



| |
|--|
| NCT Number: NCT02224755 |
| MOMENTUM 3 |
| Multi-Center Study of MagLev Technology in Patients Undergoing MCS Therapy with HeartMate 3™ |
| Version 2.0 |
| Date: 22-JAN-2016 |

Sponsor

Abbott
6035 Stoneridge Drive
Pleasanton, CA 94588
USA

Title: HeartMate III IDE Trial Statistical Analysis Plan

Device: HeartMate III Ventricular Assist Device

Sponsor: Thoratec Corporation

Date of Report: January 22, 2016

Author(s) of Report: Poornima Sood MD
Sr. Director, Clinical Affairs
Thoratec Corporation

Gerald Heatley MS
Director, Biostatistics and Data Management
Thoratec Corporation

| | | |
|----------|---|----|
| 1 | TABLE OF CONTENTS | |
| 1 | TABLE OF CONTENTS | 2 |
| 2 | Abbreviations and Definitions | 4 |
| 3 | Introduction | 4 |
| 4 | Modification to the Protocol..... | 4 |
| 5 | Study Objectives and Endpoints | 4 |
| 5.1 | Study Objective | 4 |
| 5.2 | Primary Study Endpoints | 4 |
| 5.2.1 | Short Term Indication | 4 |
| 5.2.2 | Long Term Indication | 4 |
| 5.3 | Secondary Endpoints | 4 |
| 6 | Study Population:..... | 5 |
| 7 | Study Plan..... | 6 |
| 8 | Statistical Procedures: | 6 |
| 8.1 | Analysis Population | 7 |
| 8.1.1 | Intent-to-treat: All Randomized Subjects | 7 |
| 8.1.2 | As Treated: All Treated Subjects | 7 |
| 8.1.3 | As Randomized: All Treated Subjects | 7 |
| 8.2 | Study Hypothesis | 7 |
| 8.2.1 | Short Term Support: | 7 |
| 8.2.2 | Long Term Support:..... | 8 |
| 8.3 | Randomization | 9 |
| 8.4 | Sample Size | 9 |
| 8.4.1 | Assumptions | 9 |
| 8.4.2 | Short Term Indication | 9 |
| 8.4.3 | Long Term Indication..... | 9 |
| 8.5 | Early Safety Assessment | 10 |
| 8.6 | Analysis of Primary Endpoint | 10 |
| 8.6.1 | Primary Endpoint Stratified by Components of the Composite Endpoint | 10 |
| 8.6.2 | Effect of Site Bias on the Primary Endpoint..... | 11 |
| 8.6.3 | Unblinded Interim Efficacy Analysis (Adaptive Design)..... | 11 |
| 8.7 | Subgroup Analysis | 13 |
| 8.8 | Analysis of Survival and Subject Outcome..... | 14 |
| 8.9 | Analysis of Adverse Events | 14 |
| 8.10 | Analysis of Device Malfunctions..... | 14 |
| 8.11 | Analysis of Pre-Implant Data..... | 15 |
| 8.12 | Analysis of Implant and Discharge Data | 15 |
| 8.13 | Analysis of Secondary Endpoints..... | 15 |
| 8.13.1 | Pump Hemodynamics..... | 15 |
| 8.13.2 | Laboratory values | 15 |

| | | |
|----------|--------------------------|--|
| 1 | TABLE OF CONTENTS | |
| | 8.13.3 | Rehospitalization 15 |
| | 8.13.4 | Reoperations 15 |
| | 8.14 | Analysis of Functional Status 15 |
| | 8.14.1 | NYHA..... 15 |
| | 8.14.2 | Six Minute Walk Test..... 16 |
| | 8.15 | Analysis of Quality of Life 16 |
| | 8.15.1 | EQ-5D-5L..... 16 |
| | 8.15.2 | KCCQ 16 |
| | 8.16 | Powered secondary analysis:..... 16 |
| 9 | REFERENCES | 17 |

2 ABBREVIATIONS AND DEFINITIONS

- a. BSA = Body Surface Area
- b. LVEF = Left Ventricular Ejection Fraction
- c. CI = Cardiac Index
- d. LVAS = Left Ventricular Assist System
- e. LVAD = Left Ventricular Assist Device
- f. IABP= Intra-aortic Balloon Pump
- g. IRB
- h. FDA= Food and Drug Administration

3 INTRODUCTION

The HeartMate III LVAS represents the next generation of a mechanical support device. The HeartMate III LVAD is a centrifugal mechanically levitated heart pump designed for both short and long term support in patients afflicted with advanced, refractory left-ventricular heart failure. The purpose of this clinical investigation is to evaluate the safety and effectiveness of the HeartMate III LVAS. This multicenter study is designed to determine if the HeartMate III LVAS provides survival and quality of life benefits that are non-inferior to the HeartMate II LVAS. The trial will enroll patients under one inclusion and exclusion criteria who will be evaluated for a short and long term indication. The trial design also features a pre-defined and powered secondary endpoint to study pump reliability. The study will be conducted in compliance with the Declaration of Helsinki and all IRB and FDA regulatory requirements.

