hazard ratio for treatment (aCRT ON) versus control (aCRT OFF) is equal to 1. It will be tested against the alternative hypothesis that the hazard ratio is unequal to 1.

Let $S_T(t)$ and $S_C(t)$ denote the proportion of patients in the treatment arm and control arm, respectively, that have not experienced a primary endpoint up to time t. Let $h_T(t)$ and $h_C(t)$ be the associated hazard functions. Under the proportional hazards model, $h_T(t) = HR*h_C(t)$, where HR is the hazard ratio of treatment compared to control. The null and alternative hypotheses can be expressed as:

$H_0 : HR = 1$

$H_A : HR \neq 1$

Endpoint Definition

The primary endpoint is the composite of all-cause death and any intervention for heart failure decompensation (HF event) as adjudicated by the EAC. Intervention of heart failure decompensation is defined as an event that requires "invasive intervention (i.e. IV diuretics, ultrafiltration, or equivalent) or inpatient hospitalization".

The EAC will determine if an endpoint has been met. Sites may adhere to their standard practice diagnosing heart failure, but are required to report all diagnostic assessments, tests and procedures done with supporting material as appropriate to allow the EAC to adjudicate.

The analysis will use time to event in study for each subject, defined from the date of randomization (per randomization form) to date of last follow-up or date of primary event, whichever is first. The date of the event will be the date of death or the date of initiation of treatment for heart failure decompensation as determined by the EAC. Subjects without

primary event will be followed up until study closure. Subject deaths and HF events will be collected on case report forms. Any events that occur before the date of randomization will not be included in the intent-to-treat analysis. If a subject has more than one event that contributes to the primary endpoint, the first of these events will be used in the test statistic.

The study will be event driven, based on the number of patients that experienced an event. The study will continue until a predetermined number of events have been observed, unless the DMC advises to stop earlier.

4.3.2 Analysis Methods

The hypothesis will be tested using a Cox proportional hazards regression model with a random site effect (frailty), stratified by NYHA class at enrollment. In this combination of stratified Cox regression and frailty model, the hazard for subject *i* from site *j* with NYHA class *k* at enrollment is:

$$\lambda_{ijk}(t) = \lambda_{0k}(t)\xi_j \exp(\beta X_{ijk})$$

where i =1, …, n_j; j = 1, …, J; k = 1, 2, 3.

$\lambda_{0k}(t)$ is the baseline hazard of k^{th} NYHA class. This is the part of stratified Cox regression, where the baseline hazards of NYHA class play no role in the estimation of treatment effect. That is, no between-NYHA-class comparisons of treatment effect are attempted and all information about the treatment effect comes from within NYHA class comparisons. ξ_j (j = 1, …, J) are frailties or site effects that follow a gamma distribution with mean 1 and variance θ_0 independently and identically. The variance parameter θ_0 is a measure of the heterogeneity across sites in baseline risk. As θ_0 increases, the values of ξ_j become more dispersed, inducting greater heterogeneity in the site-specific baseline hazards. This is the part of random effects or frailty model.[7] Therefore, $\xi_j\lambda_{0k}(t)$ is the baseline hazard of k^{th} NYHA class at j^{th} site, and ξ_j is constant for k = 1, 2, 3 NYHA class from the same site j. X_{ijk} is the randomization assignment (ACRT ON or ACRT OFF), and exp(β) is the hazard ratio for treatment.

The SAS PHREG procedure will be used to conduct the survival analysis of primary endpoint using code similar to:

```
PROC PHREG DATA=primary;
 CLASS NYHA site;
 MODEL time*event(0) = rand / TIES=efron;
 STRATA NYHA;
 RANDOM site/ dist = gamma;
RUN;
```

Event = 1 when a subject experienced an event that was adjudicated as primary endpoint (intervention for HF decompensation or death). Event = 0 when a subject did not experience a primary endpoint, and therefore was censored. Time to event starts from the date of randomization and ends at the date that the event occurs for those who experience

the event or the date of last documented follow-up since randomization for those who are censored. Date of last documented follow-up since randomization is the maximum value among the dates of randomization, randomization programming, any study scheduled follow-up visit, exit, adverse event onset, adverse event resolved, system modification, and healthcare utilization. Rand is an indicator variable for randomized treatment groups (1=aCRT ON, 0=aCRT OFF).

The hazard ratio for treatment arm (aCRT ON) versus control arm (aCRT OFF) will be reported with its associated p-value and 95% confidence interval. The null hypothesis will be rejected when the p-value is below the critical value for the (interim or final) analysis. The actual cumulative α level will be calculated based on the actual number of primary events included in each interim analysis. The estimated cumulative α-level is approximately 0.00002, 0.00304, 0.0193 and 0.05 at the 1st interim with 275 events, 2nd interim with 550 events, 3rd interim with 825 events and the final analysis with 1100 events, respectively (see details in section 0). It will be concluded that aCRT reduces the incidence of the primary endpoint when the null-hypothesis is rejected and the hazard ratio is below 1. Kaplan-Meier plots for incidence of primary endpoint will be presented by treatment arms, for all patients and for the separate NYHA class groups. The absolute event rate difference at 2 years post randomization will also be calculated.

The following model assumptions will be checked:

1. Proportional hazards between treatment arms: an interaction term of treatment arm by time (in log scale) will be added into the model (see the code below). If this interaction effect is not significant, the proportional hazard assumption holds. Otherwise, this interaction term should be kept in the model as a remedy for non-proportional hazards, and any secondary analysis where treatment (or treatment by subgroup factor interaction) is included should also include the corresponding interaction term with time. Being time dependent, the treatment effect will be assessed every 6 months to the end of the follow-up to monitor the direction. Significance of the treatment effect will be based on its estimate at 24 months using a contrast.

```
PROC PHREG DATA=primary;
 CLASS NYHA site;
 MODEL time*event(0) = rand rand_t / TIES=efron;
 rand_t = rand*log(time);
 STRATA NYHA;
 RANDOM site/ dist = gamma;
RUN;
```

2. Hazard ratio for treatment is constant across NYHA classes: the following code will be used to check this assumption. That is, instead of doing stratified Cox regression, NYHA class will be regarded as a covariate in a frailty model along with the interaction between treatment and NYHA class. If the interaction is not significant and the hazard ratios for treatment are similar at different NYHA classes, the assumption of stratified Cox regression part of the model in primary analysis is valid. If the interaction is significant, the assumption does not hold, and we will consider NYHA class as a covariate instead of a stratification factor in the frailty model for primary analysis and other relevant analyses, and keep the interaction term in the model. The treatment effect will be reported separately for the different NYHA classes.

```
PROC PHREG DATA=primary;
 CLASS NYHA(ref='NYHA class II') site rand(ref=0)
```

```
            / param=ref order=internal;
 MODEL time*event(0) = rand NYHA rand*NYHA / TIES=efron;
 RANDOM site/ dist = gamma;
 HAZARDRATIO rand / diff=all at(NYHA='NYHA class II');
 HAZARDRATIO rand / diff=all at(NYHA='NYHA class III');
 HAZARDRATIO rand / diff=all at(NYHA='NYHA class IV');
RUN;
```

4.3.3 Missing Data

Follow-up information for all participants is expected whether or not they continue with therapy or remain compliant (i.e. ITT principle). For survival analysis, if a subject is lost to follow up they will be censored at the last date for which the subject was known to be alive or free from the event. This means we assume non-informative censoring for lost to follow up. Here lost to follow up refers to conditional disengagement and early exit prior to study completion.

4.3.4 Sensitivity Analyses: LBBB Determination and aCRT Programming

As indicated in section **Error! Reference source not found.**, two sensitivity analyses will be performed for the primary objective. One is an ITT analysis of the randomized patients whose LBBB is confirmed by the ECG Core Lab. The purpose of this sensitivity analysis is to verify the effect of aCRT on the incidence of HF events or death among the true LBBB patients and to assess the impact of LBBB misclassification on the results. This analysis will be conducted in a similar way to that of the primary analysis:

```
PROC PHREG DATA=LBBB;
 CLASS NYHA site;
 MODEL time*event(0) = rand / TIES=efron;
 STRATA NYHA;
 RANDOM site/ dist = gamma;
RUN;
```

Here, 'event', 'time' and 'aCRT' are defined in the same way as those in the primary analysis. A p-value for treatment below the cumulative α-level of approximately 0.00002, 0.00304, 0.0193 and 0.05 at the 1st, 2nd, 3rd interim and the final analysis, respectively, is significant.

The other sensitivity analysis will also be conducted in the randomized patients with ECG Core Lab confirmed LBBB, while the purpose is to verify the effect of aCRT on the primary endpoint based on the actual aCRT programming over time. Given this being said, this analysis will not follow the ITT principle and:

- Patients with major eligibility violations will be excluded (also when the violation becomes apparent only after randomization).
- Patients with major protocol violations after randomization, such as being programmed permanently to a mode that deviated from original randomization assignment due to a system modification after randomization, will be censored at the time of protocol violations.
- Patients' actual aCRT programming will be a time-varying covariate that has 3 categories: ACRT ON, ACRT OFF, and Neither, counting from the time randomization programming is implemented.

This analysis will be conducted using the following SAS PHREG procedure:

```
PROC PHREG DATA=LBBBCross COVSANDWICH(AGGREGATE);
 ID pt;
 CLASS NYHA site pt;
 MODEL (start,stop)*event(0) = aCRTprog / TIES=efron;
 STRATA NYHA;
RUN;
```

where aCRTprog indicates patients' actual aCRT programming. Patients whose actual aCRT programming change over time get multiple records in the input dataset. Whenever there is a change, for instance, there will be one record in the randomized arm up until time of crossover and another record from time of crossover until the time that a primary endpoint is met or the time of censoring without meeting a primary endpoint. The COVSANDWICH option ensures the model accounts for multiple records coming from the same patient. In order to implement the COVSANDWICH option in SAS, random site effect will not be considered in this sensitivity analysis. A p-value for treatment below the cumulative α-level of approximately 0.00002, 0.00304, 0.0193 and 0.05 at the 1st, 2nd, 3rd interim and the final analysis, respectively, is significant.

4.3.5 Subgroup Analysis: Interaction with Treatment

Subgroup analysis of the primary endpoint will follow the ITT principle and include all randomized patients. The purpose is to test if the effect of aCRT treatment on primary endpoint would vary by pre-specified subgroup factors including:

- Geography: US/Canada vs. International
- Age
- Gender: Males vs. Females
- QRS duration: ≤150ms vs. >150ms
- AV conduction time: median as cutoff
- LBBB: Core Lab confirmed vs. Core Lab not confirmed
- NYHA class (in different model where NYHA class is a covariate not a stratification factor)
- HF etiology
- Renal dysfunction
- Diabetes

This analysis will be conducted by adding subgroup and its interaction with treatment into the model used for the primary analysis. The codes will be similar to:

```
PROC PHREG DATA=primary;
 CLASS NYHA site subgroup;
 MODEL time*event(0) = rand subgroup rand*subgroup / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

The treatment by subgroup interaction effects will be tested at α=0.05. P values will not be adjusted for multiplicity. The significance of interaction effects and confidence intervals of hazard ratios for subgroups will be presented in a Forest plot.

4.3.6 Ancillary Analysis: Covariate Adjustment and Risk Score

Subject characteristics at enrollment and baseline will be considered as covariates in an ancillary analysis. The purpose is to verify the treatment effect controlling for covariates. The ancillary analysis may only be conducted for the final report.

The subject characteristics include: from enrollment visit, gender, PR interval, QRS duration, and LVEF; from baseline form, age, physical exam status, cardiac disease classification, current HF symptoms, blood measurements, and medical history at baseline (section 4.1.1). NYHA class at enrollment will be regarded as a covariate instead, which means the corresponding model will no longer have the component for stratified Cox regression.

First, the effect of each subject characteristic on the primary endpoint will be estimated in a univariate model as follows:

