STATISTICAL ANALYSIS PLAN

A DOUBLE-BLIND, RANDOMIZED, PLACEBO-CONTROLLED, MULTIPLE-DOSE, TWO-CENTRE, SAFETY AND EFFICACY STUDY OF CO-ADMINISTRATION OF TESOFENSINE/METOPROLOL TREATMENT IN SUBJECTS WITH TYPE 2 DIABETES MELLITUS

Sponsor: Saniona, A/S
Baltorpvej 154
DK2750 Ballerup
Denmark

Author:

Steffen Selker / Junior Statistician
Profil Institut für Stoffwechselforschung GmbH
Hellersbergstraße 9
D-41460 Neuss

SIGNATURE PAGE

TM001: A double-blind, randomized, placebo-controlled, multiple-dose, two-center, safety and efficacy study of co-administration of tesofensine/metoprolol treatment in subjects with type 2 diabetes mellitus (T2DM).

We hereby declare that this Statistical Analysis Plan was prepared scientifically accurately and in full compliance with the current regulatory guidelines.

We will keep all information obtained in this Study confidential unless otherwise agreed in writing.

Dec 8, 2016

Roman V Dvorak, MD, PhD / Sponsor representative — Date:
Saniona, A/S
Baltorpvej 154
DK2750 Ballerup
Denmark

DEC 12, 2016

Grit Andersen, MD / Coordinating Investigator — Date:
Profil Institut für Stoffwechselforschung GmbH
Hellersbergstr. 9
D-41460 Neuss
Germany

Dec. 12, 2016

Steffen Selker / Trial Statistician — Date:
Profil Institut für Stoffwechselforschung GmbH
Hellersbergstr. 9
D-41460 Neuss
Germany

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AE	Adverse Event	ICF	Informed Consent Form
ATC	Anatomical Therapeutic Chemical	ICH	International Conference on Harmonisation
ANCOVA	Analysis of co-variance	IMP	Investigational Medicinal Product
AUC	Area Under the Plasma Concentration	IP	In-patient
BG	Blood Glucose	kg	Kilogram
BMI	Body Mass Index	l	Liter
BP	Blood Pressure	MedDRA	Medical Dictionary for Regulatory Activities
BMI	Body Mass Index	mg	Milligram
CI	Confidence Interval	mmol	Millimol
Cm	Centimeter	MRS	Magnetic Resonance Spectroscopy
CRF	Case Report Form	OP	Out-patient
CSR	Clinical Study Report	PE	Physical Examination
CV	Coefficient of Variation	PG	Plasma Glucose
DBP	Diastolic Blood Pressure	PHQ	Patient Health Questionnaire
dl	Deciliter	PPP	Per-Protocol Population
ECG	Electrocardiogram	SAE	Serious Adverse Event
eCRF	Electronic Case Report Form	SAP	Statistical Analysis Plan
FAS	Full Analysis Set	SBP	Systolic Blood Pressure
FPG	Fasting Plasma Glucose	SD	Standard Deviation
GCP	Good Clinical Practice	SE	Standard Error
HbA1C	Hemoglobin A1c	SOP	Standard Operating Procedure
HR	Heart Rate	T2DM	Type 2 diabetes mellitus

1 INTRODUCTION

1.1 Objectives of Statistical Analysis Plan

The statistical analysis plan (SAP) describes in detail the statistical analyses to be conducted on the data from the trial TM001. The purpose of this trial is to investigate the safety and efficacy of co-administration of tesofensine/metoprolol treatment vs. placebo in subjects with type 2 diabetes mellitus (T2DM).

The statistical analysis plan elaborates the statistical analyses outlined in the protocol and any deviations are clearly stated in the present SAP including the reason for the deviation.

The trial will be evaluated according to the specifications given in this analysis plan. However, deviations from the statistical analysis plan may be necessary in which case the nature of and the reason for the deviations will be documented and explained in the clinical study report (CSR).

The statistical analyses will be made in accordance with ICH-E9 guideline "Statistical Principles for Clinical Trials" [1].

1.2 Scope

This SAP is based on protocol 'TESO_Protocol_version 4.1_FINAL_clean.pdf' (31 Mar 2016).

2 STUDY OBJECTIVES AND ENDPOINTS

This is a two-centre, randomized, double-blind, placebo-controlled, multiple dose trial in subjects with T2DM.

The purpose of this trial is to compare the effects of co-administration of tesofensine/metoprolol treatment vs. placebo on 24-hour mean heart rate (HR) as well as on systolic and diastolic blood pressure and to demonstrate a positive effect of co-administration of tesofensine/metoprolol treatment on body weight, glycemic endpoints, liver fat, and overall tolerability and safety.

2.1 Study Objectives

Primary objective:

- To compare the effects of co-administration of tesofensine/metoprolol treatment vs. placebo on 24-hour mean HR

Secondary objectives:

- To compare the effects of co-administration of tesofensine/metoprolol treatment vs. placebo on mean systolic (SBP) and mean diastolic blood pressure (DBP)
- To demonstrate a positive effect of co-administration of tesofensine/metoprolol treatment on:
    - body weight
    - glycemic endpoints (fasting plasma glucose (FPG), 9-point plasma glucose (PG) profile and HbA1c)
    - body composition (liver fat)
- To evaluate overall safety and tolerability of co-administration of tesofensine/metoprolol treatment

2.2 Study Endpoints

Definition: baseline Day -2/-1/1

end of treatment Day 89/90/91

Primary endpoint

- Change from baseline to end of treatment in mean 24 hour HR

Secondary endpoints

- Change from baseline to end of treatment in mean SBP
- Change from baseline to end of treatment in mean DBP
- Change from baseline to end of treatment in body weight
- Change from baseline to end of treatment in Hemoglobin A1c (HbA1c)
- Change from baseline to end of treatment in FPG
- Change from baseline to end of treatment in 9-point PG profile

