STATISTICAL ANALYSIS PLAN

Protocols: WB01-202

An Exploratory Placebo-Controlled Study to Evaluate the Safety and Metabolic Effects of Food Products WBF-0010 and WBF-0011 When Administered to Subjects with Type 2 Diabetes Treated with Diet and Exercise Alone or in Combination with Metformin and/or Sulfonylurea

Original Protocol WB01-202: 25 Sept 2017
Current Protocol WB01-202: 19 June 2018

Prepared by:

Orville Kolterman, M.D.
Chief Medical Officer

James Bullard, PhD
Principal Statistician

Fanny Perraudeau, PhD
Statistician

Version: 0.2
Date: September 24, 2018

Introduction

Background
Purpose of the Analysis

Objectives

- Study Aims
- Endpoints
  - Safety and Tolerability
  - Primary Endpoints
  - Secondary Endpoints

Methods

- Design
- Interventions
  - Table 1: Study Arms
- Participants
- Randomization Schedule and Blinding Procedures
  - Sequence Generation
  - Statistical Software
- Observations and Assessments
  - Table 2: Primary endpoint schedule
  - Table 3: Secondary endpoint schedule
- Sample Size and Power
- Electronic Data Capture

General Considerations

- Timing of Analyses
- Analysis Populations
  - Enrolled Population
  - Intent-to-Treat Population (Intention-to-treat Analysis)
  - Per Protocol Population (Per-protocol Analysis)
  - Table 3: Definition of analysis sets.
- Covariates and Subgroups
  - Subject Disposition
  - Demographic and Baseline Characteristics
Concomitant Medications
Compliance and Missing Data
Treatment Compliance
Study-procedure Compliance
Interim analysis and Data Monitoring
Multi-site Analysis
Multiple Testing

Summary of Study Data
Study Biospecimen Processing
Handling of Multiple Observations
Handling of Missing Data
  Early withdrawal
  3-hour standardized meal tolerance tests
Calculation of Pharmacodynamic Parameters
  Area Under the concentration-time Curve (AUC(0-3hr))
  HOMA-IR
  Matsuda Index

Efficacy Analyses
Primary Endpoint Analyses
  Statistical Model for Primary Outcomes
Secondary Endpoint Analyses
  HADS questionnaire
  Optional Questionnaires
  FACIT Questionnaires

Safety Analyses
  Extent of Exposure
  Adverse Events
  Clinical Laboratory Evaluations
  Vital Signs

References
INTRODUCTION

Background

Diabetes mellitus and its comorbidities are currently a global, public health crisis, with approximately 30 million people diagnosed with diabetes and an estimated 80 million additional people with prediabetes in the US alone. Although diabetes is a multifactorial disease, it is widely believed that the proliferation of diabetes worldwide can be largely attributed to the adoption of the Western diet. Recently, research has shown that the use of the Western diet has significant impacts on the gut microbiome, altering both the composition and activities of that microbiome (the microbiome refers to the bacteria, viruses, fungi, phage, and other microorganisms that reside in the human intestine).

Purpose of the Analysis

These analyses will assess the results from a study designed to explore the effects of oral supplementation with two unique medical food formulations as compared to placebo by characterizing the effects of ingesting the formulations in terms of safety and tolerability endpoints (safety), effect on metabolic functions (efficacy) and change in inflammatory markers (inflammatory). The active ingredients in each formulation are commensal microbes grown to cGMP standards under anaerobic conditions. All ingredients in both formulations, as well as the placebo, have been GRAS qualified for use in humans.

OBJECTIVES

Study Aims

The overall goal of this study is to test the hypothesis that oral supplementation with a medical food designed to increase butyrate production and promote the health of the colonic mucin layer will improve metabolic health. The medical food formulations being tested in this study contain butyrate-producing organisms with or without a barrier permeability-regulating strain.
In this exploratory intervention study, we intend to evaluate these issues within patients with early stage type 2 diabetes. This includes patients diagnosed but not exposed to anti-diabetic therapy and patients treated with one of the three most commonly used initial treatment regimen(s), i.e. metformin and/or sulfonylurea. Two different medical food formulations will be tested: a formulation containing 2 butyrate-producing microbes and a second formulation consisting of the first formulation plus approximately matching colony forming units (CFUs) of both a third butyrate producer and a barrier permeability-regulating strain, which has been associated with improved gut-epithelial barrier function in rodent studies. Subjects will receive their randomized formulation twice a day for 12 weeks, followed by a 4-week washout period.