```
PROC PHREG DATA=primary;
 CLASS site SubjCha;  * CLASS only if SubjCha is categorical;
 MODEL time*event(0) = SubjCha / TIES=efron;
 RANDOM site/dist = gamma;
RUN;
```

Again, in this SAS PHREG procedure, event = 1 when a subject experiences a primary endpoint and event = 0 otherwise; and time to event starts from the date of randomization and ends at the date that the first event occurs for those who experience the event or the date of last documented follow-up since randomization for those who are censored; SubjCha is a specific baseline characteristic of interest.

Next, the subject characteristics that are significant in a univariate model will be included in a full model and backwards selection method will be used to determine significant predictors. After that, the treatment factor will be added into the model along with the interaction terms between treatment and possible effect modifiers such as gender, ECG Core Lab confirmed LBBB, AV conduction time, QRS, NYHA class at enrollment and MI. Backwards selection will be used to eliminate the interaction terms that are not significant. If an interaction is significant, the relevant subject characteristics need to remain in the model regardless whether their single effects being significant or not. Note the backwards selection will be performed manually because the SAS PHREG procedure ignores the SELECTION option for the frailty model analysis.

Based on the literatures[8,9], the following characteristics are expected in the final model: from enrollment form, gender, QRS duration, ECG Core Lab confirmed LBBB, LVEF, NYHA class, AV conduction time; from baseline form, age, BMI, SBP, rales/crackles, dyspnea, peripheral edema, HF admissions in past 6 months, creatinine, history of MI, AF, diabetes and smoking; from cardiovascular medication log, baseline diuretics dose, and ACE/ARB. If not, we might consider to add them into the final model in order to compare with the literature.

A risk score for the primary endpoint can be defined as $\mathbf{X\beta}$ with $\mathbf{X}$ being the covariates remained in the final model.

The critical values for the treatment effect in the primary analysis (estimated cumulative α-level of approximately 0.00002, 0.00304, 0.0193 and 0.05 at the 1st, 2nd, 3rd interim and the final analysis, respectively) will not be applied to this ancillary analysis.

Multiple imputation will be used to handle missing values for the covariates in the multi-variable models. The SAS MI and MIANALYZE procedures will be used to conduct multiple imputation.

4.3.7 Determination of Patients/Data for Analysis

The primary analysis will follow the ITT principle. All randomized patients will be included in analysis, and for each patient the endpoints will be attributed to the arm that the patient was randomized to irrespective of the actual treatment that the patient received. Event dates will be entered in the analysis relative to the date of randomization (which will be time zero for the survival analysis). Patients will be censored if no primary endpoint has occurred prior to the end-of-study exit visit, or prior to the last documented follow-up visit if the patient was exited early (or in case of interim analysis).

Subgroup analysis and ancillary analysis will also follow the ITT principle and include all randomized patients. However, sensitivity analyses will only include randomized patients with ECG Core Lab confirmed LBBB and ITT principle is not applicable to the sensitivity analysis per actual aCRT programming.

4.3.8 Sample Size Methods and Assumptions

In this event driven trial, the number of events for the primary objective to be met is estimated using the SEQDESIGN procedure in SAS. This procedure[10] assumes that with a total number of analysis stages K, the sequence of the standardized test statistics $\{Z_1, Z_2,\ldots, Z_K\}$ has the canonical joint distribution with information levels $\{I_1, I_2,..., I_K\}$ for the parameter θ (the response difference between treatment arm and control arm)[11]:

- $\{Z_1, Z_2,\ldots, Z_K\}$ is multivariate normal
- $Z_k \sim N(\theta\sqrt{I_k}, 1)$, $k = 1, 2,\ldots, K$
- $\text{Cov}(Z_{k1}, Z_{k2}) = \sqrt{(I_{k1}/I_{k2})}$, $1 \le k_1 \le k_2 \le K$

If the test statistic is computed from data that are not from a normal distribution, then it is assumed that the test statistic is computed from a large sample such that the statistic has an approximately normal distribution. In the SEQDESIGN procedure, the MODEL=TWOSAMPLESURV option of the SAMPLESIZE statement derives the number of events based on the log-rank statistic for testing two survival distributions. The computation details can be found in SAS/STAT® 9.2 User's Guide[12].

A post-hoc analysis of the Adaptive CRF Clinical Study showed that, in patients with normal AV conduction and LBBB indicated by medical history and having intrinsic rhythm determined by device, aCRT was associated with a lower risk of HF hospitalization or death (unadjusted hazard ratio (HR) = 0.71, 95% CI: 0.40 – 1.27) and a lower risk of intervention for HF decompensation or death (unadjusted HR = 0.67, 95% CI: 0.38 – 1.17). After adjusting for renal dysfunction, QRS duration (≤156 vs. >156) and LVEF, the adjusted HR of aCRT was 0.83 (95% CI: 0.46 – 1.49) for HF hospitalization or death and 0.76 (95% CI: 0.43 – 1.35) for intervention for HF decompensation or death. The event-free rate at 2

years post randomization in the control arm was 73.1% for HF hospitalization or death and 70.3% for intervention for HF decompensation or death.

Based on these results and given the consideration that misdiagnosis of LBBB might weaken the beneficial effect of aCRT on HF event or death, it was assumed that the true intent-to-treat hazard ratio for aCRT ON compared to aCRT OFF is 0.82 and the event-free rate at 2 years post randomization in aCRT OFF arm is 75%. Assuming time to the first HF event or death follows an exponential distribution, the hazard of aCRT OFF arm is equal to –ln(0.75)/2 = 0.14384.

The required number of HF event or death to get 90% overall study power is estimated using the following SAS SEQDESIGN procedure (also Section 5.1):

```
proc seqdesign
    boundaryscale=stdz
    errspend
    stopprob
    plots=boundary(hscale=samplesize);
    ;
  Interim: design
    alpha = 0.05
    alt = twosided
    beta = 0.10
    stop = both
    nstages = 4
    method = errfuncobf
    ;
  samplesize
    model = twosamplesurv
      ( nullhazard = 0.14384
        hazardratio = 0.81818
        acctime = 3.0
        foltime = 2.25
             );
run;
```

that assumes:

- Type I error rate: 0.05
- Type II error rate: 0.10, which corresponds to 90% power
- 3 equally spaced interim analyses plus the final analysis
- Hazard of aCRT OFF arm: 0.14384
- Hazard ratio of aCRT: 0.81818
- Enrollment duration: 3 years
- Follow-up duration: 2.25 years

The results (Section 5.1) show that a total of 1100 patients experiencing a primary endpoint will give 90% power to detect a reduction of the incidence of the primary endpoint at a significance level of 0.05, accounting for 3 equally spaced interim analyses and assuming a true intent-to-treat hazard ratio of 0.82 for aCRT ON compared to aCRT OFF.

With the inclusion of 2874 randomized patients enrolled over 3 years and followed for 2.3 years, 1100 events are expected when the true event-free rate of aCRT OFF arm is 75% at 2 years.

The study will randomize up to 3000 patients in order to accommodate attrition due to early exit and to accommodate the possibility of a somewhat lower control arm event rate. Through simulation (Section 1.1) it was confirmed that the study will have 90% power under the following assumptions:

- 3000 patients randomized in 3 years with a uniform rate
- Event-free rate of aCRT OFF arm is 75% at 2 years
- LBBB is not confirmed by the ECG Core Lab in 10% of patients
- aCRT hazard ratio is 0.78 in confirmed LBBB patients and 1.0 in unconfirmed patients
- Crossover rate at 2 years is 5% in aCRT ON arm and 8% in aCRT OFF arm
- Loss to follow-up rate is 5% at 2 years in each arm
- Final analysis is done when 1100 primary events are accrued

4.4 Secondary Objectives

Secondary objectives include the two components of the primary endpoint, clinical composite score, incidence of atrial fibrillation, quality of life, all-cause re-admission after a heart failure admission, and cost effectiveness for the healthcare system.

Interpretation of results will be guided by a formal multiple testing procedure to achieve strict control of the family wise error rate, also accounting for the interim analysis plan. Secondary objectives will be analyzed when the study has stopped for efficacy after an interim analysis or has reached the final analysis stage. As indicated in section 0, the cumulative α-level to be spent at the 1st, 2nd, 3rd interim analysis and final analysis for secondary objectives is 0.018, 0.031, 0.041 and 0.05, respectively, corresponding to a Pocock-type alpha-spending function. At the analysis stage when secondary objectives are analyzed, a Hommel procedure[5,13] will be applied to the secondary objectives (excluding the cost-effectiveness objective) to correct for multiplicity based on the cumulative α-level of that stage. Secondary objectives for which the hypothesis is rejected under the adjusted significance level of the Hommel procedure will be reported as significant.

4.4.1 Secondary Objective #1: All-cause mortality

The secondary objective #1 is to test the hypothesis that aCRT ON reduces all-cause mortality compared to aCRT OFF. All-cause mortality is one of the two components that comprise the primary endpoint.

Hypothesis

Let $S_T(t)$ and $S_C(t)$ denote the proportion of patients in the treatment arm (aCRT ON) and control arm (aCRT OFF), respectively, that have not died up to time t. Let $h_T(t)$ and $h_C(t)$ be the associated hazard functions. Under the Cox regression model that assumes proportional hazards, $h_T(t) = HR*h_C(t)$, where HR is the hazard ratio of treatment compared to control. The null and alternative hypotheses are:

H_0 : HR = 1

H_A : HR ≠ 1

Analysis Method

Similar to the primary endpoint, the hypothesis regarding the aCRT effect on all-cause mortality will be tested using a Cox proportional hazards regression model with a random site effect (frailty), stratified by NYHA class at enrollment. Section 4.3.2 has described this model in details. The analysis will be conducted using code similar to:

```
PROC PHREG DATA=mortality;
 CLASS NYHA site;
 MODEL time*death(0) = rand / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

In these codes, death = 1 given an adjudicated death and death = 0 otherwise; time to death starts from the date of randomization and ends at the date of death for those who die, or the date of last documented follow-up since randomization for those who are censored. Date of last documented follow-up since randomization is the maximum value among the dates of randomization, randomization programming, any study scheduled follow-up visit, exit, adverse event onset, adverse event resolved, system modification, and healthcare utilization.

The model assumptions will be checked using the same methods that are described in detail in section 4.3.2 for the primary endpoint. That is, the assumption of proportional hazards between the treatment arms will be checked by testing the significance of treatment by time interaction term; and, the assumption of constant hazard ratio for treatment across NYHA classes will be checked by testing the significance of treatment by NYHA class interaction and evaluating the hazard ratio for treatment within each NYHA class.

Missing Data

Similar to the analysis for primary endpoint, we assume non-informative censoring for lost to follow up in the analysis of all-cause mortality.

Sensitivity Analyses: LBBB Determination and aCRT Programming

The first sensitivity analysis is to verify the effect of aCRT treatment on all-cause mortality among the true LBBB patients. The same model used in the primary analysis will be used in this sensitivity analysis. The only difference is that patients included in the sensitivity analysis must have their LBBB confirmed by ECG Core Lab.

```
PROC PHREG DATA=mortality_LBBB;
 CLASS NYHA site;
 MODEL time*death(0) = rand / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

The second sensitivity analysis is to check the effect of aCRT on all-cause mortality based on the actual aCRT programming, given that patients' LBBB has been confirmed by the ECG Core Lab. Again, actual aCRT programming will be a time-varying covariate, meaning patients whose actual aCRT programming change over time will have multiple records in the input dataset. This will be performed using the COVSANDWICH option in the SAS PHREG procedure, and random site effect will not be considered in this sensitivity analysis.

```
PROC PHREG DATA=mortality_LBBBCross COVSANDWICH(AGGREGATE);
 ID pt;
 CLASS NYHA site pt;
 MODEL (start,stop)*death(0) = rand / TIES=efron;
 STRATA NYHA;
RUN;
```

Subgroup Analysis: Interaction with Treatment

Same as for the primary endpoint, the subgroups to be considered for all-cause mortality include:

- Geography: US/Canada vs. International
- Age
- Gender: Males vs. Females
- QRS duration: ≤150ms vs. >150ms
- AV conduction time: median as cutoff
- LBBB: Core Lab confirmed vs. Core Lab not confirmed
- NYHA class (in different model where NYHA class is a covariate not a stratification factor)
- HF etiology
- Renal dysfunction
- Diabetes

The effects of treatment by subgroup interaction on all-cause mortality will be evaluated using codes similar to:

```
PROC PHREG DATA=mortality_Subgroup;
 CLASS NYHA site Subgroup;
 MODEL time*death(0) = rand subgroup rand*subgroup / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

Ancillary Analysis: Covariate Adjustment

The process of ancillary analysis for all-cause mortality is exactly the same as that for the primary endpoint. Refer to section 4.3.6 for details on model building and model selection. Note that NYHA class at enrollment will be regarded as a covariate not a stratification factor in this ancillary analysis. Therefore, the entire process will be based on frailty models. For example, in the beginning the effect of each subject characteristic on all-cause mortality will be tested in a univariate frailty model using codes similar to the below.

```
PROC PHREG DATA= mortality;
 CLASS site SubjCha;  * CLASS only if SubjCha is categorical;
 MODEL time* death(0) = SubjCha / TIES=efron;
 RANDOM site/dist = gamma;
RUN;
```

After that, the final multivariate Cox regression model will be determined from the full model using backwards selection. Again, the backwards selection will be performed manually because the SAS PHREG procedure ignores the SELECTION option for the frailty model analysis.