Exploratory endpoints

- Change from baseline to end of treatment in liver fat
- Change in fasting insulin from baseline to end of treatment
- Change in 1,5 anhydroglucitol from baseline to end of treatment
- Change from baseline to end of treatment in waist circumference
- Change from baseline to end of treatment in Patient Health Questionnaire (PHQ)-9 score

Safety Endpoints

- Adverse events (AEs), clinical laboratory findings, electrocardiogram (ECG)
- Number and severity of hypoglycemic and hyperglycemic events
- Vital Signs, Physical Examination (PE)

3 STUDY DESIGN AND DESCRIPTION

This is a double-blind, randomized, placebo-controlled, multiple-dose, two-center, safety and efficacy study of co-administration of tesofensine/metoprolol treatment in subjects with T2DM. Study medication will be administered for ninety (90) days (+2 days after the final assessments with half-dose of metoprolol). Following all baseline assessments, eligible subjects will be randomly assigned to one of the two arms (1:1).

3.1 Study Design

Screening: Subjects who give the written informed consent will be screened for the study. For subjects on any anti-diabetic medications, treatment with all anti-diabetic medications except metformin will be washed out. The subjects will return for a baseline visit at the end of the wash-out period (7-28 days). Subjects who at screening are receiving diet treatment and/or metformin treatment only for their diabetes will be invited for the baseline visit after a minimum of 3 days (i.e. do not require a washout of other anti-diabetic medication).

Baseline: Subjects will be admitted to the Unit (investigational site) in the evening of Day -2 to undergo all baseline assessments. MRS assessment should be performed at Day -2 (in a sub-set). BP measurements will be initiated at Day -2. The 24-h HR monitoring and 9-point glucose profile assessments will start in the morning of Day -1 and will be completed in the morning of Day 1. Blood draw to establish the baseline for all measure endpoints will be drawn in the morning of Day 1.

Randomisation: Following the completion of all baseline assessments on Day 1, eligible subjects will be randomized in equal numbers to double-blind treatment with co-administration of tesofensine/metoprolol or placebo. The first dose of study medication will be administered on Day 1. Subjects will receive medication supply for 8 days and will be released from the Unit.

Treatment period: After the Day 1, subjects will visit the Unit weekly - on Days 7, 14, 21 for safety evaluations and medication supply until Day 28 (visit 6). Subjects will receive medication supply for up to 8 days (dispensed during the visits 3, 4, 5). After the Day 28, subjects will be assessed at interval of 2 weeks on Days 42 and 56 (visits 7, 8), and at interval of 19 days on Day 70 (visit 9). Study medication will be dispensed for up to 15 days during each visit 6, 7, 8 and up to 20 days at visit 9. Visits on days 28 and 56 will also include additional assessments for efficacy evaluation.

Final admission/end of treatment: In the afternoon/evening of Day 89 subjects will come in for the final admission. MRS assessment should be performed at Day 89 (in a sub-set). BP measurements will be initiated at Day 89. Final 24-h HR monitoring and 9-point glucose

profile assessments will be initiated in the morning of Day 90, which will also be the last day when subjects receive the study medication. During the Day 90 the blood will be collected for metabolic endpoint as well as safety measurements. On Day 91 the 24-h assessments will be completed and then subjects will be released.

Final follow up visit: on Day 98 phone call will be performed to check subjects' status and potential AE(s). Subjects will come in for a final efficacy and safety assessment on Day 110.

The total trial duration for each subject will be between 115 (subject without wash out period) and 138 days (subjects who needed a wash pout period).

The schedules of events from the protocol are included in the appendix of this SAP.

3.2 Study Description

3.2.1 Trial Products

Investigational Medicinal Product

Tesofensine is a serotonin-noradrenaline-dopamine reuptake inhibitor, for oral administration, which is manufactured, packed and labeled by Delpharm in Reims (France).

Metoprolol is a beta1-selective (cardioselective) adrenoceptor-blocking agent, for oral administration. It has been formulated into the extended release tablets to provide a controlled and predictable release of metoprolol for once-daily administration.

The Investigational Medicinal Product (IMP) is a kit of one tablet of tesofensine 0.5mg and one tablet of metoprolol 100 mg, both formulated for oral use once a day.

Placebo

The placebo formulation is identical in presentation and appearance to the tesofensine and metoprolol tablets but includes only the excipients; it does not contain any active ingredients.

3.2.2 Study Participants

The target population is adult subjects suffering from T2DM. Subjects will be recruited from local population around the Units. This study will be conducted at two investigational sites (Units) in Germany.

3.2.3 Screening Failures

If a subject fails during the screening phase, before randomization, (e.g. due to the subject's decision, etc.) his/her relevant data will be listed in the screening log and filed in the study documents.

3.2.4 Discontinuation

The subject has a right to discontinue study participation at any time for any reason. The reason for discontinuation will be documented in the subject's source documents and in the electronic Case Report Form (eCRF).

3.2.5 Subject Replacement

Enrolment will continue until a total of 60 subjects are randomized. Subjects excluded during the screening period (Day -28 to -5) may be considered for another selection visit at a later

date but will need to undergo complete rescreening. Randomized subjects who withdraw or are withdrawn from the study for any reason will not be replaced.

3.3 Randomization Code Creation

Profil Institut für Stoffwechselforschung GmbH, 41460 Neuss, will generate the randomization codes using the computer program RANCODE 3.6.

