
Statistical analysis plan

This supplementary material has been provided by the authors to give readers additional information about their work.
Statistical Analysis Plan

Functional Impact of GLP-1 for Heart Failure Treatment

FIGHT

A randomized, double-blind, placebo-controlled study of high-risk patients with reduced ejection fraction and post-acute heart failure syndrome (AHFS) treated for six months post-discharge.

Protocol Date
October 2, 2013

Sponsor
National Heart, Lung and Blood Institute

Protocol Principal Investigators
Kenneth Margulies, MD and Thomas Cappola, MD
# Table of Contents

1. **Overview** .............................................................................. 3  
   1.1 **Synopsis** ........................................................................ 3  
   1.2 **Study Treatments** ............................................................. 3  

2. **Study Design** ...................................................................... 3  
   2.1 **Overview** ........................................................................ 3  
   2.2 **Randomization** ............................................................... 3  
   2.3 **Data Sources** ................................................................. 3  

3. **Analysis Population and Missing Data** ................................. 3  

4. **General Methodology** .......................................................... 4  

5. **Primary Endpoint** .............................................................. 4  

6. **Secondary Endpoints** .......................................................... 4  

7. **Tertiary Endpoints** ............................................................... 5  

8. **Primary Endpoint Description** ............................................. 5  
   8.1 **Primary Endpoint** ............................................................ 5  

9. **Secondary Endpoint Descriptions** ....................................... 6  
   9.1 **Secondary Endpoint #1** .................................................... 6  
   9.2 **Secondary Endpoint #2** ................................................... 7  
   9.3 **Secondary Endpoint #3** ................................................... 8  
   9.4 **Secondary Endpoint #4** ................................................... 8  
   9.5 **Secondary Endpoint #5** ................................................... 10  
   9.6 **Secondary Endpoint #6** ................................................... 10  

10. **Tertiary Endpoint Descriptions** .......................................... 11  
    10.1 **Tertiary Endpoint #1** ..................................................... 11  
    10.2 **Tertiary Endpoint #2** ..................................................... 11  
    10.3 **Tertiary Endpoint #3** ..................................................... 12  
    10.4 **Tertiary Endpoint #4** ..................................................... 13  
    10.5 **Tertiary Endpoint #5** ..................................................... 13  

11. **Safety Analyses** ............................................................... 14  
    11.1 **Safety Endpoint #1** ....................................................... 14  
    11.2 **Safety Endpoint #2** ....................................................... 14  

12. **Interim Analysis** ............................................................... 15  

13. **Subgroup of Interest** ........................................................ 15
1. Overview

1.1 Synopsis

The Functional Impact of Glucagon-like peptide-1 (GLP-1) for Heart Failure Treatment (FIGHT) trial is testing the efficacy and safety of therapy with subcutaneous (SQ) GLP-1 agonist in the post-AHFS discharge period in order to assess clinical stability over a six month treatment and follow-up period.

1.2 Study Treatments

Placebo or GLP-1 agonist liraglutide (0.6 mg SQ daily for 7 days, 1.2 mg SQ daily from day 7 through day 30, 1.8 mg from day 30 to day 180) will be administered daily by SQ injection, initiated at discharge and administered for six months. Study treatment is administered by daily injection at any time of day, independently of meals, and can be injected subcutaneously in the abdomen, thigh or upper arm.

2. Study Design

2.1 Overview

The FIGHT study is a randomized, double-blind, placebo-controlled study in high risk post-acute heart failure patients with reduced ejection fraction. A total of 300 patients will be enrolled in the trial.

The treatments in this study are blinded. Treatment is expected to last for 24 weeks.

The over-arching hypothesis is that, compared to placebo, therapy with the GLP-1 agonist liraglutide in the post-acute heart failure syndrome discharge period will be associated with greater clinical stability as assessed by a novel global rank endpoint, which is a composite of time to death, time to heart failure hospitalization, and time-averaged proportional change in NT pro BNP through 180 days.

2.2 Randomization

Patients are randomized in a 1:1 ratio to either the Liraglutide or Placebo. The randomization scheme consists of a permuted block design with stratification by clinical site and presence of Type II diabetes.

2.3 Data Sources

A database of case report form and biomarker core lab data will be created in Inform, and the data then transferred to SAS for analysis. The randomized treatment assignment will be provided through data provided by the Axcess system, an Almac Clinical Services web-based randomization system.