4 MODIFICATION TO THE PROTOCOL

5 STUDY OBJECTIVES AND ENDPOINTS

5.1 Study Objective

The primary objective of this study is to evaluate the safety and effectiveness of the HeartMate III LVAS by demonstrating non-inferiority to the HeartMate II when used for the treatment of advanced, refractory, left-ventricular heart failure.

5.2 Primary Study Endpoints

5.2.1 Short Term Indication

Composite of survival to transplant, recovery or 6 months of LVAD support free of debilitating stroke (Modified Rankin Score > 3) or reoperation to replace the pump.

5.2.2 Long Term Indication

Composite of survival to transplant, recovery or 24 months of LVAD support free of debilitating stroke (Modified Rankin Score > 3) or reoperation to replace the pump.

5.3 Secondary Endpoints

- i. Quality of Life as measured by the EuroQoL-5D-5L (EQ-5D-5L) and Kansas City Cardiomyopathy Questionnaire (KCCQ)
- ii. Functional status as measured by the Six Minute Walk Test (6MWT) and New York Heart Association (NYHA) classification
- iii. Frequency and incidence of pre-defined anticipated adverse events

- iv. Frequency and incidence of device malfunctions
- v. Frequency and incidence of all reoperations
- vi. Frequency and incidence of re-hospitalizations

6 STUDY POPULATION:

- a. Subjects with advanced refractory left ventricular heart failure who meet the following inclusion and exclusion criteria.
- b. Inclusion Criteria
 - i. Subject or legal representative has signed Informed Consent Form (ICF)
 - ii. Age \geq 18 years
 - iii. BSA \geq 1.2 m²
 - iv. NYHA Class III with dyspnea upon mild physical activity, or NYHA Class IV
 - v. LVEF \leq 25%
 - vi. Inotrope dependent OR
CI < 2.2 L/min/m², while not on inotropes and subjects must also meet one of the following:
 - On Optimal Medical management (OMM), based on current heart failure practice guidelines for at least 45 out of the last 60 days and are failing to respond
 - Advanced heart failure for at least 14 days AND dependent on IABP for at least 7 days
 - vii. Females of child bearing age must agree to use adequate contraception
- c. Exclusion Criteria
 - i. Etiology of HF due to or associated with uncorrected thyroid disease, obstructive cardiomyopathy, pericardial disease, amyloidosis or restrictive cardiomyopathy
 - ii. Technical obstacles which pose an inordinately high surgical risk, in the judgment of the investigator
 - iii. Existence of ongoing mechanical circulatory support (MCS) other than IABP
 - iv. Positive pregnancy test if of childbearing potential
 - v. Presence of mechanical aortic cardiac valve that will not be converted to a bioprosthesis or oversewn at the time of LVAD implant
 - vi. History of any organ transplant
 - vii. Platelet count < 100,000 x 10³/L (< 100,000/ml)
 - viii. Psychiatric disease/disorder, irreversible cognitive dysfunction or psychosocial issues that are likely to impair compliance with the study protocol and LVAS management
 - ix. History of confirmed, untreated AAA > 5 cm in diameter within 6 months of enrollment
 - x. Presence of an active, uncontrolled infection
 - xi. Intolerance to anticoagulant or antiplatelet therapies or any other peri/post-operative therapy the investigator will require based upon the patients' health status

- xii. Presence of any one of the following risk factors for indications of severe end organ dysfunction or failure:
 - 1. An INR \geq 2.0 not due to anticoagulation therapy
 - 2. Total bilirubin $>$ 43 $\mu\text{mol/L}$ (2.5 mg/dl), shock liver, or biopsy proven liver cirrhosis
 - 3. History of severe chronic obstructive pulmonary disease (COPD) defined by $\text{FEV}_1/\text{FVC} < 0.7$, or $\text{FEV}_1 < 50\%$ predicted
 - 4. Fixed pulmonary hypertension with a most recent PVR \geq 8 Wood units that is unresponsive to pharmacologic intervention
 - 5. History of stroke within 90 days prior to enrollment, or a history of cerebrovascular disease with significant ($> 80\%$) carotid artery stenosis
 - 6. Serum creatinine \geq 221 $\mu\text{mol/L}$ (2.5 mg/dl) or the need for chronic renal replacement therapy
 - 7. Significant peripheral vascular disease (PVD) accompanied by rest pain or extremity ulceration
- xiii. Patient has moderate to severe aortic insufficiency without plans for correction during pump implant
- xiv. Pre albumin $<$ 150 mg/L, or Albumin $<$ 30g/L (3 g/dL)
- xv. Planned Bi-VAD support prior to enrollment
- xvi. Patient has known hypo or hyper coagulable states such as disseminated intravascular coagulation and heparin induced thrombocytopenia
- xvii. Participation in any other clinical investigation that is likely to confound study results or affect the study
- xviii. Any condition other than HF that could limit survival to less than 24 months

7 STUDY PLAN

The study will be a prospective, multi-center, non-blinded, randomized, non-inferiority study comparing the HeartMate III LVAS to the HeartMate II LVAS. Short term use of the device will be evaluated when Subjects have been supported for 6 months, transplanted or explanted for recovery, whichever occurs first. Long term use of the device will be evaluated when Subjects have been followed for 24 months, transplanted or explanted for recovery, whichever occurs first.