This document provides a detailed description of the statistical methods and procedures to be implemented during the analysis of the study (Whole Biome Protocol WB01-202). The proposed methods and approaches to the data analysis should be viewed as flexible. If the data suggest and warrant it, deviations from this plan will be considered. However, any deviations from this plan must be substantiated by sound statistical rationale, thoroughly documented prior to undertaking the analysis, and subsequently explained in detail in the clinical study report (CSR).

Endpoints

Safety and Tolerability

Safety and tolerability will be assessed based on treatment emergent adverse events and changes in safety laboratory parameters.

Primary Endpoints

The primary metabolic efficacy endpoint is the decrease from baseline to Week 12 in the total area under the glucose curve (AUC) of a standardized 3-hour meal tolerance test (MTT),

$$\Delta AUC_{Glucose} = AUC_{Glucose,Visit7} - AUC_{Glucose,Visit1}. $$

The primary inflammatory efficacy endpoint is the decrease from baseline to Week 12 in log-scaled c-reactive protein (CRP),
\[ \Delta CRP = \log_{10} \left( \frac{CRP_{\text{Visit}2}}{CRP_{\text{Visit}1}} \right) \]

Secondary Endpoints

In addition to the primary outcomes described above, the following secondary endpoints will be evaluated:

1. Decrease in the incremental area under the glucose curve ($\Delta AUC_{\text{glucose}}$) from baseline to Week 12 during standardized meal tolerance test
2. Decrease in HbA1c from baseline to Weeks 4 and 12
3. Change in the total area under the insulin curve ($\Delta AUC_{\text{insulin}}$) from baseline to Week 12 during standardized meal tolerance test
4. Change in the incremental area under the insulin curve ($\Delta AUC_{\text{insulin}}$) from baseline to Week 12 during standardized meal tolerance test
5. Decrease in fasting glucose from baseline to Weeks 4, 12, and 16
6. Change in fasting insulin from baseline to Weeks 4, 12, and 16
7. Decrease in inflammatory markers from baseline to Weeks 4 and 12:
   a. C-reactive protein (CRP)
   b. IL-6
   c. TNF-\(\alpha\)
8. Decrease in additional inflammatory markers from baseline to Week 12:
   a. IL-10
   b. TGF-\(\beta\)
9. Decrease in homeostatic model assessment for insulin resistance (HOMA-IR) from baseline to Week 12
10. Assessment of change in the Matsuda index from baseline to Week 12
11. Assessment of change in fasting lipid panel [total cholesterol, LDL cholesterol, HDL cholesterol, and triglycerides (TGs)] from baseline to Weeks 4, 12, and 16
12. Decrease in body weight from baseline to Weeks 2, 4, 12, and 16
13. Decrease in BMI from baseline to Weeks 2, 4, 12, and 16
14. Decrease in waist circumference between baseline and Week 12

15. Change in qRT-PCR -determined fecal concentrations of WB010 and WB011 -intervention strains from baseline to Weeks 4, 12, and 16

16. Change in responses to Hospital Anxiety and Depression Scale (HADS) questionnaire from baseline to Week 12

17. Change in response to FACIT D&AD and Victoria Bowel Performance Scale Questionnaire

Following completion of the Week 12 visit, subjects will discontinue study product. They will then return to the clinic for a final visit at the end of Week 16, bringing with them a frozen stool sample collected during the prior 3 days to assess the persistence of the administered microorganisms in stool ~4 weeks after discontinuing ingestion. In addition, fasting glucose, insulin, and lipid panel will be collected at Week 16 to evaluate the persistence of changes in these parameters.
METHODS

Design

The study is a balanced, parallel, double-blind, placebo-controlled, exploratory study enrolling subjects at 6 clinical sites across the US with three experimental arms; two active and one placebo arm, shown in Table 1.

Interventions

Table 1: Study Arms

<table>
<thead>
<tr>
<th>Arm</th>
<th>Intervention (WBF-0009)</th>
<th>WB-STR-0006 (CFU/day)</th>
<th>WB-STR-0005 (CFU/day)</th>
<th>EXT-STR-0001 (CFU/day)</th>
<th>WB-STR-0001 (CFU/day)</th>
<th>WB-STR-0008 (CFU/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>WBF-0010</td>
<td>3.3 x 10^7</td>
<td>1.6 x 10^9</td>
<td>2.0 x 10^7</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>3</td>
<td>WBF-0011</td>
<td>3.3 x 10^7</td>
<td>1.6 x 10^9</td>
<td>2.0 x 10^7</td>
<td>1.2 x 10^9</td>
<td>9.0 x 10^8</td>
</tr>
</tbody>
</table>

Participants

This study is designed to examine the effects oral supplementation with WBF-0010 and WBF-0011 on people with type 2 diabetes treated with diet and exercise alone or in combination with metformin and/or a sulfonylurea. Subjects meeting the inclusion and exclusion criteria will be randomized to one of three treatment arms (Table 1).