3. Analysis Population and Missing Data

All randomized patients will be included in the analysis population for assessing the primary, secondary and tertiary endpoints. However, as described in subsequent sections of this document, some patients may be excluded from certain analyses if key data elements are missing. With the extensive efforts being made in...
connection with the clinical sites to ensure data quality and completeness, it is expected that exclusion of patients for any endpoint analysis will be minimal.

There were a number of patients that had a left ventricular assist device (LVAD) placed or received a heart transplant (HTP) prior to the end of the study follow-up. Due to the affect these procedures would have on the heart failure symptoms and quality of life, any data collected post-LVAD or HTP will not be used in analyses.

The specific endpoint descriptions in Sections 8 through 10 describe the circumstances that would lead to a patient being excluded from a specific analysis.

4. General Methodology

Medians, 25th and 75th percentiles will be presented for continuous variables; the number and percentage of patients in each category will be presented for categorical variables. For all endpoints a p-value ≤0.05 will be considered statistically significant. Analyses will be performed using validated SAS software (SAS Institute, Inc, Cary, NC). Appropriate statistical models will be used to examine the effect of treatment with a GLP-1 agonist on both the primary, secondary, and tertiary outcomes in the study.

For the rank-based endpoints, a non-parametric testing strategy will be employed. For continuous endpoint variables, conventional general linear models will be used. For endpoints where the response is dichotomous (binary), the logistic regression model will be used. For time-to-event endpoints, the Cox regression model will be used.

5. Primary Endpoint

Primary Endpoint

#1: Global rank endpoint, which includes time to death, time to heart failure hospitalization, and time-averaged proportional change in NT pro BNP through 180 days

See Section 8 for a detailed description of the primary endpoint, including rules that will be followed for handling incomplete data.

6. Secondary Endpoints

Secondary Endpoints

#1: Change in cardiac structure and function from baseline to 180 days

#2: Change in 6 minute walk from baseline to 30, 90, and 180 days

#3: Change in Kansas City Cardiomyopathy Questionnaire (KCCQ) symptom scores from baseline to 180 days

#4: Individual components of the primary endpoint at 30, 90 and 180 days

#5: Number of combined events (death + heart failure (HF) hospitalization; death + HF hospitalization + emergency department (ED) visits)

#6: Global rank endpoint, which includes time to death, time to heart failure hospitalization, and change in NT pro BNP from baseline to 180 days
See Section 9 for a detailed description of each secondary endpoint, including rules that will be followed for handling incomplete data.

**7. Tertiary Endpoints**

**Tertiary Endpoints**

- **#1**: Change in AHFS biomarker panel from baseline to 30, 90 and 180 days
- **#2**: Change in glycosylated hemoglobin at 30, 90 and 180 days
- **#3**: Change in weight at 30, 90, and 180 days
- **#4**: Change in insulin resistance at 180 days
- **#5**: Change in fasting lipids at 180 days

See Section 10 for a detailed description of each tertiary endpoint, including rules that will be followed for handling incomplete data.

**8. Endpoint Descriptions**

**8.1 – Primary Endpoint**

**Endpoint Description**: Global rank endpoint, which includes time to death, time to heart failure hospitalization, and time-averaged proportional change in NT pro BNP through 180 days

**Response Variable Definition**: A rank score based on time to death (tier 1), time to adjudicated heart failure hospitalization (tier 2), and time-averaged proportional change in NT pro BNP through 180 days (tier 3). See the following table for a description of the coding and example scores. NT Pro BNP is assessed at baseline, 30, 90, and 180 days.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Example Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Death 1 day post randomization</td>
</tr>
<tr>
<td>2</td>
<td>Death 20 days post randomization</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td>X</td>
<td>End of first tier</td>
</tr>
<tr>
<td>X+1</td>
<td>Admission 1 day post randomization</td>
</tr>
<tr>
<td>X+2</td>
<td>Admission 20 days post randomization</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td>Y</td>
<td>End of second tier</td>
</tr>
<tr>
<td>Y+1</td>
<td>$\Delta = \text{maximum value}$</td>
</tr>
<tr>
<td>Y+2</td>
<td>$\Delta = \text{next smallest value}$</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td>max N</td>
<td>$\Delta = \text{smallest value}$</td>
</tr>
</tbody>
</table>

* Increases in time-averaged proportional change in NT pro BNP indicate a worse result. A day 180 change score greater than 0 ($\Delta > 0$) indicates a worse result. A day 180 change score smaller than 0 ($\Delta < 0$) indicates a better result. There are two ways that the area under the curve (AUC) can be calculated.
The proposed method uses the post-baseline NT pro BNP measures after correction for the baseline value.