4 ANALYSIS VARIABLES

4.1 Demographic and Baseline Characteristics

The following demographic variables will be collected at the screening/baseline visit:

- Age (years)
- Gender (male)
- Ethnic Origin
- Height (m)
- Body weight (kg)
- Body mass index (BMI) (kg/m^2) (calculated using height and body weight at screening)
- Waist circumference (cm)

In addition, the following baseline variables will be collected at the screening visit:

- Medical history/Concomitant illness
- Current Diabetes Treatment
- Vital signs (screening values)

4.2 Primary Endpoint

- Change from baseline to end of treatment in mean 24-hour HR:

  24-hour HR monitoring will be based on Telemetrie measurements. On Day -1 (V2) HR monitoring will start approximately 24 hours before first dosing and on Day 90 (V10) it will start one hour before last dosing.

  On both visits monitoring will be carried out to complete at least 24 hours of measurement (Days -1 to 1: baseline; Days 90 to 91: end of treatment). The data interval for analysis will be based on the beginning of the measurement and the completion of the 24 hours.

  The heart rate will be measured every minute and the mean will be recorded for every hour. There will be 24 hourly means per subject and visit.

The change from baseline to end of treatment is the difference between the mean of the 24 baseline hourly means and the mean of the 24 hourly means at end of treatment.

Additional HR measurements will be done parallel to BP measurements at V2 and V10 (see 4.3, mean systolic and mean diastolic BP). For the handling of these measurements see section 5.9.

4.3 Secondary Endpoints

- Change from baseline to end of treatment in mean systolic and mean diastolic BP:

  Over the course of three days (-2, -1, 1 for baseline and 89, 90, 91 for end of treatment) three measurements will be done at each of six different time points (Evening, Pre-Breakfast, Noon, Pre-Dinner, Midnight, Morning). For each of the six time points the mean value is calculated.

  The change from baseline to end of treatment is the difference between the mean of the six baseline values and the mean of the six values at end of treatment.

  Additional measurements of BP will be done by the subjects at home and recorded in their trial diaries (one measurement per day). For the handling of these measurements see section 5.9.

- Change from baseline to end of treatment in body weight:

  Body weight will be measured two times at baseline (Day -1 and Day 1) and one time at end of treatment (Day 90).

  The change from baseline to end of treatment is the difference between the mean of the two baseline values and the value at end of treatment.

  Additional body weight measurements will be done during various visits (V3-V9 and follow-up). For the handling of these measurements see section 5.9.

- Change from baseline to end of treatment in HbA1c:

  HbA1c will be measured from blood samples collected once at baseline (Day 1) and once at end of treatment (Day 90). These blood samples are analyzed at the central laboratory.

  The change from baseline to end of treatment is the difference between the HbA1c value at baseline and the HbA1c value at end of treatment.

Additional HbA1c measurements will be done during various visits (V6, V8 and follow-up). For the handling of these measurements see section 5.9.

- Change from baseline to end of treatment in 9-point plasma glucose profile:

The 9-point glucose profile consists of nine measurements (using a Super GL Glucose Analyzer) over the course of Day -1 and Day 1 (baseline) and nine measurements over the course of day 90 and day 91 (end of treatment) at predefined time points (before breakfast (~8h), 2 hours after breakfast (~10h), before lunch (~12h), 2 hours after lunch (~14h), before dinner (~18h), 2 hours after dinner (~20h), before going to bed (~22h), bed time (~3h), before breakfast (~8h). Since "Before breakfast" is approximately 08:00h the nine measurements are done over a course of approximately 24 hours.

9-point glucose profile as secondary endpoint will be measured as following:

For each of the visits the area under curve (AUC) will be calculated. This will be done based on the linear trapezoidal rule and actual sampling time points (nine time points in 24 hours):

$$AUC_{0-t} = \sum_{i=1}^{n} \frac{(C_i + C_{i-1})}{2}(t_i - t_{i-1}),$$

where C_i and C_{i-1} are the glucose values measured at the i-th time point (t_i) and (i-1)-th time point (t_{i-1}), respectively.

The change from baseline to end of treatment is the difference between the baseline AUC and the end of treatment AUC.

- Change from baseline to end of treatment in FPG:

FPG (as secondary endpoint) is the value of the 9-point glucose profile at time point "before breakfast" on Day -1 (baseline) and Day 90 (end of treatment).

The change from baseline to end of treatment is the difference between the FPG value at baseline and the FPG value at end of treatment.

Additional FPG measurements are done at the trial site using a Super GL Glucose Analyzer during various visits (V3-V9 and follow-up) and at home by the subjects (recorded in their diaries once each day). For the handling of these measurements see section 5.9.

4.4 Exploratory Endpoints

- Change from baseline to end of treatment in liver fat:

  Liver fat will be measured via MRS at Day -2 (baseline) and Day 89 (end of treatment) in a subset of the subjects, in total in ~40 subjects.

  The change from baseline to end of treatment is the difference between the baseline value and the value at end of treatment.

- Change in fasting insulin from baseline to end of treatment:

  Fasting insulin will be measured from blood samples collected once at baseline (Day 1) and once at end of treatment (Day 90). These blood samples are analyzed at the central laboratory.

  The change from baseline to end of treatment is the difference between the fasting insulin value at baseline and the fasting insulin value at end of treatment.

  Additional fasting insulin measurements will be done during various visits (V6, V8 and follow-up). For the handling of these measurements see section 5.9.

- Change in 1,5 anhydroglucitol from baseline to end of treatment:

  1,5 anhydroglucitol will be measured from blood samples collected once at baseline (Day 1) and once at end of treatment (Day 90). These blood samples are analyzed at the central laboratory.