Complete inclusion/exclusion criteria is detailed in the protocol WB01-202_Version2_Amendment3_19June2018.
Randomization Schedule and Blinding Procedures

Randomization to a treatment arm will be managed centrally and stratified in blocks of three at each clinical site such that study sites maintain balance across treatment arms.

Sequence Generation

For each clinical site, a sequence of length $3M (M >> 60)$, $A_{site}$, containing $M$ permutations of the group labels, $G = 1, 2, 3$ is generated. For each of the $3M$ group assignments, $a_{site,m}; m = 1, 2, \ldots, 3M$, a unique 3-character alphanumeric kit-code, $k_m$, and define $K_{site}$ as the sequence $k_1, k_2, \ldots, k_{3M}$ is generated.

The study logistics team, receives both $A_{site}$ and $K_{site}$ and labels 12 bottles for a single subject with a single kit-code according to the map: $g \mapsto \{g = a_{site,j} \mid j = 1, 2, \ldots, 3M\}$. This key is maintained confidentially by a single member of the logistics team until unblinding.

At each clinical site, when the $n^{th}$ subject is randomized to a treatment arm, they are assigned kit-code, $k_{site,n}$. Subjects are listed by their subject number $n$ and their kit-code, $k_{site,n}$. The clinical sites and study subjects are blinded to the randomization block-size (3) and arm assignments, whereas study sponsor staff are blinded to the arm assignments.

Following database lock, the member of the logistics team holding the randomization key will provide the study analysis team a copy of the key with the function,

$$k_{site,n} \mapsto a_{site,n}; a \in G,$$

to map each subject's kit-code to their group assignment.

Statistical Software

All statistical analyses will be performed using R version 3.5.1 (Feather Spray). Random sequence generation will be performed using the R function `sample`, as `replicate(M, sample(1:3))`. 
Observations and Assessments

This is a 16-week study that will include a total of 3 periods: screening, treatment and washout (Figure 1: Study Timeline). Study outcomes of interest are defined as the intra-subject change of a clinically relevant biomarker after 12-weeks of treatment (primary and select secondary outcomes; depicted in Tables 1 and 2). Additional outcomes of interest and their schedules are depicted in Table 2.

Figure 1: Study assessments by visit
Table 2: Primary endpoint schedule

<table>
<thead>
<tr>
<th>Primary Endpoints</th>
<th>Baseline Visit 1 Day 0</th>
<th>Week 4 Visit 3 Week 4</th>
<th>Endpoint Visit 7 Week 12</th>
<th>Washout Visit 8 Week 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTT Glucose</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRP</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Secondary endpoint schedule

<table>
<thead>
<tr>
<th>Secondary Endpoint #</th>
<th>Biomarker</th>
<th>Baseline Visit 1 Day 0</th>
<th>Week 4 Visit 3 Week 4</th>
<th>Endpoint Visit 7 Week 12</th>
<th>Washout Visit 8 Week 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Incremental AUC glucose</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>HbA1c</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Total AUC insulin</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Incremental AUC insulin</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Fasting Glucose</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>6</td>
<td>Fasting Insulin</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>7a</td>
<td>CRP</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>7b</td>
<td>IL6</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>7c</td>
<td>TNF-alpha</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>8a</td>
<td>IL10</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8b</td>
<td>TGF-beta</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>HOMA-IR</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Matsuda index</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Total-Cholesterol</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td>LDL-Cholesterol</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td>HDL-Cholesterol</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>11</td>
<td>Triglycerides (TG)</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>12</td>
<td>Weight</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>12</td>
<td>BMI</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>13</td>
<td>Waist Circumference</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Microbiome</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>15</td>
<td>HADS</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>FACIT D&amp;AD</td>
<td>o</td>
<td>o</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sample Size and Power

A sufficient number of individuals will be screened to enroll at least 60 subjects, with approximately 20 subjects assigned to each of the three arms. The study is conducted across 6 sites and within each site, subjects are randomized in blocks of 3 such that study sites maintain balance across treatment arms.