The time-averaged proportional change in NT Pro BNP is defined as follows:

\[(30) \times \text{Day 30 value}/\text{Baseline value} + (60) \times \text{Day 90 value}/\text{Baseline value} + (90) \times \text{Day 180 value}/\text{Baseline value}.\]

The weighting for visits will be adjusted as needed based on incomplete biomarker data. For example, if a patient is missing the Day 90 value, the revised time-average proportional change in NT pro BNP would be defined as \((30) \times \text{Day 30 value}/\text{Baseline value} + (150) \times \text{Day 180 value}/\text{Baseline value}\).

An alternative definition that will be investigated but not used for the primary result utilizes the complete area under the curve, as follows:

\[
\left((30) \times \text{Day 30 value}/\text{Baseline value} + (60) \times \text{Day 90 value}/\text{Baseline value} + (90) \times \text{Day 180 value}/\text{Baseline value}\right)/2 + \left((30) \times \text{Baseline value}/\text{Baseline value} + (60) \times \text{Day 30 value}/\text{Baseline value} + (90) \times \text{Day 90 value}/\text{Baseline value}\right)/2.
\]

Additional Covariates: None

Handling of Dropouts and Missing Data:

If the patient was missing core lab NT Pro BNP, at a specific visit but did have local lab NT Pro BNP available, the local lab result will be used in place of the missing result in the calculations. In the event the patient did not die or have a heart failure re-admission and was lacking the baseline NT Pro BNP value, the average baseline NT Pro BNP value will be imputed. Similarly, if the patient is missing the day 180 NT Pro BNP value, a LOESS smoother will be used to predict the day 180 value based on the other values for that patient. In the event that a patient had a LVAD placed yet did not have a HF hospitalization after adjudication, the hospitalization where the LVAD was reported will be considered a HF hospitalization.

Statistical Tests:

A Wilcoxon test (PROC NPAR1WAY) will be used to test whether there is a difference in the distribution of ranks between the two treatment arms.

Interpretation of Results: Lower rank scores indicate worse outcomes

9. Secondary Endpoint Descriptions

9.1 – Secondary Endpoint #1

Endpoint Description: Change in cardiac structure and function from baseline to 180 days as assessed by left ventricular end-systolic volume (LVESV) index, left ventricular end-diastolic volume (LVEDV) index, left ventricular ejection fraction, and E/e' ratio.

Response Variable Definition: Change for each response is defined as the 24 week value – baseline value. The echocardiogram is performed at baseline and day 180.

Additional Covariates: Baseline value of the response variable

Handling of Dropouts and Missing Data:

If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of
missing echocardiogram data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline EF, baseline LVEDV, baseline LVESV, baseline E, baseline medial e’, baseline lateral e’, day 180 EF, day 180 LVEDV, day 180 LVESV, day 180 E, day 180 medial e’, and day 180 lateral e’.

Two sensitivity analyses will be performed that account for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

2) Worst rank analysis: The observed change in ECHO measure will be ranked from smallest to largest. Any patient that died or had a LVAD prior to the assessment point will be assigned the worst rank possible based on the observed data.

Statistical Tests:
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in each of the response variables between the two treatment groups. The non-parametric Wilcoxon Rank Sum test (PROC NPAR1WAY) will be used for the worst rank score analysis.

Interpretation of Results: For the general linear models, a decrease in EF, an increase in LVEDV index, an increase in LVESV index, and an increase in medial and lateral E/e’ indicate worse outcomes. Lower rank scores indicate worse outcomes in the worst rank analysis.

9.2 – Secondary Endpoint #2

Endpoint Description: Change in functional status as assessed by six minute walk distance at 30, 90, and 180 days after randomization.

Response Variable Definition: Change for each response is defined as the post-baseline value – baseline value. The six minute walk is performed at baseline, and 30, 90, and 180 days.