  The change from baseline to end of treatment is the difference between the 1,5 anhydroglucitol value at baseline and the 1,5 anhydroglucitol value at end of treatment.

  Additional 1,5 anhydroglucitol measurements will be done during various visits (V6, V8 and follow-up). For the handling of these measurements see section 5.9.

- Change from baseline to end of treatment in waist circumference:

  Waist circumference will be measured two times at baseline (Day -1 and Day 1) and one time at end of treatment (Day 90).

  The change from baseline to end of treatment is the difference between the mean of the two baseline values and the value at end of treatment.

  Additional waist circumference measurements will be done during various visits (V3-

V9). For the handling of these measurements see section 5.9.

- Change from baseline to end of treatment in PHQ-9 score:

The PHQ-9 (Appendix II) was originally developed as a screener for depression in primary care and is commonly used in medical settings. The PHQ-9 demonstrated a strong correlation with a well-established measure of depression, and a score ≥ 13 demonstrated good sensitivity (83%) and specificity (72%) to that end[2]. The questionnaire consists of 9 questions with a rating corresponding to the answer of the subject (from 0 to 3). The score is the sum of all answers so it ranges from 0 to 27.

For PHQ-9 as exploratory endpoint the test results of Day -1 (baseline) and Day 91 (end of treatment) are utilized.

The change from baseline to end of treatment is the difference between the baseline value and the value at end of treatment.

Additional tests will be done during various visits (V4, V6, V7, V9). For the handling of their results see section 5.9.

4.5 Safety Endpoints

- AEs

An AE is any untoward medical occurrence in a trial subject administered an IMP and which does not necessarily have a causal relationship with the treatment. An AE can therefore be any unfavorable and unintended sign (for example, an abnormal laboratory finding), symptom, or disease temporally associated with the use of an IMP, whether or not considered related to the IMP. Any event fulfilling the criteria of an AE but occurring prior to randomization will not be entered into the CRF, but will be documented in the source data. AEs will be recorded from randomization until follow-up.

Intensity of AEs:
The following three-point rating scale will be used for rating the intensity of each AE:

Mild:	Awareness of signs or symptoms, no interference with daily activities
Moderate:	Symptoms cause discomfort with some interference with daily activities (disturbing)
Severe:	The subject is unable to work or conduct usual daily activities (disabling)

Causality of AEs:

The following five-point scale will be used for rating the causal relationship of the AE to the investigational study product:

Not related: The adverse event is clearly not related to the study treatment - i.e. another cause of the event is most plausible; and/or a clinically plausible temporal sequence is inconsistent with the onset of the event and the study treatment administration; and/or a causal relationship is considered biologically implausible.

Unlikely: In cases where sufficient information exists to establish beyond reasonable doubt that study treatment causality was not likely to be the cause of the event then such reports should be classified as unlikely related.

Possible: An event that follows a reasonable temporal sequence from administration of the study treatment follows a known or expected response pattern to the suspected study treatment, but that could readily have been produced by a number of other factors.

Probable: For inclusion in this category it is recommended that all the following minimum criteria should be complied with:

1. There should be a reasonable association in time between the administration of the study treatment and onset and duration of the reported event.

2. The description of the clinical phenomena should be consistent with, or at least plausible, given the known pharmacology and toxicology of the study treatment.

3. There should be no other equally plausible explanation(s) of the case. In particular, concurrent use of other medicinal products (and possible drug interactions) or intercurrent disease should be taken into account in the assessment.

Definitely related: The adverse event is clearly related to the study treatment – i.e. an event that follows a reasonable temporal sequence from administration of the study treatment, follows a known or expected response pattern to the suspected treatment, that is confirmed by improvement on stopping and reappearance of the event on repeated exposure and that could not be reasonably

explained by the known characteristics of the Subject's clinical state.

Outcome of AEs:
Outcome of adverse event may include at time of last observation:

Recovered / resolved

Recovering / resolving

Not recovered / not resolved

Recovered / resolved with sequelae

Fatal

Unknown

The outcome type and duration of the follow-up of subjects with AEs should be specified.

Action Taken for AEs:
AEs requiring therapy must be treated by recognized standards of medical care to protect the health and welfare of the subject. Appropriate equipment and medicines must be available to ensure the best possible treatment of emergency situations. Action(s) taken:

Study treatment withdrawn

Dose reduced

Dose increased

Dose not changed

Unknown

Not applicable

- Number and severity of hypoglycemic events

A hypoglycemic episode is confirmed if all plasma glucose values are

- ≤ 3.9 mmol/L (70 mg/dL) or
- > 3.9 mmol/L (70 mg/dL) when they occur in conjunction with hypoglycemic symptoms should be recorded by the subject. These must be transcribed into the eCRF throughout the study.

The record should include the following information:

- Date and time of hypoglycemic episode
- The plasma glucose level before treating the episode (if available)
- Whether the episode was symptomatic
- Whether the subject was able to treat him/herself

Classification of hypoglycemic episodes:

Minor hypoglycemic episode
- ≤3.9 mmol/L (70 mg/dL), or
- >3.9 mmol/L (70 mg/dL) when they occur in conjunction with hypoglycemic symptoms

Major hypoglycemic episode
- If oral carbohydrate, glucagon or IV glucose was administrated to the subject by another person

Nocturnal hypoglycemic episode
- A hypoglycemic episode occurred between 24:00 – 06:00.

- Number and severity of hyperglycemic events

Hyperglycemic events are recorded as part of the AEs.