- a. **Missing Values and Data Conventions:** Every effort will be made to collect all required data. Missing primary endpoints will be imputed using multiple imputation techniques. Missing secondary endpoints will not be imputed, except as described below.
- b. **Multiplicity Adjustments:** No adjustments for multiplicity will be made in the analysis of secondary endpoints. No statistical claims will be made concerning secondary endpoints.

8 STATISTICAL PROCEDURES:

The HM III study is a prospective, multicenter, non-blinded, randomized study to test the success of the HM III LVAS for short term hemodynamic support, such as a bridge to cardiac transplant (BTT) or myocardial recovery, or as long term support, such as destination therapy (DT).

Since the device must be surgically implanted, the investigator cannot be blinded to the treatment. Similarly, since the Subject must be trained on driveline care and responses to warning alarms, the Subject cannot be blinded to treatment. The study design will allow for a scientifically rigorous examination of the performance and safety of the device.

In general, continuous data will be presented as the number of Subjects, mean with standard deviation, and median with minimum and maximum values. Categorical data will be reported as frequencies and percentages. Adverse events will also be reported as rates per patient year. Only adverse events that occur after the start of the implant procedure will be analyzed. Survival data will be presented using the Kaplan-Meier product limit method, as well as the percentage of Subjects who successfully reach the pre-defined study primary endpoint.

Data will be analyzed using the intent-to-treat method (ITT) defined as all randomized Subjects. Every effort will be made to avoid cross-over but in the event they occur, data will also be analyzed “as randomized” and Per Protocol for efficacy analysis and “as treated” for safety analysis and all other secondary endpoints.

Statistical analysis will be performed using SAS version 9.1 or higher.

8.1 Analysis Population

8.1.1 Intent-to-treat: All Randomized Subjects

This is the primary analysis population for the primary endpoint. Subjects are analyzed under the treatment to which they were randomized.

8.1.2 As Treated: All Treated Subjects

This is the primary analysis population for the secondary endpoints. Subjects are analyzed under the treatment they actually receive.

8.1.3 As Randomized: All Treated Subjects

This population includes only the Subjects who receive the treatment they were randomized to.

8.1.4 Per Protocol:

As Randomized patients who meet all protocol inclusion and exclusion criteria and are free of major protocol violations. Major protocol violations include:

1. Patients who sign consent after randomization or patients who sign an unapproved consent form
2. Implantation of the VAD using a technique other than sternotomy
3. Missing > 2 pre-specified visits after initial discharge
4. Missing data that may affect the primary outcome:
 - i. Stroke reported but MRS never provided
 - ii. Transplant or explant prior to a study endpoint but the reason for the procedure is not reported

8.2 Study Hypothesis

8.2.1 Short Term Support:

The short term study endpoint is a composite of survival to transplant, myocardial recovery or 6 months of support (whichever occurs first) free of a debilitating

stroke (modified Rankin Score > 3) or reoperation to replace the original pump.
A Subject will be considered a success if they are:

- Electively transplanted or explanted for myocardial recovery prior to 6 months or
- Alive at 6 months, and
 - Have not experienced a stroke with a modified Rankin Score > 3, and
 - Have not received a device replacement or exchange, and
 - Have not received an urgent transplant due to a LVAS malfunction.

A Subject will be considered a failure if they:

- Expire prior to 6 months, or
- Experience a stroke with a modified Rankin Score > 3 prior to 6 months, or
- Have a device replaced or exchanged or deactivated for reasons other than myocardial recovery prior to 6 months, or
- Have received an urgent transplant due to a LVAS malfunction prior to 6 months, or
- Have withdrawn from the study for any reason prior to 6 months

HM III short term success rate will be compared to that of the HM II control in a non-inferiority manner. The null and alternative hypotheses are:

$$H_0: \pi_{HM III} \leq \pi_{HM II} - \Delta$$

$$H_A: \pi_{HM III} > \pi_{HM II} - \Delta$$

where $\pi_{HM III}$ and $\pi_{HM II}$ are the short term success rates of HM III and HM II, respectively, and where Δ is the non-inferiority margin.

8.2.2 Long Term Support:

The long term study endpoint is a composite of survival to transplant, myocardial recovery or 24 months of support (whichever occurs first) free of a debilitating stroke (modified Rankin Score > 3) or reoperation to replace the original pump.

A Subject will be considered a success if they are:

- Electively transplanted or explanted for myocardial recovery prior to 24 months or
- Alive at 24 months, and
 - Have not experienced a stroke with a modified Rankin Score > 3, and
 - Have not received a device replacement or exchange, and
 - Have not received an urgent transplant due to a device malfunction.