Additional Covariates: Baseline value of the response variable

Handling of Dropouts and Missing Data:
If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of missing walk distance data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline walk distance, day 30 walk distance, day 90 walk distance, and day 180 walk distance.

Two sensitivity analyses will be performed that account for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

2) Worst rank analysis: The observed change in walk distance will be ranked from smallest to largest. Any patient that died or had a LVAD prior to the assessment point will be assigned the worst rank possible based on the observed data.
Statistical Tests: A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups. The non-parametric Wilcoxon Rank Sum test (PROC NPAR1WAY) will be used for the worst rank score analysis.

Interpretation of Results: For the general linear models, a decrease in walk distance indicates worse outcome. Lower rank scores indicate worse outcomes in the worst rank analysis.

9.3 – Secondary Endpoint #3

Endpoint Description: Change in symptoms as assessed by the Kansas City Cardiomyopathy Questionnaire (Overall and Clinical Summary Scores) at 30, 90, and 180 days after randomization.

Response Variable Definition: Change for each response is defined as the post-baseline value – baseline value. The KCCQ is completed at baseline, and 30, 90, and 180 days.

Additional Covariates: Baseline value of the response variable

Handling of Dropouts and Missing Data: If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of missing KCCQ data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline overall summary score, baseline clinical summary score, day 30 overall summary score, day 30 clinical summary score, day 90 overall summary score, day 90 clinical summary score, day 180 overall summary score, and day 180 clinical summary score.

Two sensitivity analyses will be performed that account for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

2) Worst rank analysis: The observed change in KCCQ score will be ranked from smallest to largest. Any patient that died or had a LVAD prior to the assessment point will be assigned the worst rank possible based on the observed data.

Statistical Tests: A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups. The non-parametric Wilcoxon Rank Sum test (PROC NPAR1WAY) will be used for the worst rank score analysis.

Interpretation of Results: For the general linear models, a decrease in KCCQ score indicates worse outcome. Lower rank scores indicate worse outcomes in the worst rank analysis.

9.4 – Secondary Endpoint #4

Endpoint Description: Individual components of the primary endpoint through 180 days after randomization.
9.4.1 Response Variable Definition: Mortality through day 180.

Additional Covariates: None

Handling of Dropouts and Missing Data:
All patients will have some information regarding mortality.

Statistical Tests: The Cox regression model for survival data (PROC PHREG in SAS) will be used to test the statistical significance of differences in mortality between the treatments. Kaplan-Meier curves will be generated to graphically display the mortality rates as a function of time from randomization in each treatment.

Interpretation of Results: Higher mortality rates indicate worse outcome.

9.4.2 Response Variable Definition: Heart failure hospitalization through day 180.

Additional Covariates: None

Handling of Dropouts and Missing Data:
All patients will have some information regarding occurrence of a heart failure hospitalization.

Statistical Tests: The Cox regression model for survival data (PROC PHREG in SAS) will be used to test the statistical significance of differences in heart failure hospitalization between the treatments. Kaplan-Meier curves will be generated to graphically display the heart failure hospitalization rates as a function of time from randomization in each treatment.

Interpretation of Results: Higher heart failure hospitalization rates indicate worse outcome.

9.4.3 Response Variable Definition: Time-averaged proportional change in NT pro BNP through 180 days.
The NT Pro BNP is collected at baseline, and 30, 90, and 180 days.

Additional Covariates: Baseline value of the response variable

Handling of Dropouts and Missing Data:
If the patient is missing any of the required data such that the change cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of missing NT Pro BNP data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline NT Pro BNP, day 30 NT Pro BNP, day 90 NT Pro BNP, day 180 NT Pro BNP.

Two sensitivity analyses will be performed that account for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.
2) Worst rank analysis: The time-averaged change in NT Pro BNP will be ranked from smallest to largest. Any patient that died or had a LVAD prior to the assessment point will be assigned the worst rank possible based on the observed data.

**Statistical Tests:**
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups. The non-parametric Wilcoxon Rank Sum test (PROC NPAR1WAY) will be used for the worst rank score analysis.

**Interpretation of Results:** For the general linear models, an increase in NT pro BNP indicates worse outcome. Lower rank scores indicate worse outcomes in the worst rank analysis.