- Clinical laboratory findings

The following assessments were performed at screening, V2, V6, V8 and V10:

Lipids: total cholesterol, HDL cholesterol, LDL cholesterol and triglycerides

The following assessments were performed at screening, V2 and V10:

Hematology: blood count (erythrocytes, hemoglobin, hematocrit, leukocytes), differential blood count (neutrophils, lymphocytes, monocytes, eosinophils, basophils) and platelets

Blood chemistry: creatinine, gamma glutamyl transferase, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, GFR and urea

The following assessments were performed at screening:

Infectious serology: Hepatitis B surface antigen, Hepatitis C antibodies, HIV-1/2 Combi

- ECG

A standard 12-lead ECG will be performed at screening, V2 and V10. The following parameters were recorded:

HR

PQ Interval

QRS Interval

QT Interval

QTcB Interval

- Vital signs

An examination of the following vital signs will be performed at all trial site visits (screening, V2-V10 and follow-up) and include:

Systolic and diastolic BP

HR

Respiratory rate

Body temperature

- Physical examination

An examination of the following body systems will be performed at screening, V10 and follow-up and included:

Head, ears, eyes, nose, throat (HEENT) incl. thyroid gland

Heart, lung, chest

Abdomen

Skin and mucosae

Musculoskeletal system

Nervous system

Lymph node

Other findings

4.6 Other Assessments

For all female participants, a pregnancy test will be performed at the safety laboratory on a blood (serum) sample obtained at the screening visit. During the rest of the trial a urine pregnancy test will be performed at the trial site according to local regulations (at V2, V6, V8 and follow-up).

Concomitant medication will be recorded throughout the trial.

The study compliance will be calculated from the amount of medication taken by the subjects and the amount of days they participated in the study.

PK samples will be collected once at visit 2 (Day 1, pre dose) and once at visit 10 (Day 90, one measurement).

5 STATISTICAL METHODOLOGY

Profil Institut für Stoffwechselforschung GmbH, Neuss, Germany, will be responsible for the statistical analysis. The statistical planning and conduct of analyses of the data from this study will follow the principles defined in relevant ICH guidelines and Profil's biostatistical standard operating procedures (SOPs). All statistical analyses will be performed using SAS® (SAS Institute Inc., Cary, North Carolina, United States of America [USA]), version 9.4 or later.

5.1 Sample Size Calculation

As this is an explorative study no formal sample size calculation was done. The chosen sample size is a balance between exposing the lowest possible number of subjects to the IMP, while still being able to compare the effects of tesofensine/ metoprolol treatment vs. placebo. The overall sample size is 60 (30 per treatment arm).

5.2 Analysis Sets

The following analysis sets are defined in accordance with the ICH-E9 guideline[1]:

Full Analysis Set (FAS)

is based on the intention-to-treat principle and includes all randomized subjects. Subjects will contribute to the evaluation 'as randomized'. In exceptional cases subjects from the FAS may be excluded. The decision will be done during the blinded database release meeting. In such cases the exclusion will be justified and documented.

Per-Protocol Population (PPP):

includes all randomized subjects without any important protocol deviations. Subjects in the PPP will contribute to the evaluation 'as treated'.

Safety Analysis Set:

includes all subjects receiving at least one dose of the investigational product or its comparator. Subjects in the safety set will contribute to the evaluation 'as treated'.

Analyses of primary, secondary and exploratory endpoints will be based on the FAS. In case of differences among FAS and PPP an additional analysis based on the PPP will be done, if deemed relevant. This decision will be done at the DBR meeting. The analysis of the safety endpoints will be based on the Safety Analysis Set.

Before data are released for statistical analysis, a blinded review of all data will take place to identify protocol deviations that may potentially affect the results. This review will be performed without revealing which IMP the subjects are assigned to. The blinding of the IMPs will be maintained for everyone involved in allocating subjects to the analysis sets until data are released for statistical analysis. Furthermore, outliers will be identified by data review according to ICH-E9[1], and a fake randomization. In addition, protocol deviations, which may potentially affect the results, will be identified and it will be evaluated if subjects and/or data should be excluded from the analysis. Obviously erroneous data points may be excluded from the analyses. The decision to exclude data points from the statistical analysis is the joint responsibility of the sponsor, the investigator and the trial statistician.

The subjects or observations to be excluded and the reason for their exclusion will be documented and signed by those responsible prior to database release. The documentation will be stored together with the remaining study documentation. The subjects and observations excluded from analysis sets, and the reason for this will be described in the CSR.

5.3 Coding

Adverse events, medical history, and concurrent conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual used version number will be documented in the database. Concomitant medication will be coded using the Anatomical Therapeutic Chemical (ATC) Classification System.

5.4 Missing, Unused and Spurious Data

All analyses, listings, tables and plots will be performed using all available data unless otherwise stated. By default missing data will not be imputed. The final decision regarding imputation of missing data will be made at the DBR meeting.

If samples are missing that are important for the derivation of the endpoints, and if a reanalysis is not possible/decided, certain endpoints will not be calculated and they will be excluded from analysis.

The subjects, observations or endpoints to be excluded and the reason for their exclusion will be documented prior to database lock. The documentation will be stored together with the remaining trial documentation. The subjects and observations excluded from analysis sets, and the reason for this will be described in the CSR.

5.5 Statistical Methods

Demographic and anthropometric information, medical history/concomitant illness and the results of fasting plasma glucose will be summarized by means of descriptive statistics

(number, mean, SD, minimum, median and maximum) or frequency tables for the whole population and stratified for treatment arm and trial site. Individual listings will be presented. The summaries of the demographic and background data will be presented for the Safety Analysis Set.