A Subject will be considered a failure if they:

- Expire prior to 24 months, or
- Experience a stroke with a modified Rankin Score > 3 prior to 24 months, or
- Have a device replaced or exchanged or deactivated for any reason other than myocardial recovery prior to 24 months, or
- Have received an urgent transplant due to a LVAS malfunction prior to 24 months, or
- Have withdrawn from the study for any reason prior to 24 months

The 24 month success rate of the HM III Subjects will be compared to the HM II control. The null and alternative hypotheses are:

$$H_0: \pi_{HM III} \leq \pi_{HM II} - \Delta$$
$$H_A: \pi_{HM III} > \pi_{HM II} - \Delta$$

where $\pi_{HM III}$ and $\pi_{HM II}$ are the long term success rates of HM III and HM II, respectively, and where Δ is the non-inferiority margin.

8.3 Randomization

Subjects will be randomized in a 1:1 fashion (1 HM III: 1 HM II). The randomization will be stratified by study center and blocked to maintain the 1:1 ratio over time. Randomization will be implemented through the Electronic Data Capture (EDC) system. Study centers will be allowed a maximum of 50 randomized Subjects. Subjects will be considered enrolled in the study upon randomization and will be included in the intent-to-treat analysis.

8.4 Sample Size

8.4.1 Assumptions

- Larger gap between the rotor and pump housing in the HM III may result in less thrombus than HM II
- Larger gap between the rotor and pump housing in the HM III may result in less pump replacement due to ingested thrombus than HM II
- HM III modular driveline may reduce pump replacements due to driveline damage or fatigue.

8.4.2 Short Term Indication

Based on a review of recent INTERMACS and Thoratec data, it is assumed that the HM II will achieve a composite success rate of 85% at 6 months. It is also assumed that the HM III will have a composite success rate of 87% due to less pump replacements at 6 months caused by thrombus or driveline issues. It will take 138 HM III and 138 HM II Subjects (276 total Subjects) to achieve 80% power to prove the HM III is non-inferior to HM II when the margin of non-inferiority is -10% ($=\Delta$ in the above null and alternative hypotheses) using the Farrington-Manning risk difference approach to non-inferiority at a one-sided alpha = 0.025.

INTERMACS HM II data from 26 sites likely to participate in the HM III IDE study was reviewed. Eight hundred and twenty (820) patients were implanted with the HM II in 2012 at these sites and 52 (6%) received a transplant or explant due to myocardial recovery prior to 6 months. In order to have sufficient data to evaluate the 6 month success rate, an additional 9 Subjects will be randomized per arm (6% of 138) to account for these early outcomes. This results in a total of 147 Subjects randomized in each arm (294 total Subjects).

8.4.3 Long Term Indication

Based on the results from the HM II Destination Therapy IDE study, it is assumed that 50% of the HM II Subjects will successfully achieve the composite primary endpoint. It is also assumed that the HM III will have a composite success rate of 55% due to less pump replacements at 24 months caused by thrombus or

driveline issues. It will take 174 HM III and 174 HM II Subjects (348 total Subjects) to achieve 80% power to prove the HM III is non-inferior to HM II when the margin of non-inferiority ($=\Delta$ in the above null and alternative hypotheses) is -10% using the Farrington-Manning risk difference approach to non-inferiority at a one-sided alpha = 0.025.

INTERMACS HM II data from 26 sites likely to participate in the HM III IDE study was reviewed. Eight hundred and twenty (820) patients were implanted with the HM II in 2012 at these sites and 52 (6%) received a transplant or explant due to myocardial recovery prior to 6 months. In order to have sufficient data to evaluate the 24 month success rate, an additional 9 Subjects will be randomized per arm (6% of 138) to account for these early outcomes. This results in a total of 183 Subjects randomized in each arm (366 total Subjects). It is assumed that 75 HM 3 patients will achieve 730 days of support on their original pump by the time the long-term cohort has completed the long-term follow up. This will provide sufficient data to evaluate the clinical reliability of the pump. If the HM 3 is proven to be non-inferior to the HM II for the long-term indication but 75 HM 3 pumps have not achieved 730 days of support, the submission of data to the Agency will be delayed until 75 pumps with durations of at least 730 days have been evaluated and included in the dataset.

8.5 Early Safety Assessment

The HM III IDE study will include an early safety assessment in lieu of a feasibility study. The first 10 HM III Subjects enrolled in the study will be included and their data analyzed when they have achieved 30 days of support. A table describing the 30 days status of the Subjects will be prepared. Adverse events will be presented as the percentage of Subjects who experience the event, the number of events and the event rate per 30 days. The data will be presented to the DSMB and FDA for a recommendation to continue the study and to expand to remaining study centers. All Subjects included in the early safety assessment will continue to be followed per protocol and will be included in the final Short Term and Long Term analysis.

8.6 Analysis of Primary Endpoint

The HM III will be considered non-inferior to the HM II for both short and long term indications if the lower two-sided 95% confidence limit of the risk difference in the composite success between treatment arms (HM III minus HM II) is greater than -10% ("negative 10%", where 10% is the non-inferiority Δ in the above null and alternative hypotheses). Once non-inferiority is inferred, the data will be analyzed for superiority at a one-sided 0.025 level of significance using closed testing methods via the z-test of proportions using the normal approximation to the binomial distribution. Since the short term and long term evaluations are two distinct endpoints, no adjustment for multiple comparisons is required. The long-term evaluation will only be made if the HM 3 is non-inferior to the HM II for the short-term indication, thus no adjustment to alpha will be made.