Determination of Patients/Data for Analysis

Similar to the primary endpoint, the primary, subgroup and ancillary analysis for all-cause mortality will follow the ITT principle and include all randomized patients; sensitivity analyses will only include randomized patients with ECG Core Lab confirmed LBBB and ITT principle is not applicable to the sensitivity analysis per actual aCRT programming.

4.4.2 Secondary Objective #2: Intervention for heart failure decompensation

The secondary objective #2 is to test the hypothesis that aCRT ON reduces the rate of intervention for heart failure decompensation compared to aCRT OFF. It is the other component of the two that comprise the primary endpoint.

Hypothesis

Per CIP, intervention for heart failure decompensation (HF event) is an event that requires "*invasive intervention (i.e. IV diuretics, ultrafiltration, or equivalent) or inpatient hospitalization*". The terms "intervention for heart failure decompensation" and "HF event" are exchangeable also in the SAP.

Considering the presence of death as a competing risk, the cumulative incidence function (also called subdistribution function) of intervention for heart failure decompensation in the treatment arm, $F_T(t)$, and control arm, $F_C(t)$, are the proportion of patients that have had the first intervention for heart failure decompensation prior to time t.

In a competing risks setting, the effect of covariates can be modeled in two distinct manners[14 15]. Covariates can be modeled as having a multiplicative effect on either (1) the cause-specific hazard (Cox regression), or (2) the derivative of the cumulative incidence function (Fine-Gray regression).

In the primary analysis, Cox regression will be used for competing risk analysis where the hazard of the event of interest is modeled and there is no simple connection between covariate effect from Cox model and cumulative incidence curves.

Let $h_T(t)$ and $h_C(t)$ be the associated hazard functions. Under the Cox regression model, $h_T(t) = HR*h_C(t)$, where HR is the hazard ratio of treatment compared to control for HF event. The null and alternative hypotheses can be expressed as:

H_0 : HR = 1

H_A : HR ≠ 1

Analysis Method

The crude cumulative incidence curves of intervention for heart failure decompensation will be displayed by treatment and NYHA class combinations using the SAS autocall macro %CUMINCID:

```
%cumincid(data=HFevent, time=TtoHFevent, status=HFeventCens,
event=1, compete=2, censored=0, strata=randNYHA)
```

where TtoHFevent starts from the date of randomization and ends at the date that the first HF event occurs for those who experience the HF event, or the date of death for those who died before experiencing a HF event, or the date of last documented follow-up for those who are censored. HFeventCens = 0 when a subject did not experience a HF event or death and therefore was censored; HFeventCens = 1 when a subject experienced a HF event the first time since randomization; HFeventCens = 2 when a subject died without experiencing a HF event. The competing risk exists when HFeventCens = 2.

The hypothesis regarding the aCRT effect on HF event will be tested using a Cox proportional hazards regression model with a random site effect (frailty), stratified by NYHA class at enrollment. Section 4.3.2 has described this model in details. The analysis will be conducted using code similar to:

```
PROC PHREG DATA=HFevent;
 CLASS NYHA site;
 MODEL time*HFevent(0) = rand / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

In these codes, HFevent = 1 if a subject experienced a HF event and HFevent = 0 otherwise; time to HF event starts from the date of randomization and ends at the date that the first HF event occurs for those who experienced at least one HF event, or the date of last documented follow-up or date of death for those who are censored. Date of last documented follow-up since randomization is the maximum value among the dates of randomization, randomization programming, any study scheduled follow-up visit, exit, adverse event onset, adverse event resolved, system modification, and healthcare utilization.

The model assumptions will be checked using the same methods that are described in detail in section 4.3.2 for the primary endpoint. That is, the assumption of proportional hazards between the treatment arms will be checked by testing the significance of treatment by time interaction term; and, the assumption of constant hazard ratio for treatment across NYHA classes will be checked by testing the significance of treatment by

NYHA class interaction and evaluating the hazard ratio for treatment within each NYHA class.

Missing Data

Non-informative censoring is assumed for lost to follow up in the analysis of intervention for heart failure decompensation

Secondary Competing Risk Analysis: Fine-Gray Regression

In addition to the Cox regression as the primary analysis for HF event, Fine-Gray regression will be conducted as a secondary competing risk analysis.

Let $h_T(t)$ and $h_C(t)$ be the subdistribution hazard of $F_T(t)$ and $F_C(t)$, respectively. Under the Fine-Gray model[16] that assumes proportional subdistribution hazards, $h_T(t) = SHR*h_C(t)$, where SHR is the subdistribution hazard ratio of treatment compared to control for intervention for heart failure decompensation. The null and alternative hypotheses are:

H_0 : SHR = 1

H_A : SHR ≠ 1

The hypothesis regarding the aCRT effect on incidence of intervention for heart failure decompensation will be tested using a stratified Fine-Gray model where random center effect is not considered. The code will be something like below:

```
data Risk;
 NYHAc=2; rand=1; output;
 NYHAc=2; rand=0; output;
 NYHAc=3; rand=1; output;
 NYHAc=3; rand=0; output;
 NYHAc=4; rand=1; output;
 NYHAc=4; rand=0; output;
run;
proc phreg data=HFevent plots(overlay=stratum)=cif ;
 class NYHAc rand;
 model TtoHFevent*HFeventCens(0)= rand / eventcode=1;
 strata NYHAc;
 hazardratio 'Pairwise' rand / diff=pairwise;
 baseline covariates=Risk out=out1 cif=_all_ / rowid=rand;
run;
```

where EVENTCODE=1 indicates the event of interest is intervention for heart failure decompensation (i.e. HFeventCens = 1). HAZARDRATIO statement requests the subdistribution hazard ratio and its 95% confidence interval of treatment. BASELINE statement with COVARIATES= option display the cumulative incidence curves of aCRT ON and aCRT OFF for each NYHA class.

The following model assumption will be checked.

1. Constant subdistribution hazard ratios of treatment across NYHA classes: instead of stratified Fine-Gray, NYHA class will be regarded as a covariate in a Fine-Gray model along with the interaction between treatment and NYHA class. If the interaction is not

significant, this assumption would be valid. Otherwise, NYHA class will be regarded as a covariate instead of a stratification factor in the Fine-Gray model.

```
proc phreg data=HFevent;
     class NYHAc rand;
     model TtoHFevent*HFeventCens(0)= rand|NYHAc / eventcode=1;
     hazardratio 'Pairwise' rand / diff=pairwise;
run;
```

In addition, an R package of 'crrSC' which performs competing risk analysis for stratified data or clustered data may be considered.

.

Sensitivity Analyses: LBBB Determination and aCRT Programming

The first sensitivity analysis is to verify the effect of aCRT treatment on HF event among the true LBBB patients. The same model used in the primary analysis will be used in this sensitivity analysis. The only difference is that patients included in the sensitivity analysis must have their LBBB confirmed by ECG Core Lab.

```
PROC PHREG DATA=HFevent_LBBB;
 CLASS NYHA site;
 MODEL time*HFevent(0) = rand / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

The second sensitivity analysis is to check the effect of aCRT on HF event based on the actual aCRT programming, given that patients' LBBB has been confirmed by the ECG Core Lab. Again, actual aCRT programming will be a time-varying covariate, meaning patients whose actual aCRT programming change over time will have multiple records in the input dataset. This will be performed using the COVSANDWICH option in the SAS PHREG procedure, and random site effect will not be considered in this sensitivity analysis.

```
PROC PHREG DATA=HFevent_LBBBCross COVSANDWICH(AGGREGATE);
 ID pt;
 CLASS NYHA site pt;
 MODEL (start,stop)*HFevent(0) = rand / TIES=efron;
 STRATA NYHA;
RUN;
```

Subgroup Analysis: Interaction with Treatment

The same subgroups that are considered for the primary endpoint and 1st secondary endpoint will be considered for intervention for heart failure decompensation. These include:

- Geography: US/Canada vs. International
- Age
- Gender: Males vs. Females
- QRS duration: ≤150ms vs. >150ms
- AV conduction time: median as cutoff
- LBBB: Core Lab confirmed vs. Core Lab not confirmed

- NYHA class (in different model where NYHA class is a covariate not a stratification factor)
- HF etiology
- Renal dysfunction
- Diabetes

The effects of treatment by subgroup interaction on intervention for heart failure decompensation will be evaluated using codes similar to:

```
PROC PHREG DATA=HFevent_Subgroup;
 CLASS NYHA site Subgroup;
 MODEL time*HFevent(0) = rand subgroup rand*subgroup / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

Ancillary Analysis: Covariate Adjustment

The process of ancillary analysis for intervention for heart failure decompensation follows similar rules as those for the primary endpoint. NYHA class at enrollment will be regarded as a covariate instead of a stratification factor in this ancillary analysis. Thus, the entire process will be based on frailty models. In the beginning the effect of each subject characteristic on intervention for heart failure decompensation will be tested in a univariate frailty model using codes similar to the below:

```
PROC PHREG DATA= HFevent;
 CLASS site SubjCha;  * CLASS only if SubjCha is categorical;
 MODEL time* HFevent(0) = SubjCha / TIES=efron;
 RANDOM site/dist = gamma;
RUN;
```

After that, the final multivariate Cox regression model will be determined from the full model using backwards selection. Note the backwards selection will be performed manually because the SAS PHREG procedure ignores the SELECTION option for the frailty model analysis.

Analysis of Recurrent HF Events

In this analysis, the incidence of intervention for heart failure decompensation will be summarized and compared between the two treatment arms using a multistate Markov model. This model is also referred to as a stratified Andersen-Gill model.[17] This model was recommended because it allows a flexible modelling strategy that incorporates important features in the analysis of intervention for HFdecompensation and death and in the meantime extends relevant characteristics of other models for analysis of recurrent events.[18] Stratification of NYHA class and random center effect will not be considered in this analysis.

The multi-state model stratifies a Cox model by transition states as follows.

$$\alpha_k(t \mid X) = \alpha_{k_0}(t)\exp(X'\beta)$$

for covariates X, time t and transition states from k = 1,…, K.

These transition states include first intervention for HF decompensation, second intervention for HF decompensation, etc., death, death after first intervention of HF decompensation, death after second intervention for HF decompensation, etc.

The dataset should be arranged similar to:

PTID	Time1	Time2	Transition	Status	trtGroup
M0000001	0	60	R->H1	0	1
M0000001	0	60	R->D	0	1
M0000002	0	14.2	R->H1	1	2
M0000002	0	14.2	R->D	0	2
M0000002	14.2	60	H1->H2	0	2
M0000002	14.2	60	H1->D	0	2
M0000003	0	24.7	R->H1	0	2
M0000003	0	24.7	R->D	1	2
M0000004	0	38	R->H1	1	1
M0000004	0	38	R->D	0	1
M0000004	38	48.5	H1->H2	0	1
M0000004	38	48.5	H1->D	0	1

In this dataset (which will be referred to as mmsurvdata),

- The following definitions apply:
  - Time1 and Time2 are the start and end time of a period at risk for an event, which can be either a HF endpoint or a censoring event
  - The events being modeled are Randomization (R), first intervention for HF decompensation (H1), second intervention for HF decompensation (H2) and Death (D)
  - A subject from R can transition to H1 (R->H1) or D (R->D) and a subject from H1 can transition to H2 (H1->H2) or D (H1->D)
  - Status indicates whether a possible transition occurred (Status=1) or no longer at risk to present transition (Status=0)
  - trtGroup indicates whether the subject is aCRT ON (trtGroup =1) or aCRT OFF (trtGroup =2)
- The subject histories in the example dataset shown above were as follows:
  - Subject M0000001 made it 60 months with no intervention for HF decompensation or death
  - Subject M0000002 had an intervention for HF decompensation at 14.2 months and made it the remainder of the 60-month follow-up with no additional intervention for HF decompensation nor death
  - Subject M0000003 died at 24.7 months without a prior HF event
  - Subject M0000004 had a HF event at 38 months and had accrued 48.5 total months of follow-up

The definition above accounts for up to two HF events. For the actual analysis, the breadth of transition events will be dependent on the data. The value N for

H[N-1]->H[N]

where N is the maximal number of HF events modeled, will require that an adequate number of subjects experience at least N HFevents. The definition of adequate will be guided by the data, but will ultimately be left to the discretion of the analysts.

R code for implementing the multistate Markov model will be similar to:

```
> library(survival)
> fit1 <- coxph(Surv(time1, time2), status) ~ factor(trtGroup) + cluster(pt) +
strata(transition), data= mmsurvdata)
> print(cox.zph(fit1))
```

Determination of Patients/Data for Analysis

Again, the primary, subgroup and ancillary analysis for intervention for HF decompensation will follow the ITT principle and include all randomized patients; while sensitivity analysis will only include randomized patients with ECG Core Lab confirmed LBBB.

4.4.3 Secondary Objective #3: Improvement on clinical composite score (CCS)

The secondary objective #3 is to test the hypothesis that aCRT ON increases the proportion of patients that improve on the Clinical Composite Score (CCS) compared to aCRT OFF, at 6 months of follow-up.

Hypothesis

Let P_T(improved) and P_C(improved) denote the expected probability of having improved clinical composite score at 6 months of follow-up post randomization in the aCRT ON and aCRT OFF arm, respectively. The odds ratio (OR) for aCRT treatment is the odds of having improved CCS at 6 months post randomization in the aCRT ON arm against that in the aCRT OFF arm. The corresponding null and alternative hypotheses are:

H_0 : OR = 1

H_A : OR ≠ 1

Endpoint Definition

Per CIP, the Clinical Composite Score classifies patients according their clinical status at 6 months of follow-up post randomization into categories Improved, Unchanged, and Worsened.[19]

- A patient is classified Worsened in case of death, hospitalization for worsening heart failure, worsened NYHA class compared to Baseline (using last observation carried forward), or worsened status on the Global Assessment Score. Also patients that exit the study or cross over because of worsening heart failure are classified Worsened.
- A patient is classified Improved when not Worsened and there is an improvement in NYHA class at 6 months compared to Baseline, or Global Assessment Score.
- Patients that are not Worsened or Improved are Unchanged.

Analysis Method

In the primary analysis, improvement on CCS will be regarded as a dichotomous endpoint with two categories: Improved and Not Improved (i.e. Unchanged and Worsened). The probability of having improved CCS at 6 months post randomization follows a binomial distribution. The impact of aCRT treatment on improvement of CCS will be evaluated using a marginal generalized linear model similar to the following:

```
proc glimmix data=CCS;
class site rand NYHA;
model ImprovedCCS = rand NYHA / dist=binomial link=logit;
random intercept / subject=site;
run;
```

where RANDOM statement indicates random intercept effect; SUBJECT=site assumes observations within each site are correlated.

Missing Data

This is not applicable because the derivation of CCS ensures non-missing data.

Determination of Patients/Data for Analysis

The analysis follows the ITT principle and includes all the randomized subjects.

4.4.4 Secondary Objective #4: Incidence of atrial fibrillation (AF)

The secondary objective #4 is to test the hypothesis that aCRT ON reduces the incidence of AF compared to aCRT OFF.

Hypothesis

Occurrence of AF will be recorded in device data. The endpoint will be the first day after randomization on which there is >6 hrs of AF reported in the device's Cardiac Compass Trends.

Considering the presence of death as a competing risk, the cumulative incidence function of AF event in the treatment arm, $F_T(t)$, and control arm, $F_C(t)$, are the proportion of patients that have had the first >6 hrs AF episode prior to time t.

In the primary analysis, Cox regression will be used for competing risk analysis where only the hazard of the event of interest is modeled and there is no simple connection between covariate effect from Cox model and cumulative incidence curves.

Let $h_T(t)$ and $h_C(t)$ be the associated hazard functions. Under the Cox regression model, $h_T(t) = HR*h_C(t)$, where HR is the hazard ratio of treatment compared to control for occurrence of the first >6 hrs AF episode. The null and alternative hypotheses can be expressed as:

$H_0 : HR = 1$

$H_A : HR \neq 1$

Analysis Method

The same methods used for HF event in secondary objective #2 will be used for the AF endpoint here.

The crude cumulative incidence curves of AF endpoint will be displayed by treatment and NYHA class combinations using %CUMINCID:

```
%cumincid(data=AF, time=TtoAF, status=AFcens, event=1, compete=2,
censored=0, strata=randNYHA)
```

where TtoAF starts from the date of randomization and ends at the date that the first >6 hrs of AF was reported in the device's Cardiac Compass Trends for those who experienced this event, or the date of death for those who died before experiencing a AF event, or the date of last documented follow-up (including the date of last available device data) for those who are censored. AFcens = 0 when a subject did not experience an AF event or death and therefore was censored; AFcens = 1 when a subject experienced an AF event the first time since randomization; AFcens = 2 when a subject died without experiencing an AF event. The competing risk of AF event exists when AFcens = 2.

In addition, device measured AF burden and treatment for AF will be summarized. Device measured AF burden is defined as the daily time in AF as captured in Cardiac Compass averaged over all days with data. For AF reported as an adverse event, its symptoms and treatment taken documented in the CRFs will be listed.

The hypothesis regarding the aCRT effect on AF will be tested using a Cox proportional hazards regression model with a random site effect (frailty), stratified by NYHA class at enrollment. Section 4.3.2 has described this model in details. The analysis will be conducted using code similar to:

```
PROC PHREG DATA=AF;
 CLASS NYHA site;
 MODEL time*AF(0) = rand / TIES=efron;
 STRATA NYHA;
 RANDOM site/dist = gamma;
RUN;
```

In these codes, AF = 1 given the occurrence of >6 hrs AF episode and AF = 0 otherwise; time to AF starts from the date of randomization and ends at the date when the first >6 hrs of AF was reported in the device's Cardiac Compass Trends for those who experienced this event, or the date of last documented follow-up or date of death for those who are censored.

The model assumptions will be checked using the same methods that are described in detail in section 4.3.2 for the primary endpoint. That is, the assumption of proportional hazards between the treatment arms will be checked by testing the significance of treatment by time interaction term; and, the assumption of constant hazard ratio for treatment across NYHA classes will be checked by testing the significance of treatment by NYHA class interaction and evaluating the hazard ratio for treatment within each NYHA class.

Missing Data

Non-informative censoring is assumed for lost to follow up in the analysis of defined AF event.

Secondary Competing Risk Analysis: Fine-Gray Regression

In addition to the Cox regression as the primary analysis for AF, Fine-Gray regression will be conducted as a secondary competing risk analysis.

Let $h_T(t)$ and $h_C(t)$ be the subdistribution hazard of $F_T(t)$ and $F_C(t)$, respectively. Under the Fine-Gray model[14] that assumes proportional subdistribution hazards, $h_T(t) = SHR*h_C(t)$, where SHR is the subdistribution hazard ratio of treatment compared to control for occurrence of the first >6 hrs AF episode. The null and alternative hypotheses are:

H_0 : SHR = 1

H_A : SHR ≠ 1

The hypothesis regarding the aCRT effect on incidence of the first >6 hrs AF will be tested using a stratified Fine-Gray model:

```
data Risk;
 NYHAc=2; rand=1; output;
 NYHAc=2; rand=0; output;
 NYHAc=3; rand=1; output;
 NYHAc=3; rand=0; output;
 NYHAc=4; rand=1; output;
 NYHAc=4; rand=0; output;
run;
proc phreg data=AF plots(overlay=stratum)=cif ;
 class NYHAc rand;
 model TtoAF*AFcens(0)= rand / eventcode=1;
 strata NYHAc;
 hazardratio 'Pairwise' rand / diff=pairwise;
 baseline covariates=Risk out=out1 cif=_all_ / rowid=rand;
run;
```

Determination of Patients/Data for Analysis

The analysis follows the ITT principle and includes all the randomized subjects.

4.4.5 Secondary Objective #5: Quality of life measured by KCCQ

The secondary objective #5 is to test the hypothesis that the change in quality of life, measured by the Kansas City Cardiomyopathy Questionnaire (KCCQ), in the aCRT ON group is better than the change in the aCRT OFF group

Hypothesis

The KCCQ questions are categorized into clinically relevant domains[20]

- Physical limitation (question 1: 6 subquestions),
- Frequency of symptoms (questions 3, 5, 7 and 9),
- Severity of symptoms (questions 4, 6 and 8),
- Symptom stability (question 2),
- Self-efficacy and knowledge (questions 10 and 11),
- Social limitation (question 15: 4 subquestions) and
- Quality of life (questions 12, 13 and 14).

A score for each domain will be calculated. Also, total symptom score is the average of symptom frequency score and symptom severity score. In addition, two summary scores will be calculated:

- Clinical Summary Score: mean of physical limitation score and total symptom score;
- Overall Summary Score: mean of physical limitation score, total symptom score, quality of life score, and social limitation score.

Hypothesis testing will be performed on the overall summary score. Summary statistics will be provided for the other scores by study arm.

The KCCQ questionnaire is collected at baseline and at the 3, 6, 12, 18, and 24 months follow-up visits. The overall summary score (range, 0 to 100; higher scores indicate better health status) will be assessed. Let α_T and α_C denote the fixed treatment effect on KCCQ overall summary score for aCRT ON and aCRT OFF group, respectively. The null and alternative hypotheses are:

$H_0 : \alpha_T = \alpha_C$

$H_A : \alpha_T \neq \alpha_C$

Analysis Method

General linear mixed models will be used to analyze the continuous KCCQ overall summary score collected over time. In the primary analysis, the model will include the fixed effects of treatment, visit, treatment by visit interaction and baseline KCCQ score, and two random effects - random site effect and random subject effect:

$$Y_{ijkg} = \mu + \alpha_i + \beta_j + \tau_{ij} + \lambda_{kg} + \gamma_k + \eta_{i(g)} + \varepsilon_{ijkg}$$

Here:

- μ is the common intercept of the model
- α_i is the ith fixed treatment effect, i=T, C;
- β_j is the jth fixed visit effect, j=1, 2, …, J;
- τ_{ij} is the fixed interaction effect between ith treatment and jth visit;
- λ_{kg} is the fixed baseline KCCQ effect of gth subject in kth site;
- γ_k is the random site effect for the kth site, k= 1, 2, …, K;
- $\eta_{i(g)}$ is the random subject effect of gth subject in ith treatment group, g=1, 2, …, G;
- ε_{ijkg} is the random error for the gth subject from kth site having ith treatment at jth visit.

The analysis will be conducted using the SAS MIXED procedure similar to below:

```
proc mixed data=KCCQ method=reml covtest;
id pt;
class pt site rand visit;
model KCCQscore = rand visit rand*visit KCCQ_base
                    / solution ddfm=satterth covb chisq;
random site pt(rand);
run;
```

where KCCQscore is a continuous longitudinal endpoint that includes all post-randomization KCCQ measurements, visit is the categorical version of the visit number, rand*visit is the interaction between treatment and visit, and KCCQ_base is the KCCQ score at baseline. If the interaction between treatment and visit is not significant, it will be eliminated from the model and the main effect of treatment will be used to test the hypotheses; otherwise; if the treatment by visit interaction turns out to be significant, then treatment effect at each visit will be evaluated and the hypotheses will be tested using CONTRAST or LEMESTIMATE statement in the SAS MIXED procedure.

Missing Data

Missing data for individual KCCQ questions will be handled as suggested by the questionnaire authors[19]. Missing values within each domain are assigned the average of the answered items within that same domain. Scale scores are transformed to a 0 to 100 range by subtracting the lowest possible scale score, dividing by the range of the scale and multiplying by 100.

To investigate whether patients with missing KCCQ data are missing at random (MAR), the baseline characteristics between patients with missing data on KCCQ and those with complete data will be compared. If there is no evidence to suggest that patient data are not missing at random, we will cautiously conclude that patient data are MAR. Then, multiple imputation techniques will be employed to account for missing data.

Multiple imputation[21 22] models will be developed to estimate the overall scores for all subjects/visits with missing overall scores. Imputation may be stratified if differences are detected in baseline characteristics between those with or without missing data. The SAS MI and MIANALYZE procedures will be used to conduct multiple imputation.

Determination of Patients/Data for Analysis

The analysis follows the ITT principle and includes all the randomized subjects.

4.4.6 Secondary Objective #6: Quality of life measured by EQ-5D

The secondary objective #6 is to test the hypothesis that the change in health outcome, measured by the EQ-5D, in the aCRT ON group is better than the change in the aCRT OFF group.

Hypothesis

The EQ-5D index will be assessed at baseline and at the 3, 6, 12, 18, and 24 months follow-up visits. Derivation of EQ-5D index will follow instructions provided by the EuroQol Group (www.euroqol.org).

Let α_T and α_C denote the fixed treatment effect on EQ-5D index for aCRT ON and aCRT OFF group, respectively. The null and alternative hypotheses are:

$$H_0 : \alpha_T = \alpha_C$$

$$H_A : \alpha_T \neq \alpha_C$$

Analysis Method

The same general linear mixed models used for KCCQ overall summary score in section 4.4.5 will be used to analyze the EQ-5D score collected over time. Again, in the primary analysis, the model will include the fixed effects of treatment, visit, treatment by visit interaction and baseline EQ-5D score, and two random effects – random site effect and rand subject effect. The analysis will be conducted using the SAS MIXED procedure similar to below:

```
proc mixed data=EQ5D method=reml covtest;
id pt;
class pt site rand visit;
model EQ5Dscore = rand visit rand*visit EQ5D_base
                    / solution ddfm=satterth covb chisq;
random site pt(rand);
run;
```

where EQ5Dsocre is a continuous longitudinal endpoint that includes all post-randomization EQ5D measurements, visit is the categorical version of the visit number. rand*visit is the interaction between treatment and visit, and EQ5D_base is the EQ5D score at baseline. If the interaction between treatment and visit is not significant, it will be eliminated from the model and the main effect of treatment will be used to test the hypotheses; otherwise; if the treatment by visit interaction turns out to be significant, then treatment effect at each visit will be evaluated and the hypotheses will be tested using CONTRAST or LEMESTIMATE statement in the SAS MIXED procedure.

Missing Data

Missing data for EQ-5D score will be handled similarly as those for KCCQ score. Mechanism of missing will be investigated to see if EQ-5D data are missing at random (MAR). Multiple imputation will be used with the SAS MI and MIANALYZE procedures to estimate the EQ-5D scores for subjects that had missing data on this outcome.

Determination of Patients/Data for Analysis

The analysis follows the ITT principle and includes all the randomized subjects.

4.4.7 Secondary Objective #7: All-cause re-admission after an heart failure admission within 30 days of the index event

The secondary objective #7 is to test the hypothesis that aCRT reduces the incidence of all-cause re-admissions after a heart failure (HF) admission within 30 days of the index event.

Hypothesis

Assuming the number of all-cause re-admissions after a HF admission within 30 days of the index event follows a Negative Binomial distribution, $P_T(Y=k)$ and $P_C(Y=k)$ denote the expected probability of having k all-cause re-admissions after a HF admission within 30 days of the index event, for k=0, 1, 2, …, in the aCRT ON arm and aCRT OFF arm, respectively. The rate ratio (RR) for aCRT treatment is the ratio of the number of all-cause re-admissions after a HF admission within 30 days of the index event per patient year, in the aCRT ON arm relative to that in the aCRT OFF arm. The corresponding null and alternative hypotheses are:

H_0 : RR = 1

H_A : RR ≠ 1

Endpoint Definition

For this secondary objective, an index event must satisfy the following conditions:

- It is a heart failure admission;
- It is not within 30 days of the discharge from a previous index event;

Analysis Method

The analysis unit is patient. A subject could have multiple HF admissions that are index events. The total number of 30-day all-cause readmissions is aggregated over all index events and summarized into a single count per patient. The 30-day all-cause re-admission rate after an index HF hospitalization will be compared between aCRT ON and aCRT OFF arm using a Negative Binomial model with random site effect and NYHA at enrollment as a covariate. An offset based on the total follow-up experience of all patients will be included. If the last heart failure admission occurred after 30 days of the discharge from the previous index event, however, the patient died or exited within the 30 days after the last heart failure admission, the total follow-up time will count up to one day before the last heart failure admission. The SAS code will be similar to:

```
proc glimmix data=HFIndexReadmit;
class site rand NYHA;
model nReAdmit30Days = rand NYHA / dist=negbin link=log
                                   offset=ln_followyrs;
random intercept / subject=site;
run;
```

Missing Data

Since the endpoint of this objective is a variable for counts starting from 0, missing data is not expected.

Determination of Patients/Data for Analysis

The analysis follows the ITT principle and includes all the randomized subjects.

4.4.8 Secondary Objective #8: Healthcare System Cost Effectiveness

The secondary objective #8 is to assess cost-effectiveness of CRT devices with the aCRT algorithm relative to traditional CRT devices. Per CIP, analysis method for this objective will be defined in the Health Economic Analysis Plan. Additional statistical analyses, if any, that are needed to assist with the health economic analysis will be listed in the Appendix of this SAP. However the relevant results including those from health economic analysis will not be included in any DMC report or the Final Report.

4.5 Ancillary Objectives

4.5.1 [redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted] The [redacted]

[redacted]

[redacted]

[redacted]

[redacted]

5 APPENDIX

5.1 Sample Size Estimation by SAS: SEQDESIGN code and output

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

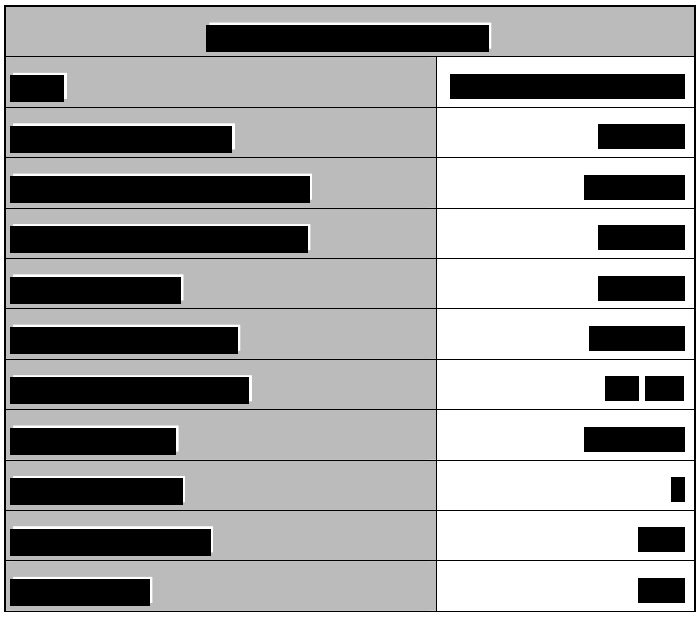
[redacted]

Design Information	
Null Ref ASN (	

[redacted]

[redacted]

[redacted]



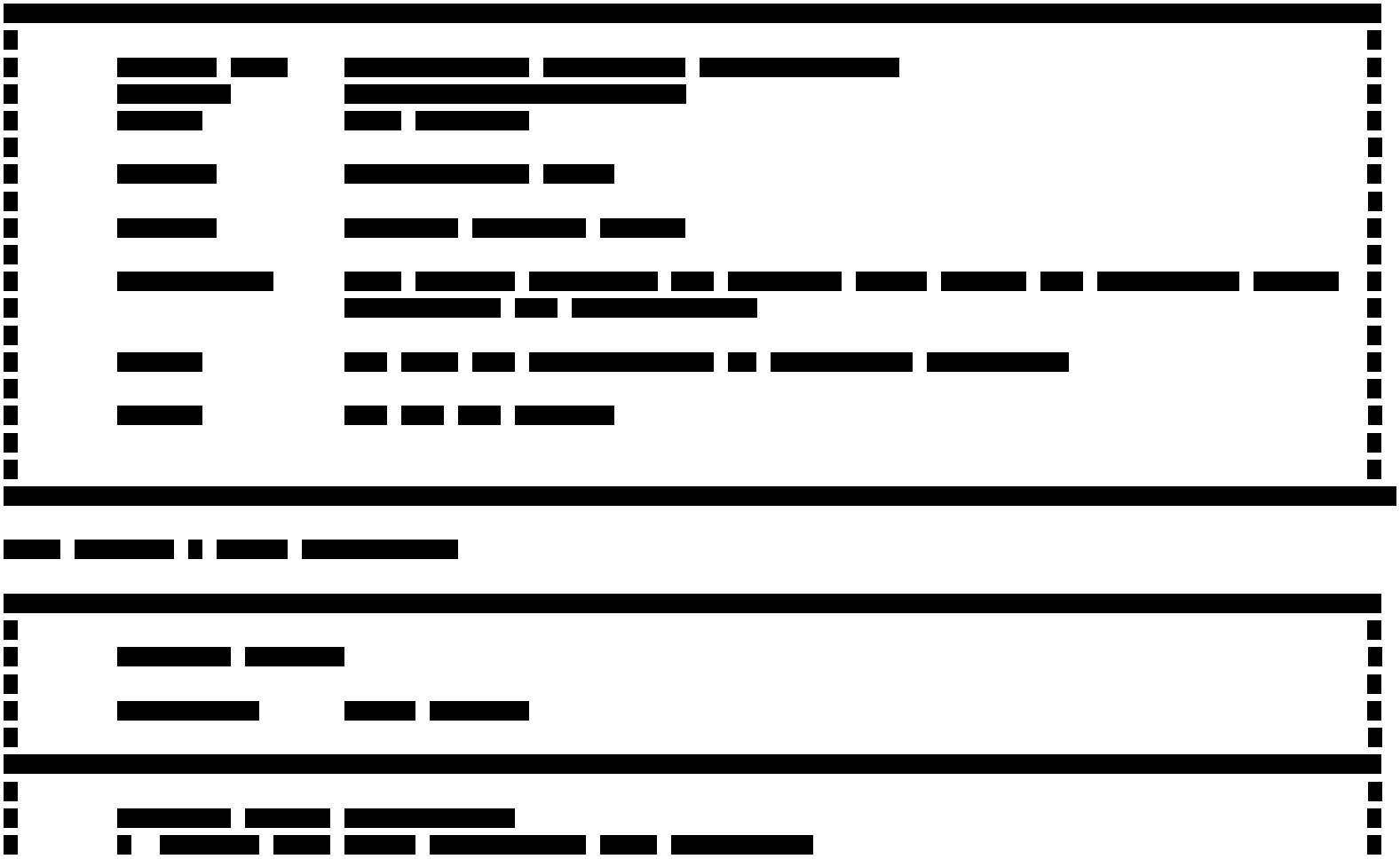
[redacted]

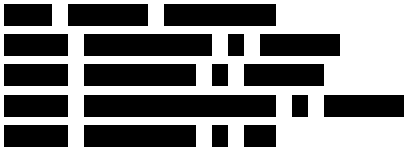
[redacted]

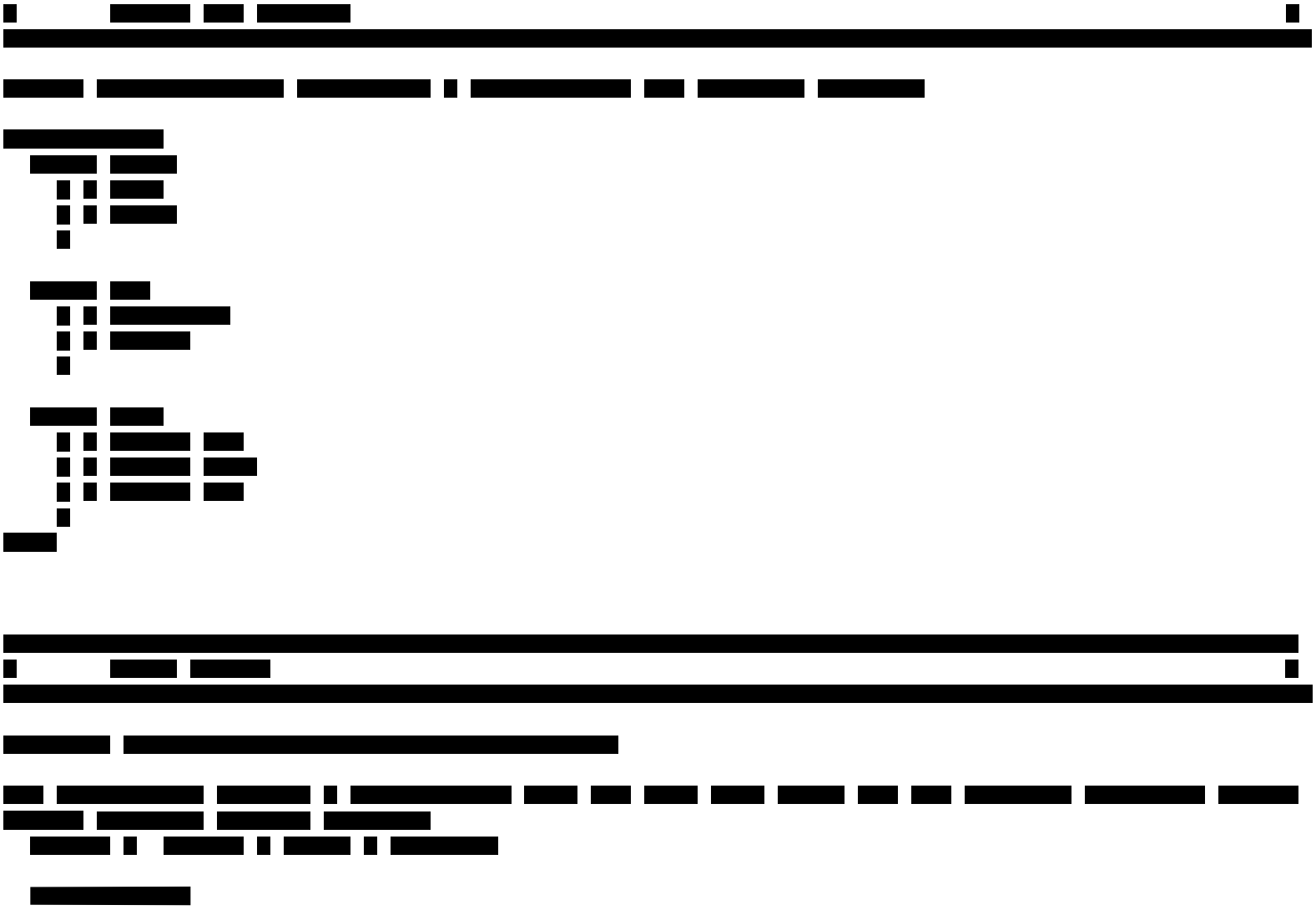
[redacted]

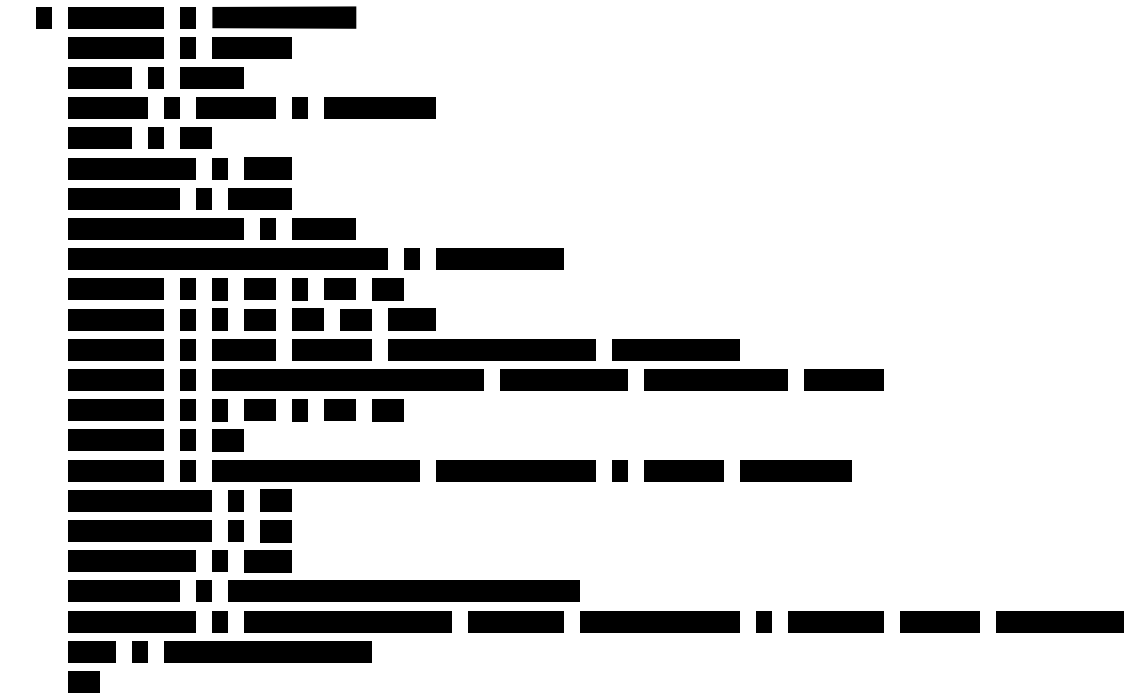
5.2 Sample Size Simulation by SAS

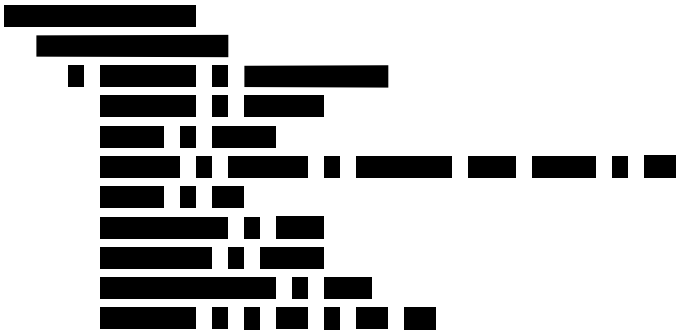
The following SAS codes were used to simulate events and estimate power in an event-driven fashion:



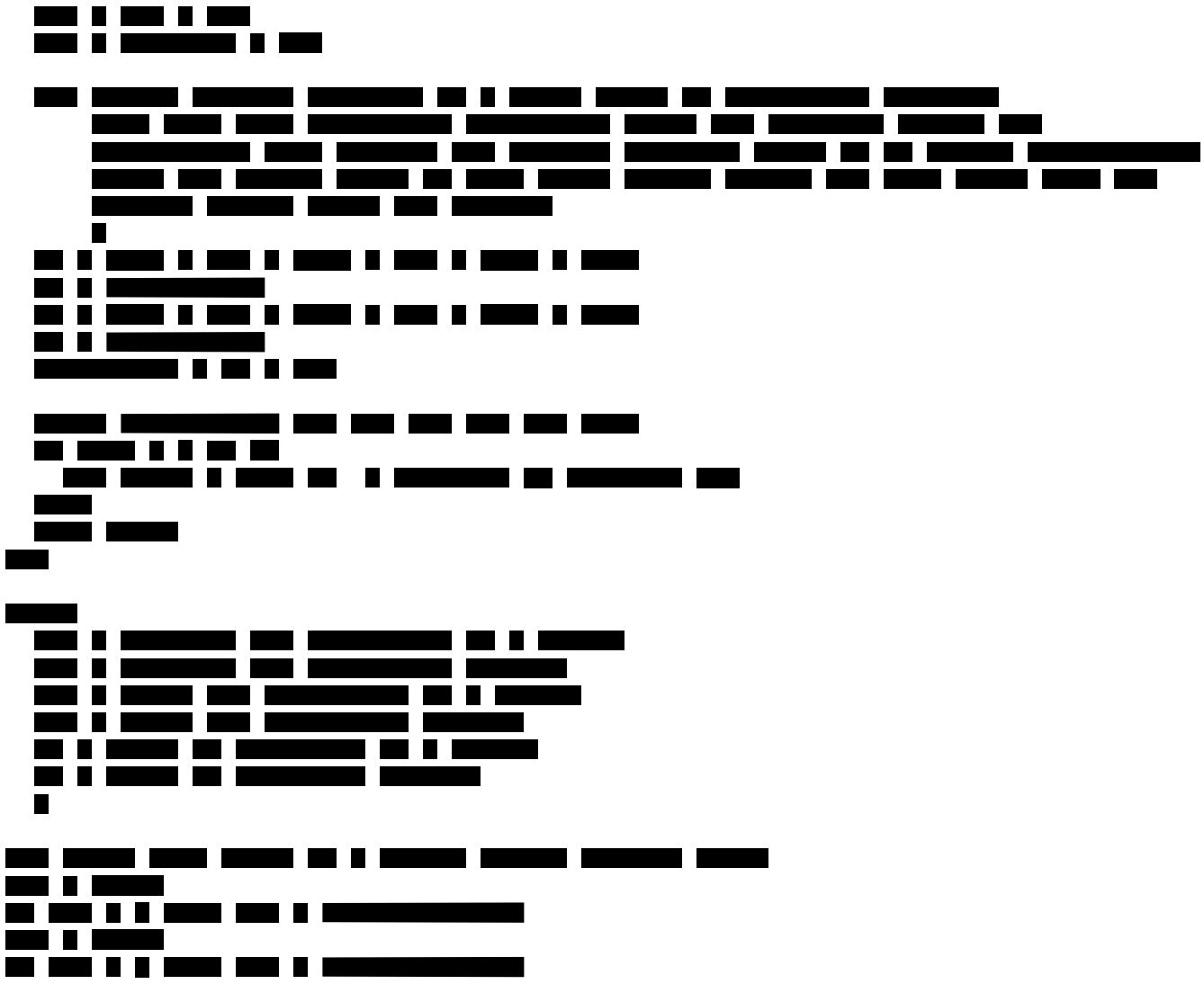


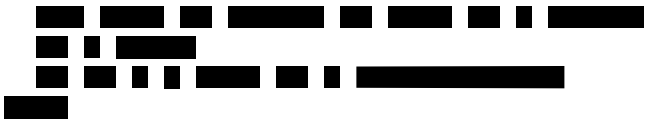