The analysis of all primary, secondary and exploratory endpoints will be based on the FAS. In case of differences among FAS and PPP an additional analysis based on the PPP will be done, if deemed relevant. This decision will be done at the DBR meeting. The endpoints will be summarized by treatment and visit, overall and stratified by trial site, if significant, using descriptive statistics. Continuous endpoints are summarized by the arithmetic mean, median, standard deviation, coefficient of variation (CV), minimum and maximum value. Categorical endpoints are summarized by the number (N) and percentage (%). Percentages of subjects will be based on non-missing values. Moreover, complete listings of individual values for all endpoints will be provided.

The analysis of the safety endpoints will be based on the Safety Analysis Set.

5.6 Analysis and Presentation of the Primary Endpoint

The primary endpoint will be derived from the individual heart rate measurements of the 24-h profile (see 4.2).

The primary endpoint will be checked for normal distribution via a Shapiro-Wilk test and will be analyzed with a parametric model. If the endpoint is not normal distributed then an additional non-parametric test will be done

The primary endpoint will be compared between treatment arms by means of an analysis of covariance (ANCOVA) model (proc MIXED) using change from baseline to end of treatment as dependent variable, treatment and trial site as fixed effects and the value from baseline as covariate. The residual errors are assumed i.i.d. and normally distributed.

In case of normally distributed parameter an additional analysis will be performed without trial site as fixed effect if this effect is not significant. But if it is significant the analysis with trial site as an effect will be kept and additionally the analysis without trial site as an effect will be provided per site.

Within the model least square mean (LS-mean) for each treatment as well as the difference of the means between the treatment arms, the corresponding 95% confidence intervals (CIs) and p-values will be calculated.

In case of not normally distributed parameter, an additional analysis will be performed by non-parametric techniques. For between treatment comparisons Wilcoxon Rank Sum test will

be used based on a two-sided alpha level of 0.05. In addition, the estimate of Hodges and Lehmann and the corresponding 95% nonparametric CI will be shown.

The individual data will be listed by visit, day and time point. A summary will be made by visit and treatment arm, overall and stratified by trial site, if significant, presenting mean, standard deviation, minimum, median and maximum.

For each subject the individual curves for the 24h- profiles will be plotted overlaid by treatment and visit.

Furthermore mean value curves of the treatment arms will be plotted overlaid in one figure, overall and stratified by trial site, if significant.

In addition box plots of baseline/end of treatment displaying the mean, quartiles, and minimum and maximum observation for 24h- mean HR will be presented for each treatment arm, overall and stratified by trial site, if significant.

Additionally figures, which present the HR/pulse values over the course of the whole trial (Mean HR from 24h measurement for V2 and V10; vital signs HR from the other visits), will be plotted. These figures will be presented for each subject individually and as mean value graphs overlaid by treatment. The mean graphs will be plotted overall and stratified by trial site, if significant.

5.7 Analysis and Presentation of the Secondary & Exploratory Endpoints

- Change from baseline (Days -2/-1/1) to end of treatment (Days 89/90/91) in mean SBP
- Change from baseline (Days -2/-1/1) to end of treatment (Days 89/90/91) in mean DBP
- Change from baseline (Days -1/1) to end of treatment (Day 90) in body weight
- Change from baseline (Day 1) to end of treatment (Day 90) in HbA1c
- Change from baseline (Days -1/1) to end of treatment (Days 90/91) in 9-point plasma glucose profiles (AUC)
- Change from baseline (Day -1) to end of treatment (Day 90) in FPG
- Change from baseline (Day -2) to end (Day 89) of treatment in liver fat
- Change in fasting insulin from baseline (Day 1) to end of treatment (Day 90)
- Change in 1,5 anhydroglucitol from baseline (Day 1) to end of treatment (Day 90)

- Change from baseline (Day -1/1) to end of treatment (Day 90) in waist circumference
- Change from baseline (Day -1) to end of treatment (Day 91) in PHQ-9 score

For all endpoints except PHQ-9 scores the individual data will be listed by visit, day and time point (data from visits and entries from diaries). Summaries will be made by visit (visit data only, no diary entries) and treatment arm, overall and stratified by trial site, if significant, presenting mean, standard deviation, minimum, median and maximum. The analysis will be conducted as described for the primary endpoint for visit data only.

PHQ-9 score will be listed individually including change from baseline and be summarized via frequencies by treatment, overall and stratified by trial site, if significant. The analysis of the PHQ-9 score will only be non-parametric (Wilcoxon Rank Sum test as described for the primary endpoint (not normally distributed)) because of its ordinal scale.

Box plots of baseline/end of treatment displaying the mean, quartiles, and minimum and maximum observation for all endpoints except PHQ-9 score will be presented for each treatment arm, overall and stratified by trial site, if significant.

For BP and FPG figures, which present the values over the course of the whole trial, will be plotted (visits and diary entries). These figures will be presented for each subject individually and as mean value graphs overlaid by treatment arm. The mean graphs will be plotted overall and stratified by trial site, if significant.

For HbA1c, body weight, waist circumference and PHQ-9 figures, which present the values over the course of the whole trial, will be plotted (visits). These figures will be presented for each subject individually and as mean value graphs overlaid by treatment arm. The mean graphs will be plotted overall and stratified by trial site, if significant.

5.8 Analysis and Presentation of Safety Endpoints

- AEs
- Number and severity of hypoglycemic events
- Clinical labs (Hematology, Blood Chemistry)
- ECG
- Vital signs
- Physical examination

The safety endpoints will be based on the Safety Analysis Set. Continuous data will be summarized using non-missing counts, mean, standard deviation, median, minimum, and maximum. Categorical data will be summarized using counts and percentages.