### 9.5 – Secondary Endpoint #5

**Endpoint Description:** Number of combined events (death + HF hospitalization, death + HF hospitalization + ED visit).

**Response Variable Definition:** Both the matched and unmatched approach to the win ratio of Pocock will be used based on the description in the manuscript (reference 28 in the FIGHT protocol). In general terms, each patient will be placed into one of the hierarchical groups based on the comparison of active versus placebo patients. The actual categorization will differ depending on whether the matched or unmatched approach is used. The hierarchy for the comparisons is as follows:

- (a) Liraglutide patient had death first
- (b) Placebo patient had death first
- (c) Liraglutide patient had HF rehospitalization first
- (d) Placebo patient had HF rehospitalization first
- (e) Liraglutide had ED visit first – only for endpoint including ED visit
- (f) Placebo patient had ED visit first – only for endpoint including ED visit
- (g) None of the above

**Additional Covariates:** None

**Handling of Dropouts and Missing Data:**
All patients will have some information regarding mortality. In the event that a patient had a LVAD placed yet did not have a HF hospitalization after adjudication, the hospitalization where the LVAD was reported will be considered a HF hospitalization.

**Statistical Tests:** The win ratio approach of Pocock will be used to evaluate each endpoint. Calculation of the test statistics, p-values and confidence intervals is detailed in the Pocock manuscript (reference 28 of FIGHT protocol).

**Interpretation of Results:** Lower win ratio scores indicate a worse outcome

### 9.6 – Secondary Endpoint #6
**Endpoint Description:** Global rank endpoint, which includes time to death, time to heart failure hospitalization or ED visit, and time-averaged proportional change in NT pro BNP at 180 days

**Response Variable Definition:** A rank score based on time to death (tier 1), time to heart failure hospitalization or ED visit (tier 2), and time-averaged proportional change in NT pro BNP at 180 days (tier 3). The only difference between this endpoint and the primary endpoint is addition of ED visit to tier 2. See section 8.1 for a description of the handling of dropouts and missing data, statistical tests and interpretation of results.

10. **Tertiary Endpoint Descriptions**

10.1 – **Tertiary Endpoint #1**

**Endpoint Description:** Change in core lab biomarkers from baseline to 30, 90 and 180 days. The specific list of biomarkers will be finalized in the near future. The list will include at least aldosterone, cystatin C, and hsCRP.

**Response Variable Definition:** Change for each response is defined as the post-baseline value – baseline value. The core lab biomarkers are collected at baseline and 30, 90, and 180 days.

**Additional Covariates:** Baseline value of the response variable

**Handling of Dropouts and Missing Data:**
If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of missing biomarker data: treatment arm (liraglutide, placebo), stratification (diabetes, no diabetes), age, sex, baseline lab value, day 30 lab value, day 90 lab value, and day 180 lab value.

One sensitivity analysis will be performed that accounts for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

**Statistical Tests:**
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups.

**Interpretation of Results:** For the general linear models, an increase in cystatin C or hsCRP indicates a worse outcome. Since changes in aldosterone could be considered beneficial or troublesome, the interpretation of those results will be made in the context of the observed data.

10.2 – **Tertiary Endpoint #2**

**Endpoint Description:** Change in glycosylated hemoglobin (HbA1c) from baseline to 30, 90 and 180 days.
Response Variable Definition: Change for each response is defined as the post-baseline value – baseline value. The HbA1c is collected at baseline and 30, 90, and 180 days.

Additional Covariates: Baseline value of the response variable

Handling of Dropouts and Missing Data:
If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of missing HbA1c data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline lab value, day 30 lab value, day 90 lab value, and day 180 lab value.

One sensitivity analysis will be performed that accounts for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

Statistical Tests:
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups.

Interpretation of Results: For the general linear models, an increase in HbA1c indicates a worse outcome.

10.3 – Tertiary Endpoint #3

Endpoint Description: Change in weight from baseline to 30, 90 and 180 days.

Response Variable Definition: Change for each response is defined as the post-baseline value – baseline value.

Additional Covariates: Baseline weight

Handling of Dropouts and Missing Data:
If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of missing weight data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline weight, day 30 weight, day 90 weight, and day 180 weight.

One sensitivity analysis will be performed that accounts for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

Statistical Tests:
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups.
Interpretation of Results: For the general linear models, an increase in weight indicates a worse outcome.