Electronic Data Capture

OpenClinica is used by the clinics to record and maintain the pertinent data from each subject for the trial throughout the study beginning with the screening visit. Data are collected using electronic case report forms which can be found in the supporting materials (https://docs.google.com/document/d/1U8YKn03FT0_e0WPywNyv1E22LB_53L7FtnjD9BL5NU/edit?usp=sharing).
GENERAL CONSIDERATIONS

Timing of Analyses

The final analyses will be performed after all enrolled study subjects have completed the 16-week study, study sites have resolved all open queries, clinical laboratories have processed and delivered all non-stool based lab results performed throughout the study and the database has been locked. qPCR processing will be performed blinded by the personnel at Whole Biome.

Study database locking procedure is documented in detail in WB01-202 SOP_OTH_001 - Database Lock.

Analysis Populations

Enrolled Population

The Enrolled Population for WB01-202 will consist of all subjects who signed the informed consent to participate in the study, who passed screening, and randomized at the baseline visit.

Intent-to-Treat Population (Intention-to-treat Analysis)

The Intent-to-Treat Population, ITT, for WB01-202 will consist of all enrolled subjects of WB01-202 who were randomized to a study arm at Visit 1 (Randomization/Baseline) and ingested at least one portion of study product.

Two ITT datasets are defined to investigate the potential of non-random dropout on conclusions inferred from primary endpoint analysis. Both ITT as-is and ITT datasets contain the same subjects, however, for the ITT datasets, the same procedure applied in per-protocol analysis is used. These two analysis sets provide insight into the effect of non-random dropout while accounting for potential issues in data integrity for a small number of subject's meal tolerance test results.

Per Protocol Population (Per-protocol Analysis)

The Per Protocol Population, PP, will consist of all ITT subjects who have had adequate exposure (greater than 85% of overall anticipated doses and
greater than 85% during the last 4 weeks of ingestion of study product) to the randomized study product during the 12-week treatment period and have adequately complied with the protocol as assessed by the Sponsor prior to database lock. Specifically, subjects should have completed the Visit 1 and Visit 7 procedures without major protocol deviations, and have achieved overall compliance to the assigned treatment as defined below in Treatment Compliance.

In Table 3: Definition of analysis sets, we define the distinct analyses sets and their relationship to various analyses to be presented in the final report.

Table 3: Definition of analysis sets.

<table>
<thead>
<tr>
<th>Name</th>
<th>Subjects in Analysis Set</th>
<th>MTT, ITT Processing Procedure</th>
<th>Missingness</th>
<th>Endpoint</th>
<th>Additional Supporting Endpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITT as-is</td>
<td>All randomized subjects who received at least one portion of study product</td>
<td>None</td>
<td>LOCF/NOCB</td>
<td>Safety</td>
<td>ΔAUC_{glucose}, ΔCRP</td>
</tr>
<tr>
<td>ITT</td>
<td>All randomized subjects who received at least one portion of study product</td>
<td>LOCF/NOCB</td>
<td>NA</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Per Protocol (PP)</td>
<td>ITT - {Withdraw, Study Product Compliance, Major Study Procedure Violation}</td>
<td>See 3-hour standardized meal tolerance tests</td>
<td>NA</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>PP Metabolic</td>
<td>PP - {Missing ΔAUC_{glucose}}</td>
<td>Drop</td>
<td>ΔAUC_{glucose}</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>PP Inflammatory</td>
<td>PP - {Missing ΔCRP}</td>
<td>Drop</td>
<td>ΔCRP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PP Secondary Endpoint (Y)</td>
<td>PP - {Missing Y}</td>
<td>Drop</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Covariates and Subgroups

Subject Disposition

Subject disposition and enrollment will be summarized by treatment group and clinical site for the Randomized Population.
Demographic and Baseline Characteristics

Demographic and baseline characteristics will be summarized descriptively by treatment group and clinical site for ITT Populations.

Demographic and baseline characteristics include, but are not limited to: gender, age, race, ethnicity, body weight, height, body mass index (BMI), HbA1c, fasting plasma glucose, and metformin and sulfonylurea usage. Dosage of metformin and sulfonylurea will be available in the concomitant medication table.

Concomitant Medications

Concomitant medications will be tabulated for ITT subjects containing subject numbers, medication name and dose, treatment frequency, and start and stop dates if available.

Concomitant medications include prior concomitant and new concomitant medications, in which the prior concomitant medications refer to the medications that have a start date prior to receiving the first randomized portion of study product and continue to be taken on or after that, and the new concomitant medications refer to the medications that were initiated after receiving the first randomized portion of study product.