8.6.1 Primary Endpoint Stratified by Components of the Composite Endpoint

Differences in success rates between HM III and HM II will be performed for each component of the composite endpoint to evaluate if a single component is influencing the outcome. Specifically, for each component, two-sided 95%

confidence intervals of the differences and a z-test of proportions will be generated using the normal approximation to the binomial distribution.

8.6.2 Effect of Site Bias on the Primary Endpoint

In order to determine if a few superior investigational sites are influencing the primary endpoint results, a comparison of results across sites will be performed. Specifically, for each of the short term and long term outcomes, the significance of the treatment-by-site interaction effect will be assessed using logistic regression with the main effects for treatment and site, and with a treatment-by-site interaction effect. The treatment-by-site interaction effect will be tested at the 0.15 level of significance. A non-significant interaction or an interaction that is significant but only quantitative and not qualitative in nature will support the pooling of Subjects across sites for the primary analyses. Given that a number of sites will contribute only small numbers of Subjects, we will pool sites with less than 5 Subjects for the analysis.

8.6.3 Unblinded Interim Efficacy Analysis (Adaptive Design)

After 74 of the planned 147 patients per treatment group required for the analysis on the short term primary outcome are treated and followed for 6 months, an interim unblinded analysis comparing treatments on the 6-month short term primary outcome will be carried out. The short term primary outcome is the binary composite endpoint of survival to elective transplant if transplant is prior to 6 months, survival to myocardial recovery prior if recovery is prior to 6 months, or survival for 6 months of LVAD support free of a debilitating stroke (modified Rankin Score > 3) and free of reoperation to replace the pump. There will be no provision to stop the study at interim stage for overwhelming effectiveness and hence no adjustment of the significance level for the final analysis. The first purpose of the interim analysis is to calculate the power for non-inferiority, conditioned on the difference between treatments with respect to short term outcome rates and on the non-inferiority margin of 10% (or 0.10). Specifically, at the interim analysis, the conditional power for rejecting the following null hypothesis in favor of the alternative (i.e. for obtaining a non-inferiority conclusion at the final analysis for the short-term outcome) will be calculated:

$$H_0: \pi_{\text{HM III}} \leq \pi_{\text{HM II}} - \Delta$$

$$H_A: \pi_{\text{HM III}} > \pi_{\text{HM II}} - \Delta$$

where $\pi_{\text{HM III}}$ and $\pi_{\text{HM II}}$ are the short term success rates of HM III and HM II, respectively, and where Δ is the non-inferiority margin and is fixed at 10% (or 0.10). The conditional power will be calculated under the assumption that the interim observed estimate of the treatment difference is the true treatment difference. It will be calculated using the following formula as discussed in Lan and Wittes (1988):

$$1 - \Phi\left[\frac{Z_{1-\alpha} - B_\tau / \tau}{\sqrt{1 - \tau}}\right]$$

where

- a. $Z_{1-\alpha}$ is the $(1-\alpha)*100\%$ percentile of the standard normal distribution (i.e., the critical value used to assess non-inferiority at the final analysis at overall one-sided significance level α); here, with one-sided $\alpha = 0.025$, $Z_{1-\alpha}$ is set to 1.96.

- b. τ is the information fraction (= proportion of patients in the first interim analysis = r/M where r is the number of patients per group in the interim analysis and M is the planned number of patients per group for the final analysis). Here, it is expected that $r=74$, $M=147$, and $\tau=0.503$.
- c. $B_\tau = Z_\tau \sqrt{\tau}$ where Z_τ = the Farrington-Manning non-inferiority test statistic calculated on the interim observed data; specifically

$$Z_\tau = \frac{\hat{p}_{HMIII} - \hat{p}_{HMII} - (-\Delta)}{\sqrt{\frac{\tilde{p}_{HMIII}(1-\tilde{p}_{HMIII})}{n_{HMIII}} + \frac{\tilde{p}_{HMII}(1-\tilde{p}_{HMII})}{n_{HMII}}}}$$

where

- i. \hat{p}_{HMIII} = the interim observed estimate of the true success rate (π_{HMIII}) for HM III (=number of HM III subjects with success divided by number of HM III subjects in the interim analysis).
- ii. \hat{p}_{HMII} is the similarly defined interim observed estimate of the true success rate (π_{HMII}) for HM II.
- iii. \tilde{p}_{HMIII} and \tilde{p}_{HMII} are the interim maximum likelihood estimates of π_{HMIII} and π_{HMII} calculated under the above null hypothesis, as shown in Farrington and Manning (1990).
- iv. n_{HMIII} and n_{HMII} are the sample sizes used for the interim analysis from HM III and HM II, respectively.
- v. Δ is the non-inferiority margin of 0.10.

- d. Φ is the cumulative distribution function of the standard normal distribution.