10.4 – Tertiary Endpoint #4

Endpoint Description: Change in insulin resistance as assessed by Homeostasis Model Assessment-Insulin Resistance (HOMA-IR) from baseline to 180 days. The HOMA-IR derivation uses local lab glucose and insulin.

Response Variable Definition: Change for each response is defined as the post-baseline value – baseline value. Glucose is measured at baseline and 30, 90, and 180 days. Insulin is measured at baseline and 180 days.

Additional Covariates: Baseline value of the response variable

Handling of Dropouts and Missing Data:
If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of missing insulin resistance data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline glucose, baseline insulin, day 30 glucose, day 90 glucose, day 180 glucose and day 180 insulin.

One sensitivity analysis will be performed that accounts for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

Statistical Tests:
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups.

Interpretation of Results: For the general linear models, an increase in HOMA-IR indicates a worse outcome.

10.5 – Tertiary Endpoint #5

Endpoint Description: Change in local lab fasting lipids from baseline to 180 days. The measures of interest include total cholesterol, LDL, HDL, and triglycerides.

Response Variable Definition: Change for each response is defined as the post-baseline value – baseline value.

Additional Covariates: Baseline value of the response variable

Handling of Dropouts and Missing Data:
If the patient is missing any of the required data such that a difference cannot be calculated, multiple imputation will be used to address the missing data. The following variables will be used in the calculation of
missing lipid data: treatment arm – (liraglutide, placebo), stratification – (diabetes, no diabetes), age, sex, baseline LDL, HDL, and triglycerides and day 180 LDL, HDL and triglycerides.

One sensitivity analysis will be performed that accounts for missing data differently.

1) Observed data: If the patient is missing any of the required data such that a difference cannot be calculated, the patient is excluded from that particular analysis.

Statistical Tests:
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in response variables between the two treatment groups.

Interpretation of Results: For the general linear models, an increase in LDL, triglycerides, and total cholesterol indicates a worse outcome. A decrease in HDL indicates a worse outcome.

11. Safety analysis – see section 15.4 of protocol

11.1 – Safety Endpoint #1
Safety Endpoint #1: Comparison of serious adverse events through 180 days

Response Variable Definitions: Each unique serious adverse event (SAE) type based on preferred term within body system. If a patient has more than one event of the same type, the patient is only counted once. The overall rate of serious adverse events will also be compared by treatment arm.

Additional Covariates: None

Handling of Dropouts and Missing Data:
It is assumed that all patients will have some assessment of serious adverse events made through their individual follow-up durations. No data are expected to be missing.

Statistical Tests:
Statistical comparisons will be based on the Fisher’s mid-p-value.

Interpretation of Results: Higher SAE rates indicate a worse outcome.

11.2 – Safety Endpoint #2
Safety Endpoint #3: Change in local laboratory values at 30, 90 and 180 days.

Response Variable Definitions: Change in laboratory values will be calculated as post-baseline value – baseline value. The list of laboratory values includes sodium, potassium, chloride, bicarbonate, BUN, creatinine, GFR, total calcium, total protein, albumin, ALT, AST, alkaline phosphatase, total bilirubin, magnesium, glucose, uric acid, hemoglobin, hematocrit, red blood cell count, red cell distribution width, white blood cell count, and platelet count.

Additional Covariates: Baseline lab value

Handling of Dropouts and Missing Data:
The analyses will be conducted using only observed data with no adjustment for incomplete data.

Statistical Tests:
A general linear model (PROC GLM in SAS) will be used to estimate and statistically compare the differences in the change in local laboratory values measures between the two treatment groups.

Interpretation of Results: Since changes in many labs could be considered beneficial or troublesome, the interpretation of those results will be made in the context of the observed data.

12. Interim Analyses

Interim data analysis for efficacy will not be conducted due to the relatively small size and short duration of this clinical trial. Safety data will be periodically assessed by the Data and Safety Monitoring Board (DSMB) based on the reporting of adverse events and local laboratory changes. There are no pre-specified guidelines for determining stopping rules due to a safety concern. The clinical opinion from the DSMB deliberations will be sole determinant.

13. Subgroup of Interest

Selected analyses will be performed examining the effects of diabetes at baseline. Models will either adjust for presence of diabetes or will be performed in diabetes subgroups, as appropriate. Per National Heart, Lung, and Blood Institute requirements, the primary endpoint will also be assessed within each gender.