Compliance and Missing Data

Treatment Compliance

Prior to database lock, a memo-to-file will be created to document the selection of the Per Protocol Population based on blinded data review of product and study-procedure compliance.

Study-procedure Compliance

Subjects are required to fast for at least 8 hours prior to Visit-1, 3 and 7. Subjects who have not fasted adequately are removed from the per-protocol analysis.

For each subject, both plasma and serum blood glucose are collected at time 0 of the MTT. As plasma glucose and serum glucose are expected to
agree to +/- 10 mg/dL when collected according to the current protocol, values differing by more than 60 mg/dL are deemed physiologically implausible and may represent potential sample handling and/or processing issues, e.g., mislabeling time t=0 in the MTT.

If preprandial plasma glucose and preprandial serum glucose are present and their absolute difference is greater than 60 mg/dL and the serum glucose value is more physiologically aligned with the remainder of the MTT values, the serum glucose value will be used in place of plasma glucose for the time zero MTT measurement.

Protocol deviations will be tabulated by ITT subject containing treatment group, clinical site, and a description of the protocol deviation. The protocol deviations will be reviewed by clinical and statistical personnel prior to database lock to determine each subject's eligibility for inclusion in the Per Protocol Population.

**Interim analysis and Data Monitoring**

There is no formal interim analysis planned for WB01-202. WB01-202 will be conducted in double-blind fashion with database lock planned at the completion of the 16-week study period. Accordingly, no planned alpha spending function will be implemented.

**Multi-site Analysis**

A total of six clinical sites enrolled subjects for this study. The original protocol was amended during the course of the study to add two additional clinical sites due to lower than expected enrollment rates at the original 4 sites. The two sites that were added are more similar to one another than they are to the other 4 sites.

Although we are not powered to perform stratified analysis by each site, we will perform stratified analysis by creating the derived variable, \textit{site-class}, taking on the value 0 for subjects enrolled at Youngstown, 1 for subjects enrolled at the Juno sites and 2 otherwise. This stratification is based on a cohort split criterium: Youngstown is not a traditional clinical site, Juno sites is a predominantly immigrant population and the other three sites are more traditional clinical sites. Details concerning the use of
the site-class derived factor in primary and secondary analyses can be found in the section Efficacy Analyses.

Multiple Testing

Inferential statistical testing will be conducted to control for the study-wise error rate of 10% (\(\alpha = .10\)). To control for the study-wise error rate and minimize the impact of multiple comparisons, p-values are assessed sequentially; interpretation ends with the first p-value larger than \(\alpha\).

The fixed sequence of hypothesis tests are presented in the section: Efficacy Analyses.
**SUMMARY OF STUDY DATA**

For descriptive statistical summaries, the mean, sample size (n), standard deviation (SD), standard error (SE), median, minimum (min), and maximum (max) values will be calculated for continuous variables. For logarithm-transformed data, the geometric mean and standard error of the geometric mean will also be provided. For categorical variables, frequency and counts in each category will be provided.

Min and max values will be rounded to the precision of the original value. Means, least squares (LS) means, and medians will be rounded to 1 decimal place greater than the precision of the original value. SDs, SEs, and 95% confidence intervals (CIs) will be rounded to 2 decimal places greater than the precision of the original value. Percentages for summarizing categorical data will be rounded to one decimal place.

P-values will be presented with 4 decimal places and values less than 0.0001 will be presented as <0.0001.

**Study Biospecimen Processing**

All fecal samples were frozen immediately after collection and shipped frozen to Whole Biome where they will be processed and prepared for quantitation of probiotic strains by qPCR. After unblinding, participant’s fecal concentrations of strains will be determined, provided to the study analysis team and summarized in a stand alone research report which will be attached as an appendix to the clinical study report. Bacterial fecal concentrations will be presented and assessed as an efficacy predictor.

All clinical blood samples are handled by a centralized laboratory, Quest Labs; Quest will provide personally de-identified and encrypted rolling electronic access to laboratory results. The de-identified encrypted results are stored in a 21 CFR Part 11 compliant, read-only, audited, secured, limited-access data server.

**Handling of Multiple Observations**

All values, scheduled or unscheduled, will be presented in data listings. However, multiple observations will be pre-processed before summarizing or analyzing the data.
If an unscheduled visit occurs in a visit window with an existing scheduled visit, the assessment at the scheduled visit will be used for data summary and analysis. If no scheduled visit assessment exists for a visit window but at least one unscheduled visit assessment is available within that visit window, then the latest unscheduled visit within the visit window will be used for data summary and analysis.