Any entries concerning AEs, hyperglycemia, hypoglycemia or hypotension symptoms, which are documented in the diaries, are treated the same as the records from the trial sites (AEs, hyperglycemia, hypotension at AEs; hypoglycemia at hypoglycemic episodes).

Treatment emergent AEs will be summarized by treatment arm, intensity, and relation to IMP. The descriptive statistics presented for each system organ class and preferred term will be the number of subjects with event (N), the percentage of subjects exposed with event (%), and the number of events (E). The following frequency tables will be provided by treatment arm:

- overview of incidence of AEs
- summary of AEs by MedDRA system organ class and MedDRA preferred term
- summary of AEs by categories: all, serious, non-serious (all / mild / moderate / severe), related (all / probable / possible / unlikely) and AE withdrawals

All AEs will be listed by subject, including demographic information (age and gender), treatment, system organ class and preferred term.

Individual listings of serious and non-serious AEs, and AEs leading to withdrawal and deaths will be presented by treatment arm, including the system organ class and the preferred term, the time of onset and duration, the time of last drug administration, the intensity, and relationship to the investigational product, the outcome of the AE, the action taken on the trial product, and the action taken to treat the AE.

Frequency analysis of AEs will be done using a logistic regression model (proc LOGISTIC) with total number of AEs as dependent variable and treatment and trial site as fixed effects. The odds ratio between the treatments as well as its respective p-value and 95% CI will be calculated. If 1 is included in the CI no significant difference between the treatment arms could be found.

Hyperglycemic episodes are part of the AEs.

Hypoglycemic episodes will be listed by subject, including demographic information (age and gender), treatment, including information if the episode was symptomatic and/or serious, the intensity, and relationship to the investigational product, the outcome and the action taken with the trial treatment as well as due to the event.

Summary of hypoglycemic events will be done by treatment arm and classification (minor, major and nocturnal.

Frequency analysis of hypoglycemic events will be done using the same model and approach as the analysis of the AEs.

Laboratory safety variables will be summarized by descriptive statistics (number, mean, SD, minimum, median and maximum) and by visit. All laboratory values will be listed by visit and subject number including demographic information and flagging of values outside normal rang. A listing of abnormal values will be presented in an end of text (EOT) listing.

All ECG results will be summarized using descriptive statistics by visit or by treatment at the treatment period, respectively. Abnormal evaluations together with the Investigator's comments will be listed. All individual data will be listed.

Vital signs will be summarized using descriptive statistics by visit. All individual data will be listed.

For physical examination all data (numbers and percentages) will be shown by a frequency table. All individual data will be listed.

5.9 Analysis and Presentation of Other Assessments

The summaries of the Other Assessments will be done based on the Safety Analysis Set or Full Analysis Set, respectively.

Concomitant medication data and pregnancy test data will be presented by individual listings. In addition, tabular summaries (by treatment) will be used.

HR data, which is recorded parallel to BP measurements at V2 and V10, will be presented by individual listings.

BP and FPG measurements, which are recorded in the subject's diaries, will be presented by individual listings and summarized by diary period (V2-V3, V3-V4, V4-V5, V5-V6, V6-V7, V7-V8, V8-V9 and V9-V10) and treatment arm.

Body weight measurements, which are not used for the calculation of the secondary endpoint, will be presented by individual listings and summarized by visit and treatment arm.

Waist circumference measurements, which not used for the calculation of the exploratory endpoint, will be presented by individual listings and summarized by visit and treatment arm.

PHQ-9 data, which is not used for the calculation of the exploratory endpoint, will be presented by individual listings and summarized by visit and treatment arm.

The measurements of HbA1c, fasting insulin and 1,5 anhydroglucitol, which are done from blood samples collected at visits 6 and 8, will be presented by individual listings and summarized by visit and treatment arm.

The study compliance will be presented by individual listings. In addition, tabular summaries (by treatment) will be used.

The measurements of tesofensine/metabolite and metoprolol concentration (PK data) will be presented by individual listings and summarized by visit and treatment arm. The summary additionally includes standard error (SE) and 95% CI of the mean. Furthermore scatterplots of the values from Day 90 will be presented by treatment arm one for tesofensine and metabolite and one for metoprolol.

If grouping (separated groups with low concentration values and high concentration values, respectively) is found within the PK data, further analyses may be done by group. This decision will be done at the DBR meeting.

5.10 Subject Disposition

Subject disposition will be tabulated for the whole population and stratified for trial site including the numbers of screened subjects, screening failures, subjects exposed to study product (overall and by treatment arm), subjects completing the study and subjects in the FAS, PPP and Safety Analysis Set.

For screening failures the reasons for exclusion will be recorded.

Subjects withdrawn from the study will be listed including the primary reason for withdrawal.

5.11 Interim Analysis

No interim analysis is planned.

5.12 Changes from the Protocol

The definition of the FAS was added. It was agreed to do the analyses based on the FAS.

The presentation of tesofensine/metabolite and metoprolol concentrations (PK data) was added.

6 ITEMS FOR STATISTICAL DOCUMENTATION

The statistical documentation to be included in an appendix to the CSR will include:

- Statistical Analysis Plan
- Relevant output from statistical analyses

7 SOFTWARE

All statistical analyses will be performed using SAS® (SAS Institute Inc., Cary, North Carolina, United States of America [USA]), version 9.4 or higher.

8 REFERENCES

1. ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials E9. International, Conference on Harmonisation E9 Expert Working Group, http://www.ich.org/LOB/media/MEDIA485.pdf. 5-Feb-1998.

2. Beard C, Hsu K, Rifkin LS, Busch AB, Björgvinsson T. Validation of the PHQ-9 in a psychiatric sample. J. Affect. Disord. 2016; 15:193:267-73.