Following the “promising zone” algorithm in Chen, Demets, Lan (2004), if the conditional power is <50% or >80% at the interim stage, the study will continue as is (there will be no stoppage of the study for futility nor will there be a sample size increase). If the conditional power is between 50-80% (the promising zone), the sample size will be re-estimated to achieve conditional power of 80% for the 6-month short term endpoint. . The total revised sample size per group, M' required to achieve a conditional power of 80% to assess non-inferiority at the final analysis using a non-inferiority margin of 0.10 (or 10%) is found by solving the following equation for M' (this formula is from Wang et al (2002), but it can also be derived algebraically from the above Lan and Wittes (1988) formula for conditional power by setting the above conditional power formula to 0.80 and solving for M'):

$$Z_{0.20} \sqrt{1 - \frac{r}{M'}} + Z_\tau \sqrt{\frac{M'}{r}} - Z_{1-\alpha} = 0$$

where $Z_{0.20}$ is the 20th percentile of the standard normal distribution (= -0.84) and where all other variables in this equation are defined above. This formula assumes the interim observed effect size between treatments is the true effect size between treatments.

Note that the final sample size for the 6-month short-term outcome will NOT be decreased below the currently planned final sample size for the 6-month short-term outcome. According to Chen, Demets and Lan (2004), in order to maintain Type I error at the nominal level, the sample size will not be increased to beyond a percentage of the original sample size as determined by the bound R , where R satisfies:

$$\frac{\sqrt{1+R}(\sqrt{1+R}-1)}{\sqrt{(1+R)-t}} = \left(\frac{|Z_{0.20}|}{Z_{1-\alpha}} \right).$$

With $|Z_{0.20}| = 0.84$ and $Z_{1-\alpha} = 1.96$, it can be shown that $R = 1.513$, which means the sample size can increase up to 151.3% of the original planned sample size (up to 369 per group) without inflating Type I error.

Also, any sample size increase for the 6-month short term outcome will also be applied to the final sample size used to assess the long-term outcome (there will be no interim unblinded analysis on the long-term outcome). The following table displays the sample size scenarios for each of the long-term and short-term outcome under various situations:

| Endpoint | Sample Size Per Group at Interim | Sample Size Per Group at Final Analysis | |
|--------------------|---|--|--|
| | | In the Event of No Sample Size Increase at Interim | In the Event of a Sample Size Increase of X at Interim |
| 6-month Short Term | After 74 patients per group are treated and followed for 6 months | After 147 patients per group are treated and followed for 6 months | After 147 + X patients per group are treated and followed for 6 months |
| 24-month Long Term | This endpoint will not be analyzed in the interim. | After 183 patients per group are treated and followed for 24 months. | After 183 + X patients per group are treated and followed for 24 months. |

The analyses on sample size re-estimation will be carried out by the independent statistician (non-voting) and presented only to the DSMB by the independent biostatistician. The DSMB will inspect the results and inform the sponsor of their final decision, without necessarily stating the reason for the decision. E.g., the DSMB will inform the sponsor “Continue the study as is” without informing the sponsor as to whether the reason is because the conditional power is <50% or because the conditional power is >80%; if the DSMB recommends an increase in sample size, the DSMB will not give a reason for the increase or any details behind the recommendation until the study is complete.

8.7 Subgroup Analysis

Once the analyses comparing the treatment arms are complete, a series of subgroup analyses will be performed, assessing treatment difference within each subgroup. Each subgroup will be evaluated for the primary composite endpoint, survival, adverse events,

device malfunction, quality of life and functional status, as described above. Subgroups will include but may not be limited to:

- Gender: Males vs Females
- Race: Caucasian/White vs African-American/ Black vs Other
- Age: 18 – 59 vs 60 – 75 vs > 75 years
- Intended Use at implant as defined by Appendix 6
- VO2 max
- INTERMACS Profile
- A detailed evaluation of patients who survive to 730 days or more including the influence of the site, age, and INTERMACS profile
- Learning Curve Analysis: The first 2 patients implanted with the HeartMate 3 in centers with 4 or more HeartMate 3 patients will be compared to the later HeartMate 3 patients. Comparison will include survival and adverse events.
- Time Dependent Analysis: Patients implanted during the first 7 months of the trial will be compared to patients implanted after 7 months. Analysis will include survival and adverse events

The purpose of the subgroup analyses is not to reach a statistically significant result within each subgroup, but rather to assess consistency of treatment difference across subgroups.

8.8 Analysis of Survival and Subject Outcome

Overall survival will be assessed for each of the two treatments using the Kaplan-Meier product-limit method. Differences between treatments in survival distributions will be analyzed using a logrank test. Subjects surviving will be censored at last known follow-up time point.

A competing outcome graph will be prepared at 6 months for short term results and 24 months for long term results.

8.9 Analysis of Adverse Events

All pre-defined adverse events will be captured. Tables will be created for HM III and HM II AEs that show the by-treatment incidences of all adverse events and the by-treatment event rate per patient year of support. Serious adverse events (SAEs) will be analyzed in a similar manner as AEs. Differences in event rates between the treatment arms will be analyzed using Fisher's Exact test or Poisson regression, as appropriate.