In some cases, a subject's laboratory results are invalid due to procedural aspects of collecting a sample, e.g., sample hemolyzed. In these cases, a subject may have the labs redone and a second set of results will exist for one or more of the clinical tests for that visit. If available, and no more than a week after the original visit, these lab results will be used for efficacy analyses (Per Protocol Population).

Prior to database lock, this procedure will be assessed and will be modified as appropriate with any modifications to the procedure fully documented.

**Handling of Missing Data**

**Early withdrawal**

Subjects who withdraw from the study prior to completing all study procedures through Study Termination, but have data collected for at least 60% of baseline and endpoint outcomes individually, will have their missing values imputed using LOCF. Imputed values are used in ITT analyses.

**3-hour standardized meal tolerance tests**

Incorrect or mis-annotated data may occur during the collection of the Baseline/Visit 1 and Endpoint/Visit 7 time-series measures (MTT, ITT). This involves incorrectly labeling and/or transcribing one or more time points while processing the clinical samples. If physiologically impossible glucose and insulin trajectories are observed for a subject at a particular visit, an attempt is made to correct the 'collection-time' of each glucose and insulin value of the profile.

All corrections to MTT, ITT time-series measurements will be performed on blinded data and both the Principal Investigator and the Statistician will
review all changes. All changes will be documented in detail in the final study report and locked before unblinding.

Subject to the procedures described in section Calculation of Pharmacodynamic Parameters, primary and secondary efficacy endpoints will be assessed in the Per Protocol Population using the correctly annotated glucose trajectories.

In some cases, glucose measurements are declared invalid/missing because a non-physiologically possible change occurs between two adjacent time points, likely due to errors in handling the samples. When the difference in glucose concentration is greater than 200 mg/dL between two consecutive time points, the measurement with the greatest absolute difference from the arithmetic mean of the other 4-6 collected values, that measurement is treated as missing. See below for detailed description of how missing values are handled when computing primary and secondary MTT related outcomes.

**Calculation of Pharmacodynamic Parameters**

A 3-hour standardized meal tolerance test (MTT) is conducted before administration of study product at Visit1/Baseline and again at Visit7/End after 12 weeks of study product administration.

Blood samples for the measurement of plasma glucose and serum insulin are to be collected at 0, 30, 60, 90, 120, 180 minutes relative to the ingestion of the standardized liquid breakfast (with t=0 as the time the subject begins to consume the liquid).

The following pharmacodynamic parameters will be calculated for plasma glucose and serum insulin endpoints:

**Area Under the concentration-time Curve (AUC(0-3hr))**

$AUC_{(0-3hr)}$ is computed using trapezoidal quadrature as implemented in the DescTools R package. Both incremental ($iAUC_{(0-3hr)}$) and total ($AUC_{(0-3hr)}$) AUC will be calculated. Incremental AUC represents the area under the glucose curve ignoring the area beneath the fasting concentration (baseline) value. Glucose and insulin values below the lower-limit of detection (<LLOQ) will be treated as missing. The LLOQ values for glucose
and insulin concentrations in this study are 20 mg/dL and 2 μIU/mL, respectively.

Missing insulin and glucose values collected during the MTT will be handled as follows:

1. If both preprandial serum and plasma glucose values are missing, glucose AUC_{(0-3hr)} will not be computed.
2. If preprandial plasma glucose is missing and preprandial serum glucose is present, plasma glucose will be replaced by serum glucose.
3. If the glucose or insulin value(s) for time points 180 or (120 and 180) are missing, iAUC_{(0-3hr)} and AUC_{(0-3hr)} will be calculated without including the missing points. The calculation of iAUC_{(0-3hr)} and AUC_{(0-3hr)} will be performed over the same time interval for both baseline and endpoint evaluations so that AUC are comparable between baseline and endpoint.
4. If the glucose or insulin value(s) are missing but there are non-missing values before and after the missing value(s), the iAUC_{(0-3hr)} and AUC_{(0-3hr)} will be calculated as if the sample(s) was (were) not planned at that time point. The calculation of iAUC_{(0-3hr)} and AUC_{(0-3hr)} will be performed over the same data points for both baseline and endpoint evaluations so that AUC are comparable between baseline and endpoint.
5. If 3 or more values out of the 6 collected values are missing, the iAUC_{(0-3hr)} and AUC_{(0-3hr)} will not be calculated.

HOMA-IR

The effects of insulin resistance will be examined through HOMA-IR calculated using the following formula

\[ \text{HOMA-IR} = \frac{G_0 I_0}{405} \]

where \(G_0\) is the fasting glucose concentration (mg/dL) and \(I_0\) is the fasting insulin concentration (μIU/mL).