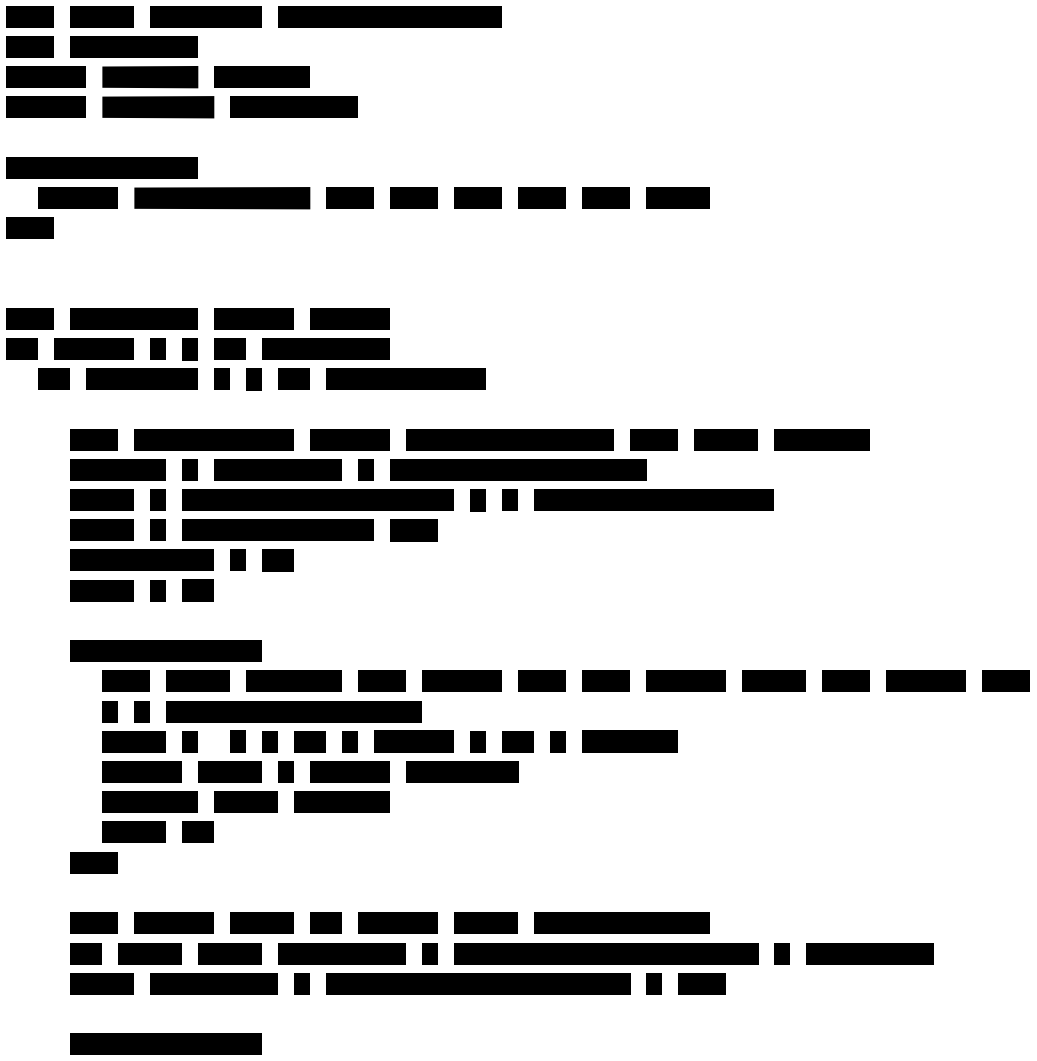


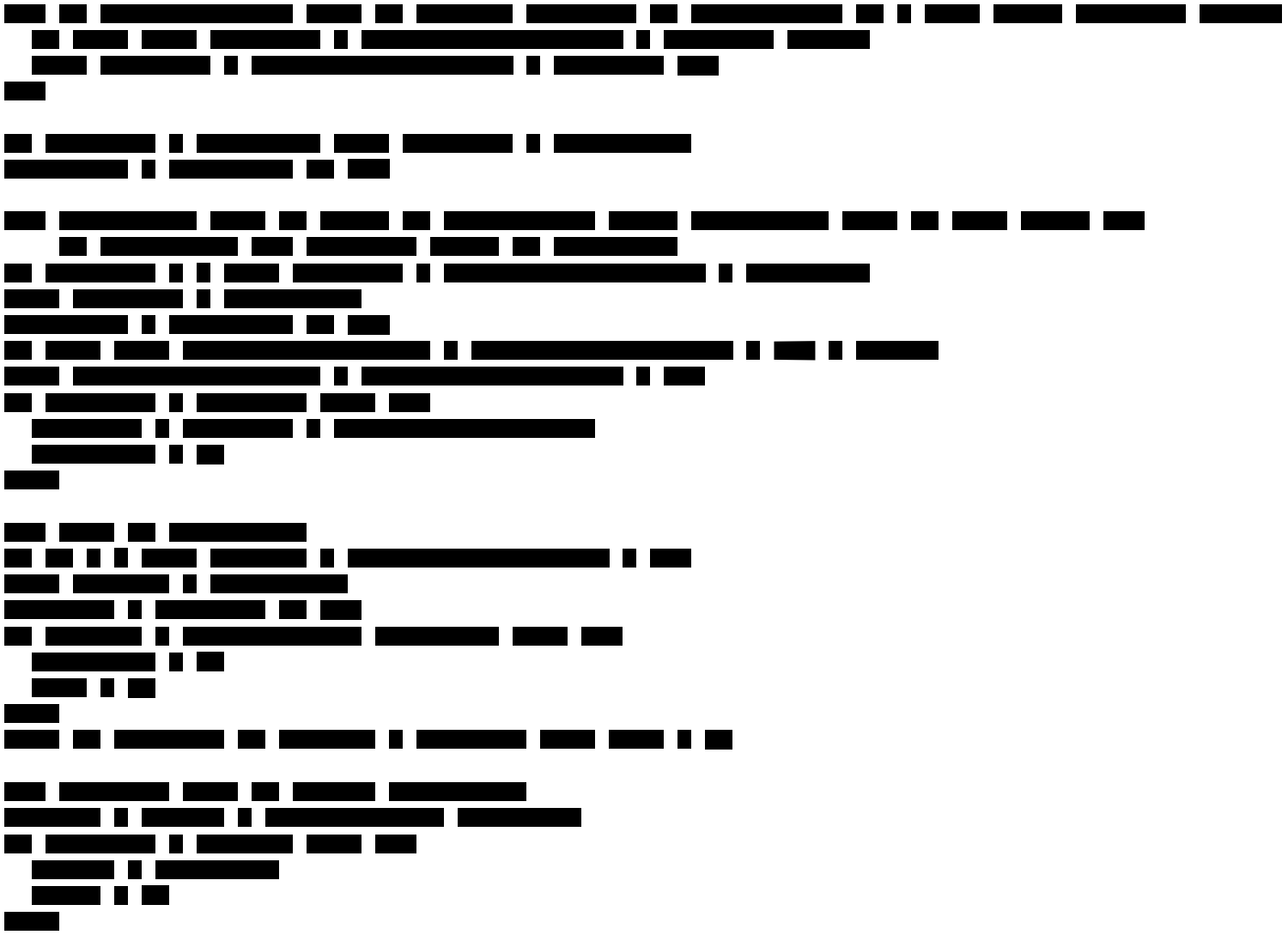


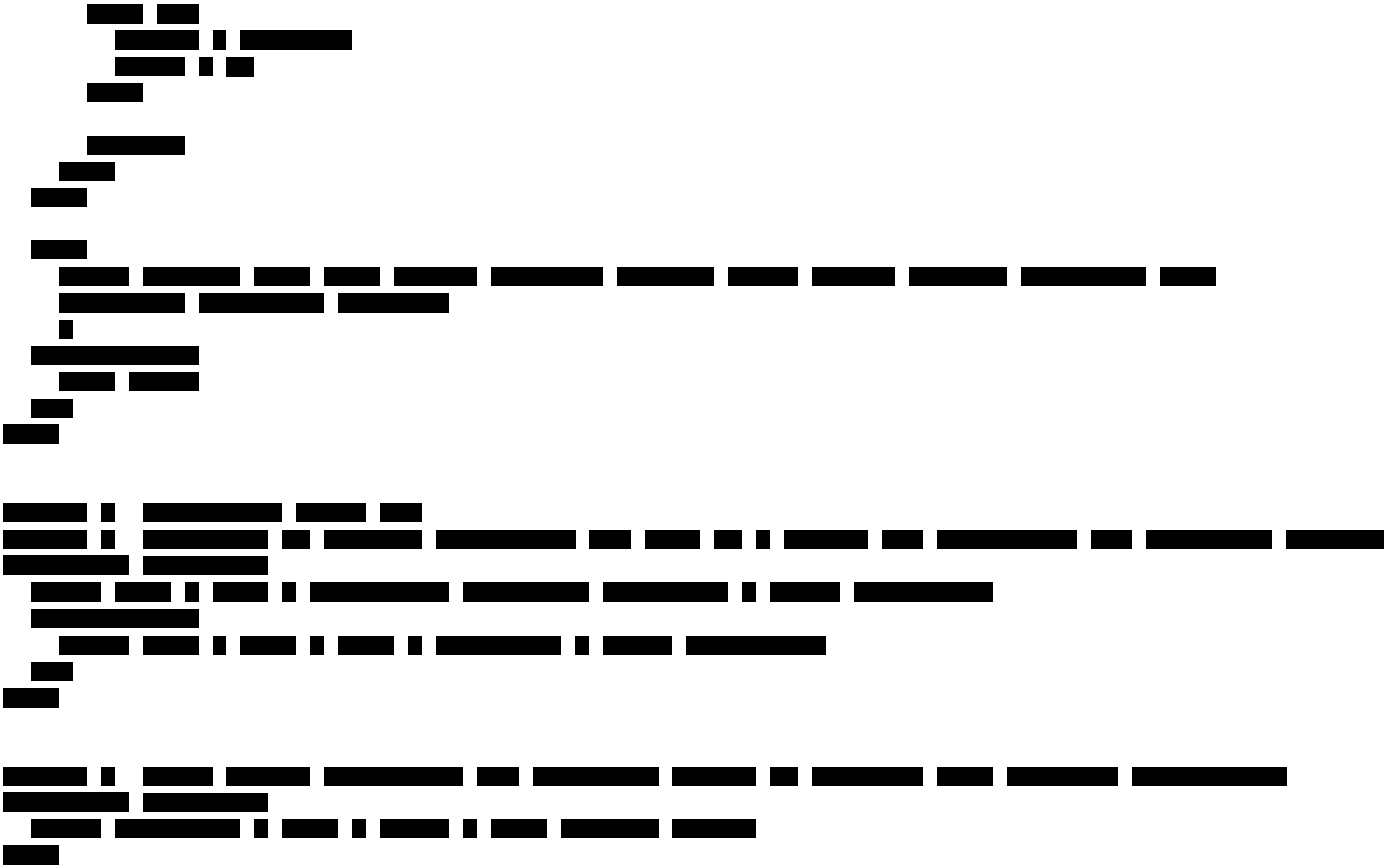
[redacted]

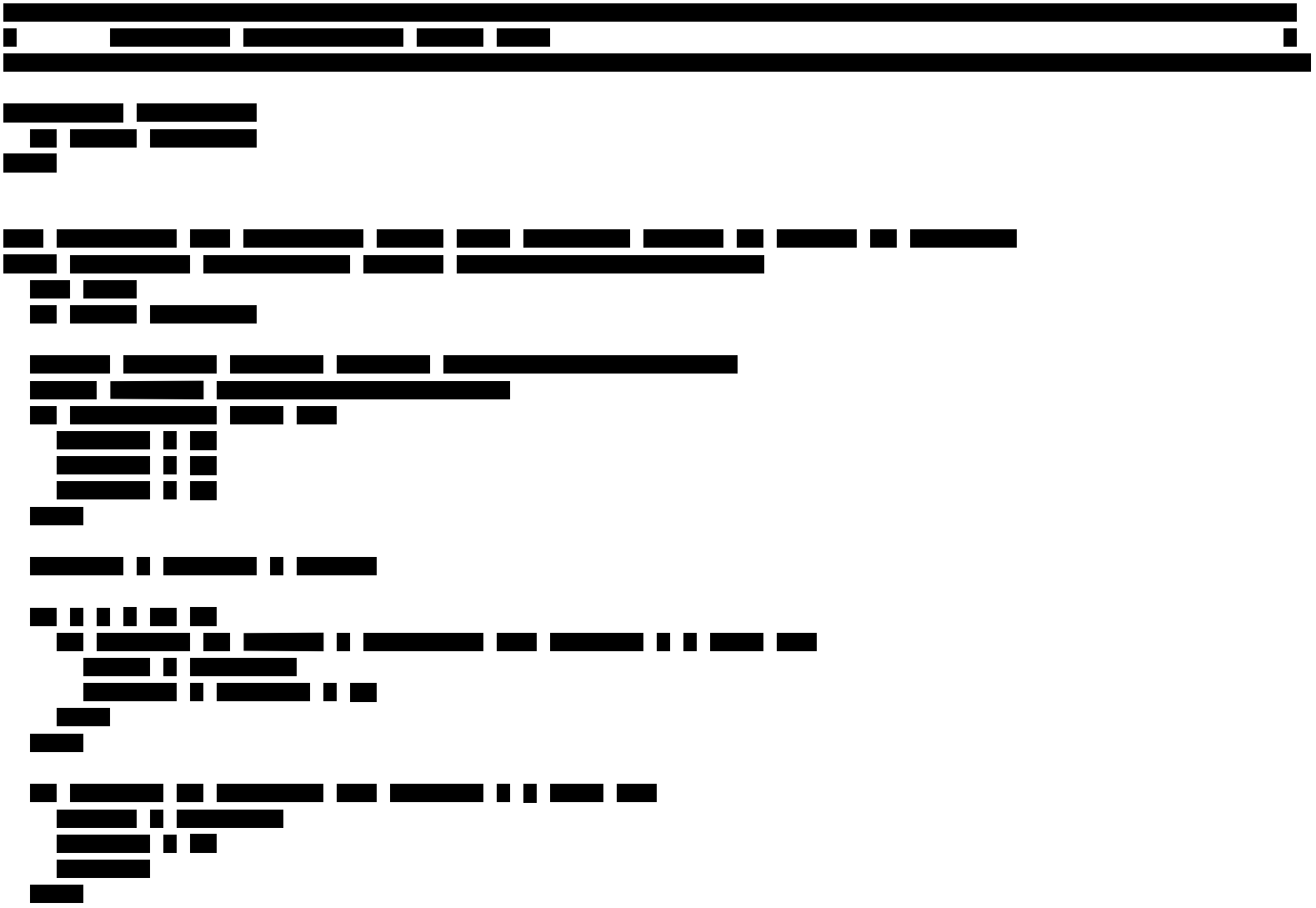




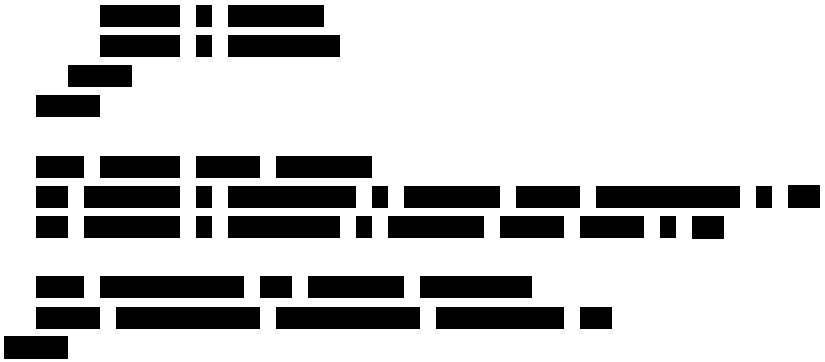


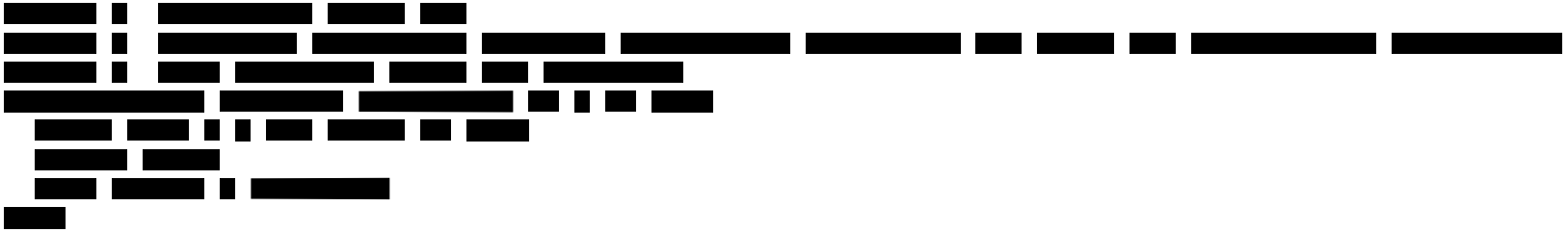




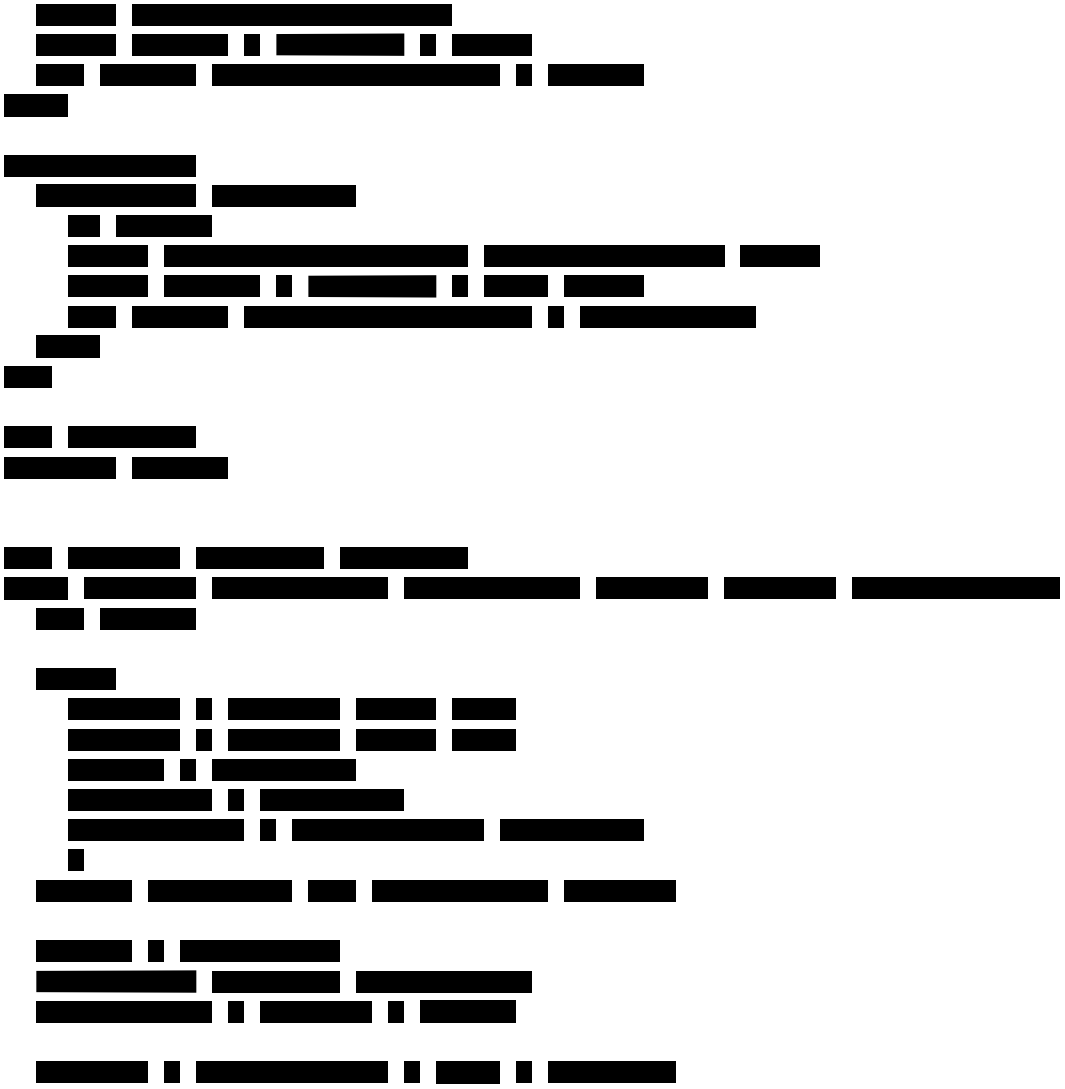


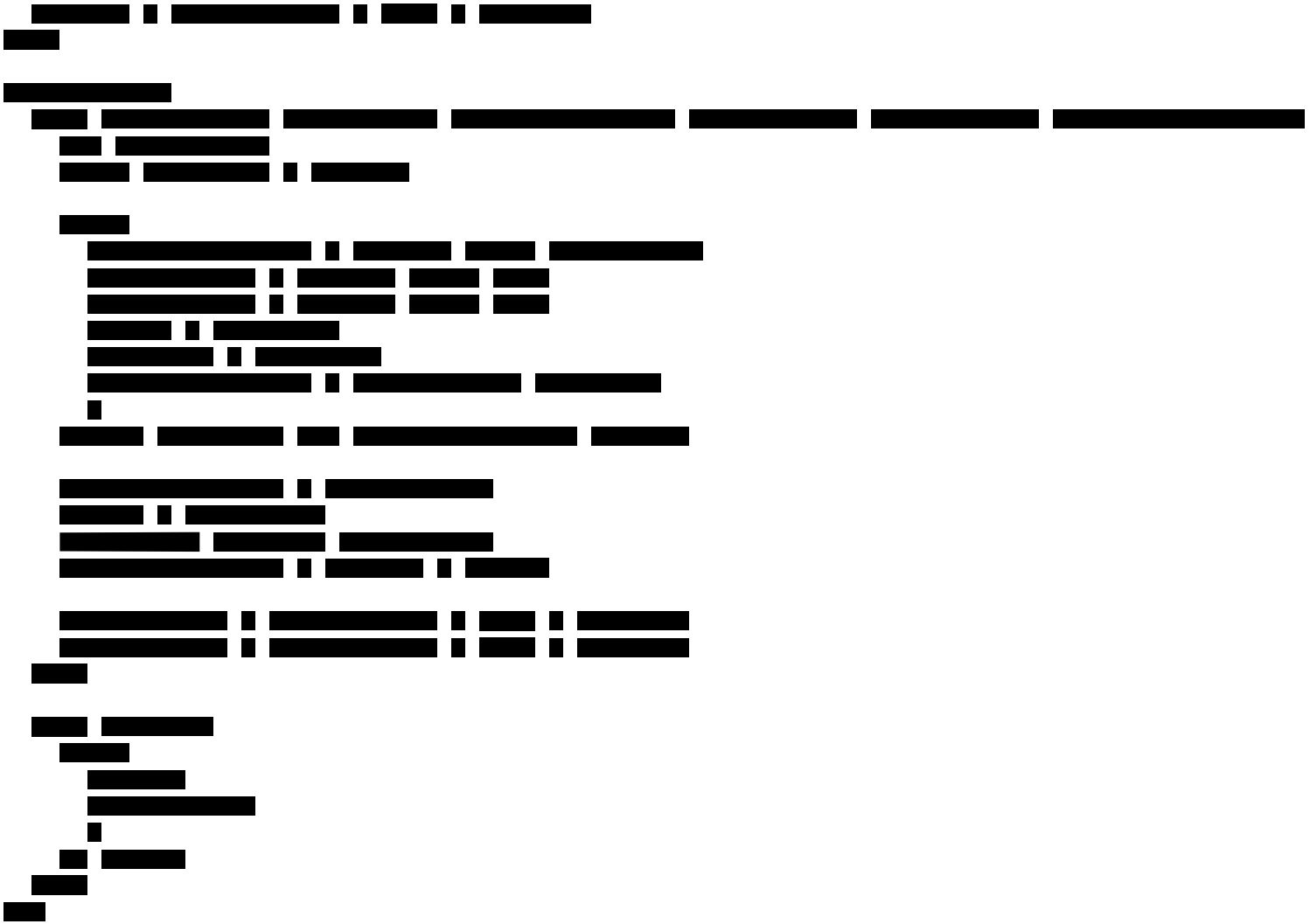
[redacted]





[redacted]





5.3 References

[1] Fisher LD, Dixon DO, Herson J, Frankowski RK, Hearron MS, Peace KE. Intention to treat in clinical trials. In: Peace KE, editor. Statistical issues in drug research and development. New York: Marcel Dekker; 1990. pp. 331–50.

[2] Lan KKG and DeMets DL. Discrete Sequential Boundaries for Clinical Trials. *Biometrika*, 70 (1983), 659–663.

[3] O'Brien PC and Fleming TR. A Multiple Testing Procedure for Clinical Trials. *Biometrics*, 35 (1979),549–556.

[4] Glimm E, Maurer W and Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. *Statistics in Medicine* 2009; **29**:219-28.

[5] Hommel G. A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*. 1988; 75:383-386.

[6] U.S. Department of Health and Human Services Food and Drug Administration Center for Biologics Evaluation and Research (CBER) Center for Drug Evaluation and Research (CDER) Center for Devices and Radiological Health (CDRH). Guidance for clinical trial sponsors – establishment and operation of clinical trial data monitoring committee. March 2006 OMB Control No. 0910-0581.

[7] Glidden DV and Vittinghoff E. Modelling clustered survival data from multicenter clinical trials. *Statistics in Medicine*. 2004; 23(3):369-388.

[8] Pocock ST, Wang D, Pfeffer MA, Yusuf S, McMurray JJV, Swedberg KB, Ostergren J, Michelson EL, Pieper, KS, and Granger CB on behalf of the CHARM investigators. Predictors of mortality and morbidity in patients with chronic heart failure. *European Heart Journal*. 2006; 27(1):65-75.