8.10 Analysis of Device Malfunctions

All suspected HM III device malfunctions will be reported. Thoratec will ask that all explanted devices be returned for analysis. Data on device malfunctions will be analyzed and tables will be created that report the following:

- Events that are confirmed by analysis of the device by Thoratec engineers
- The component of the device involved
- Days to the malfunction
- Action taken in response to the malfunction
- Reoperations due to malfunction
- Death due to malfunction

8.11 Analysis of Pre-Implant Data

Tables will be created to define the study population at baseline. Tables will include demographics, all laboratory assessments, all hemodynamic assessment, cardiac history, INTERMACS profile, and concurrent interventions (Cardiac Resynchronization Therapy (CRT), Automatic Internal Cardiac Defibrillator (AICD), IABP, Inotropes, etc). The intended use of the device at implant will also be collected, as defined by INTERMACS (Appendix 6). Baseline data will be compared between treatment groups using unpaired t-tests or Fisher's exact test as appropriate.

8.12 Analysis of Implant and Discharge Data

Time on cardiopulmonary bypass during implant surgery will be collected and reported as a median, quartiles and range. All concurrent procedures carried out during implant surgery will be reported. Length of Stay (LOS) will be defined as the time from implant to discharge. LOS will be reported as a mean with standard deviation, median, quartiles and range. The time on cardiopulmonary bypass and LOS will be compared between treatment groups using the Wilcoxon Rank Sum test.

8.13 Analysis of Secondary Endpoints

Secondary endpoints will each be tested at the two-sided 0.05 level of significance. There will be no adjustment for multiple comparisons across the secondary endpoints. There will be no imputation of missing data for the secondary endpoints.

8.13.1 Pump Hemodynamics

The mean flow and pump index (flow/BSA) with standard deviation for the HM III and HM II Subjects will be plotted over time. At each time point, treatments will be compared using the unpaired t-test.

8.13.2 Laboratory values

Mean laboratory values with standard deviations for HM III and HM II Subjects will be plotted over time. At each time point, treatments will be compared using the unpaired t-test.

8.13.3 Rehospitalization

Time to rehospitalization and the reason for rehospitalization will be reported. Time in and out of the hospital will be reported for the HM III and HM II Subjects. Treatments will be compared on time to re-hospitalization using the log-rank test. Subjects not re-hospitalized will be censored at last known follow-up.

8.13.4 Reoperations

Time to reoperation, frequency of reoperation, and the reason for the surgery will be reported for HM III and HM II Subjects. Treatments will be compared on time to re-operation using the log-rank test. Subjects not re-operated will be censored at last known follow-up.

8.14 Analysis of Functional Status

8.14.1 NYHA

The Subjects NYHA Functional Status will be assessed by an independent assessor at baseline and then at 3, 6, 12, 18, and 24 months. At each visit, treatments will be compared on NYHA functional status and on the change from baseline functional status using the Wilcoxon Rank Sum test. The Short Term indication will be limited to the 3 and 6 month assessments. The Long Term

indication will include all assessments until 24 months or outcome, whichever occurs first.

8.14.2 Six Minute Walk Test

Subjects may not be able to walk due to heart failure, especially at baseline. Subjects unable to walk due to heart failure will receive a score of 0 meters. For all other reasons for missing data the score will remain missing and not be included in the analysis. The Six Minute Walk test will be conducted at baseline and then at months 3, 6, 12, 18 and 24 post implant. Data will be analyzed using mixed modeling by comparing the distances walked over time to the baseline distance. The Short Term indication will be limited to the 3 and 6 month assessment. The long term indication will include all assessments until 24 months or outcome, whichever occurs first.

8.15 Analysis of Quality of Life

Quality of Life will be measured using the EuroQol (EQ-5D-5L) and the Kansas City Cardiomyopathy Questionnaire (KCCQ).

8.15.1 EQ-5D-5L

The EQ-5D-5L VAS and total score will be assessed at baseline and then at 3, 6, 12, 18, and 24 months. Data will be analyzed using mixed modeling by comparing the EQ-5D-5L score at each assessment interval to the baseline score. The Short Term indication will be limited to the 3 and 6 month assessments. The Long Term indication will include all assessments until 24 months or outcome, whichever occurs first. In addition, the percentage of each component of the EQ-5D-5L will be graphically presented over time.

8.15.2 KCCQ

The KCCQ score will be assessed at baseline and then at 3, 6, 12, 18, and 24 months. Data will be analyzed using mixed modeling by comparing the KCCQ score at each assessment interval to the baseline score. The Short Term indication will be limited to the 3 and 6 month assessments. The Long Term indication will include all assessments until 24 months or outcome, whichever occurs first.

8.16 Powered secondary analysis:

In addition to the primary outcome, the study will be powered to test if HM III pump reliability is superior to HM II by analyzing the incidence of pump replacements. Randomization described in Section 13.4 will continue beyond the Subjects needed to power the primary endpoint until a sufficient sample size has been enrolled to test the secondary endpoint. The null and alternative hypotheses are:

$$H_0: \pi_{HM\ III} \geq \pi_{HM\ II}$$

$$H_A: \pi_{HM\ III} < \pi_{HM\ II}$$

where $\pi_{HM\ III}$ is the HM III pump replacement rate and $\pi_{HM\ II}$ is the HM II pump replacement rate.