Matsuda Index

The Matsuda index will be calculated using the following formula
Matsuda Index = \frac{10,000}{\sqrt{(G_0 I_0)(G_{mean} I_{mean})}}

where $G_0$ is the fasting glucose concentration (mg/dL), $I_0$ is the fasting insulin concentration (uIU/mL), $G_{mean}$ is the mean glucose concentration during MTT (mg/dL), $I_{mean}$ is the mean insulin concentration during MTT (uIU/mL), 10,000 is a simplifying constant to get numbers from 0 to 12, and the square root is a correction of the nonlinear values distribution.
Efficacy Analyses

All efficacy analyses will be performed for both the ITT (ITT-as-is and ITT) and Per Protocol Populations (Per-protocol Inflammatory and Per-protocol Metabolic).

Primary Endpoint Analyses

Descriptive statistics (n, mean, SD, SE, median, min, and max) will be used to summarize values of the primary endpoints by treatment group and clinical site. These summaries will be presented for the following populations: Per-protocol Inflammatory, Per-protocol Metabolic and Intention to Treat populations (ITT-as-is and ITT).

Statistical Model for Primary Outcomes

Modeling of $\Delta$AUC$_{glucose}$ and $\Delta$CRP is performed using a generalized linear modeling framework assuming conditional normality of response variables.

<table>
<thead>
<tr>
<th>Model</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{\text{unadjusted}}(O)$</td>
<td>$O \sim \text{TREATMENT}$</td>
</tr>
<tr>
<td>$M_{\text{full}}(O)$</td>
<td>$O \sim \text{TREATMENT + (SITECLASS + SEX + MET + SUL + AGE)}$</td>
</tr>
</tbody>
</table>

Table 4: Upper and lower model formulae. Models are specified in the formula language of the R programming environment.  

Where $M_{\text{unadjusted}}(O)$ and $M_{\text{full}}(O)$ define the unadjusted and full models where the outcome, O, is either $\Delta$AUC$_{glucose}$ or $\Delta$CRP and $\varepsilon$ is a normally distributed random variable with zero mean and unknown variance $\sigma^2$. For each outcome $O \in \{\Delta$AUC$_{glucose}$, $\Delta$CRP$, stepwise model selection by AIC using the function stepAIC from the R package MASS is performed. We set $M_{\text{unadjusted}}(O)$ and $M_{\text{full}}(O)$ as lower and upper models respectively in the stepAIC procedure.

---

1 The original SAP had incorrectly written this statistical model as, $O = \text{TREATMENT} \times \text{TREATMENT} \times \text{SITECLASS} + \text{SEX} + \text{MET} + \text{SUL} + \text{AGE} + \varepsilon$, failing to include the parenthesis and including an error term.
This procedure uses the Akaike information criterion (AIC) to select a final model, \( \mathcal{M}_{\text{adjusted}}(O) \). The adjusted model, \( \mathcal{M}_{\text{adjusted}}(O) \), provides estimates of the two treatment effects, \( \beta_{BF^0010}^{(O)} \) and \( \beta_{BF^0011}^{(O)} \), for \( O \in \{ \Delta AUC_{\text{glucose}}, \Delta CRP \} \), as compared to placebo and provides 95% confidence intervals and unadjusted p-values for the null hypotheses of no treatment effect.

The statistical testing outline is described in section: Hypothesis Testing Significance Level. In Table 5, we define the sequence of null hypotheses to be evaluated.

To test these hypotheses, we compute t-statistics using maximum-likelihood estimates of \( \hat{\beta}_f^{(O)} \) and \( SD(\hat{\beta}_f^{(O)}) \) obtained from the model \( \mathcal{M}_{\text{adjusted}}(O) \). The t-statistic is defined as,

\[
t_f^{(O)} = \frac{\hat{\beta}_f^{(O)}}{SD(\hat{\beta}_f^{(O)})}.
\]

With critical value, \( t_{0.05, df_{\text{adjusted}}(O)} \). We reject the null hypothesis when our estimate, \( \hat{\beta}_f^{(O)} \), is greater than \( t_{0.05, df_{\text{adjusted}}(O)} \times SD(\hat{\beta}_f^{(O)}) \).