[9] Conraads VM, Spruit MA, Braunschweig F, Cowie MR, Tavazzi L, Borggrefe M, Hill MR, Jacobs S, Gerritse B, van Veldhuisen DJ. Physical activity measured with implanted devices predicts patient outcome in chronic heart failure. *Circulation: Heart Failure*. 2014; 7(2):279-287.

[10] Yuan Y. Group sequential analysis using the new SEQDESIGN and SEQTEST procedures. SAS Institute Inc.; 2009, paper 211.

[11] Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. New York: Chapman & Hall; 2000.

[12] The SEQDESIGN Procedure: applicable two-sample tests and sample size computation. SAS/STAT® 9.2 User's Guide.

[13] Dmitrienko A and d'Agostino R. Traditional multiplicity adjustment methods in clinical trials. *Statistics in Medicine*. 2013; 32:5172-218.

[14] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine*. 2007; 26:2389-2430.

[15] Dignam JJ, Zhang Q, Kocherginsky MN. The use and interpretation of competing risk regression models. *Clinical Cancer Research*. 2012; 18(8):2301-2308.

[16] Fine J and Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999; 94:496-509.

[17] Cook RJ and Lawless JF. Analysis of repeated events. *Statistical Methods in Medical Research*. 2002; 11:141-166.

[18] Castaneda, J and Gerritse, B, Appraisal of Several Methods to Model Time to Multiple Events per Subject: Modelling Time to Hospitalizations and Death. *Revista Columbiana de Estadistica* 2010; 33: 43-61.

[19] Packer M. Proposal for a new clinical end point to evaluate the efficacy of drugs and devices in the treatment of chronic heart failure. *Journal of Cardiac Failure*. 2001; 7(2):176-182.

[20] Green CP, Porter CB, Bresnahan DR and Spertus JA. Development and Evaluation of the Kansas City Cardiomyopathy Questionnaire: A New Health Status Measure for Heart Failure. *JACC* 2000; 35: 1245-1255.

[21] Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York.

[22] Schafer, J. L. (1997), *Analysis of Incomplete Multivariate Data,* New York: Chapman and Hall.