Based on data contained in Thoratec's device tracking database, 7% of the HM II Subjects receive a pump replacement by 24 months. If we assume that HM III pump replacements will be reduced to 3% at 24 months, then 1028 Subjects (514 per arm) will be needed to prove superiority with a power of 80% and alpha = 0.05 (2-sided).

Once the 366 Subjects needed for the long term indication are enrolled, Thoratec will continue to randomize 662 more Subjects (331 HM III and 331 HM II) for this secondary analysis. Data for this secondary analysis is not needed for the short or long term indications, but rather will be used to provide additional labeling information.

The Subjects will be followed for 24 months or to outcome, whichever occurs first, and analyzed using the Fisher's exact test. Treatments will also be compared on time-to-pump replacement using the log-rank test where Subjects without a replacement are censored at last known follow-up. Otherwise, there will be no imputation of missing data for this analysis.

9 REFERENCES

5. Farrington CP, Manning G. Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. *Stat Med* 1990;9:1447-54.
6. White paper: Wang C, Keller DS, Lan KKG (2002). Sample size re-estimation for binary data via conditional power. Joint Statistical Meetings – Biopharmaceutical Section.
7. Lan KKG and Wittes J. The B-Value: A tool for monitoring data. *Biometrics* 1998;44:579-585.
8. Chen YH, DeMets DL, Lan KKG. Increasing the sample size when the unblinded interim result is promising. *Stat Med* 2004; 23:1023-1038.

Summary of Changes to the HeartMate III IDE Trial Statistical Analysis Plan

| Protocol Version | Date | Revision Summary |
|------------------|----------|---|
| 1.0 | 10/14/14 | Original Protocol |
| 2.0 | 01/22/16 | <ul style="list-style-type: none"> • Updated <i>Analysis Population</i> Section (8.1) by adding definition of the <i>Per Protocol</i> population (sub-section 8.1.4), which are As Randomized patients who meet all protocol inclusion and exclusion criteria and are free of major protocol violations. • Updated <i>Sample Size - Long Term Indication</i> Section (8.4.3) by adding that 75 HM 3 patients that will achieve 730 days of support will be evaluated for the clinical reliability of the pump. If the HM 3 is proven to be non-inferior to the HM II for the long-term indication but 75 HM 3 pumps have not achieved 730 days of support, the submission of data to the Agency will be delayed until 75 pumps with durations of at least 730 days have been evaluated and included in the dataset. • Updated <i>Analysis of Primary Endpoint</i> Section (8.6) by adding the following statement “The long-term evaluation will only be made if the HM 3 is non-inferior to the HM II for the short-term indication, thus no adjustment to alpha will be made.” • Added a reference to Lan and Wittes (1988) for the conditional power calculation formula in the <i>Unblinded Interim Efficacy Analysis (Adaptive Design)</i> Section (8.6.3). • Added a reference to Chen, Demets, Lan (2004) to the “promising zone” algorithm in the <i>Unblinded Interim Efficacy Analysis (Adaptive Design)</i> Section (8.6.3). • Added a reference to Wang et al (2002) to the revised sample size calculation formula in the <i>Unblinded Interim Efficacy Analysis (Adaptive Design)</i> Section (8.6.3). Also added the following statement in addition to the aforementioned reference: “it can also be derived algebraically from the above Lan and Wittes (1988) formula for conditional power by setting the above conditional power formula to 0.80 and solving for M.” • The statement that the sample size will not be increased to >1000 per group in the <i>Unblinded Interim Efficacy Analysis (Adaptive Design)</i> Section (8.6.3) was updated with the details on calculation and references below: <ul style="list-style-type: none"> According to Chen, Demets and Lan (2004), in order to maintain Type I error at the nominal level, the sample size will not be increased to beyond a percentage of the original sample size as determined by the bound R, where R satisfies: $\frac{\sqrt{1+R}(\sqrt{1+R}-1)}{\sqrt{(1+R)-t}} = \left(\frac{ Z_{0.20} }{Z_{1-\alpha}}\right)$ <ul style="list-style-type: none"> With $Z_{0.20} = 0.84$ and $Z_{1-\alpha} = 1.96$, it can be shown that $R = 1.513$, which means the sample size can increase up to 151.3% of the original planned sample size (up to 369 per group) without inflating Type I error. • Added Learning Curve Analysis (the first 2 patients implanted with the HM 3 in centers with 4 or more HM 3 patients compared to later HM 3 patients for survival and adverse events), Time Dependent Analysis (patients implanted during the first 7 months of the trial will be compared to patients implanted after 7 months for survival and adverse events), a detailed evaluation of patients surviving to 730 days or more to the <i>Subgroup Analysis</i> Section (8.7). • Two additional references to Lan and Wittes and to Chen, DeMets, and Lan were added to the <i>Reference</i> Section (9). |