<table>
<thead>
<tr>
<th>Test Order</th>
<th>Null Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \beta_{BF^0011}^{(\Delta AUC_{\text{glucose}})} = 0 )</td>
</tr>
<tr>
<td>2</td>
<td>( \beta_{BF^0010}^{(\Delta AUC_{\text{glucose}})} = 0 )</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_{BF^0011}^{(\Delta CRP)} = 0 )</td>
</tr>
<tr>
<td>4</td>
<td>( \beta_{BF^0010}^{(\Delta CRP)} = 0 )</td>
</tr>
</tbody>
</table>

Table 5: Primary outcome sequential test ordering (total AUC and CRP).

Secondary Endpoint Analyses

All secondary endpoint analyses will be performed for the Per Protocol Population (Per-protocol Analysis) with observed data. For each outcome, we report the number of subjects by treatment arm participating in the
analysis. In cases where the number of missing values is large or skewed by clinic, the analyzed population will be further characterized, e.g., HADS questionnaires.

Secondary efficacy endpoints are computed using data collected at one or more of Visit 1, Visit 3, Visit 7 and Visit 8. These data are modeled using generalized linear mixed models (GLMM) or generalized linear models (GLM) as appropriate. Skewed data distribution will be log-transformed as appropriate. For each of the secondary outcomes, an equivalent model selection as that described above will be performed. However, sequential testing and presentation of unadjusted and FDR-adjusted p-values for each hypothesis tested.

HADS questionnaire

The responses to the Hospital Anxiety and Depression Scale (HADS) questionnaire will be summarized by treatment group as an exploratory assessment of potential impacts on symptoms of anxiety and depression. The Hospital Anxiety and Depression Scale is a well-validated 14-item questionnaire that has been widely used in studies of medical populations [5].

The questionnaire assesses both anxiety and depression; anxiety is the sum of the odd questions and depression is the sum of the even questions. Values of 0-7 are considered normal, scores of 8-10 are indicative of a state of being anxious/depressed and scores >10 indicate an actual mood diagnosis. We model the change in HADS score from Baseline/Visit1 to Endpoint/Visit7 using the selection and modeling framework described above in: Statistical Model for Primary Outcomes.

Optional Questionnaires

Two optional questionnaires, the FACIT D&AD and the Victoria Bowel Performance Scale, administered at Week 12 and 16, were added to the protocol mid-study and, therefore, are available for only the subset of subjects finishing the study after the protocol modification. These questionnaires provide an exploratory assessment of potential differences in gastrointestinal symptoms experienced by subjects.
FACIT Questionnaires

Two FACIT questionnaires are provided as optional surveys to be filled out at weeks 12 and week 16; FACIT-AD: For patients with Abdominal symptoms and FACIT-D: For patients with Diarrhea. The responses to the questionnaires will be summarized for the population responding along with the characteristics of the population. Victoria Bowel Performance Scale Questionnaire

The BPS is a bipolar, nine-point ordinal scale from -4 diarrhea to +4 constipation. This is a self-assessment of fecal consistency. The population responding to this questionnaire will be characterized and the responses summarized.

**SAFETY ANALYSES**

**Extent of Exposure**

The extent of exposure to study product will be summarized separately for each study arm. This value is reported as the number of subjects exposed for at least, 1, 2 and 3 months.

<table>
<thead>
<tr>
<th>formulation</th>
<th>1 month</th>
<th>2 months</th>
<th>3 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBF-0010</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WBF-0011</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summarizing extent of exposure

**Adverse Events**

All reported terms (investigator descriptions) for adverse events (AEs) will be grouped by organ system. Treatment-emergent adverse events (TEAEs) are defined as adverse events with an onset or worsening after ingestion of the first randomized portion of study product. Number and percentage of subjects with TEAEs will be summarized by treatment group, system organ classification, and preferred term. AEs occurring prior to receiving the first randomized portion of study product (e.g., during the Screening and Pre-Treatment Periods) will be provided separately in listings.
TEAEs leading to discontinuation from the study, TEAEs possibly or probably related to study treatment, serious TEAEs, death, TEAEs by severity (or intensity), and TEAEs with maximum severity (or intensity) and causality, will be summarized.

Additionally, TEAEs of special interest, e.g., specific gastrointestinal symptoms, will be summarized separately if deemed appropriate.

**Clinical Laboratory Evaluations**

All clinical laboratory results will be listed by site, treatment group, subject, and visit, including scheduled and unscheduled/repeat measurements. Laboratory assessments that are outside of normal ranges and/or with potential clinical importance will be flagged. Baseline values, the values at each visit, and changes from baseline values will be summarized descriptively for each of the quantitative laboratory assessments.

**Vital Signs**

For vital signs baseline values, the values at each visit, and changes from baseline values will be summarized by treatment group.
REFERENCES