9 APPENDICES

9.1 Appendix I: Study Flow Chart: Visits 1 - 6 (all visits are ±1 day)

Visit	1 Screening OP	2a Baseline IP	2a Baseline IP	2b Randomisation IP	3, 4, 5 OP	6 OP
Day	-28 to -5*	-2	-1	1	7,14,21	28
ICF, Inclusion/ Exclusion	X	X [f]				
Physical examination, height, BMI	X					
Med. History/ demographics	X					
Vital signs	X	X	X	X	X	X
PHQ-9 Score	X		X		X[g]	X
Admission		X				
Release				X		
Randomization				X		
Administration of first dose of IMP at the Unit				X		
24-h HR measurement			X initiation	X completion		
BP measurement		X initiation	X	X completion		
Weight, waist circumference	X		X	X	X	X
Hematology	X			X		
Blood chemistry	X			X		
Infectious serology	X					
ECG	X	X				
FPG	X[a]		X	X	X	X
HbA1c	X			X		X
1,5-anhydroglucitol, lipids, insulin	X[j]			X		X
9-p glucose profile			X	X		

			initiation	completion		
MRS (liver fat) [b]		X				
PK sample				X		
Lifestyle advice				X		
Medication dispensed				X	X	X
AEs				X	X	X
Hypoglycaemia			X	X	X	X
Hyperglycemia			X	X	X	X
Medication accountability					X	X
Concomitant medication	X		X	X	X	X
Blood pregnancy test [c]	X					
Urine pregnancy test [c]			X			X
Diary dispensation	X			X	X	X
Diary review		X	X	X	X	X
Diary collection		X			X	X
Glucometer dispensation	X					
BP measurement device dispensation [h]				X[i]		
Subject card dispensation	X					
End of study form						

[a]: During the whole study the subjects will be asked to check the FPG every day

[b] In a sub-set (Neuss)

[c]: Female subjects of childbearing potential only

[f] Only re-check of exclusion and inclusion criteria

[g] Only Day 14

[h] Only subjects who do not have an adequate one available

[i] Subjects will be requested to measure BP every day until the follow-up visit

[j] Lipids only

IP = in-patient; OP = out-patient; Ph = phone

*There will be up to additional 4 weeks in-between screening and admission in those subjects who will need to have their anti-diabetic medication washed out.

9.2 Appendix II: Study Flow Chart: Visits 7 - 12 (all visits are ±1 day)

Visit	7 OP	8 OP	9 OP	10 IP	10 IP	10 IP	11 F-up Phone	12 F-up OP
Day	42	56	70	89	90	91	98	110
Physical examination, height, BMI[a]					X			X
Vital signs	X	X	X	X	X	X		X
PHQ-9 Score	X		X			X		
Admission				X				
Release						X		
24-h HR measurement					X initiation	X completion		
BP measurement				X initiation	X	X completion		
Weight, waist circumference	X	X	X		X			X[b]
Hematology					X			
Blood chemistry					X			
ECG				X				
FPG	X	X	X		X	X		X
HbA1c		X			X			X
1,5-anhydroglucitol, lipids, insulin		X			X			X[c]
9-p glucose profile					X initiation	X completion		
MRS (liver fat)[b]				X				
PK sample					X			
Urinary pregnancy test		X						X
Medication dispensed	X	X	X			X[d]		
AEs	X	X	X		X	X	X	X
Hypoglycaemia	X	X	X		X	X	X	X
Hyperglycaemia	X	X	X		X	X	X	X

Medication accountability	X	X	X		X			
Concomitant medication	X	X	X		X	X	X	X
Diary dispensation	X	X	X	X				
Diary review	X	X	X	X	X			X
Diary collection	X	X	X	X				X
End of study form								X

[b] In a subset (Neuss)
[d] Dispensation of two half-doses of active / placebo metoprolol
[e] Except lipids and insulin
[g] Conducting physical examination (except height and BMI)
[h] Only body weight measurement

9.3 Appendix III: The Patient Health Questionnaire (PHQ-9; Version in German)

Gesundheitsfragebogen für Patienten (PHQ-9)

Wie oft fühlten Sie sich im Verlauf der letzten 2 Wochen durch die folgenden Beschwerden beeinträchtigt?	**Überhaupt nicht**	**An einzelnen Tagen**	**An mehr als der Hälfte der Tage**	**Beinahe jeden Tag**
a. Wenig Interesse oder Freude an Ihren Tätigkeiten	☐ 0	☐ 1	☐ 2	☐ 3
b. Niedergeschlagenheit, Schwermut oder Hoffnungslosigkeit.	☐ 0	☐ 1	☐ 2	☐ 3
c. Schwierigkeiten ein- oder durchzuschlafen oder vermehrter Schlaf	☐ 0	☐ 1	☐ 2	☐ 3
d. Müdigkeit oder Gefühl, keine Energie zu haben	☐ 0	☐ 1	☐ 2	☐ 3
e. Verminderter Appetit oder übermäßiges Bedürfnis zu essen	☐ 0	☐ 1	☐ 2	☐ 3
f. Schlechte Meinung von sich selbst; Gefühl, ein Versager zu sein oder die Familie enttäuscht zu haben	☐ 0	☐ 1	☐ 2	☐ 3
g. Schwierigkeiten, sich auf etwas zu konzentrieren, z.B. beim Zeitunglesen oder Fernsehen	☐ 0	☐ 1	☐ 2	☐ 3
h. Waren Ihre Bewegungen oder Ihre Sprache so verlangsamt, dass es auch anderen auffallen würde? Oder waren Sie im Gegenteil „zappelig" oder ruhelos und hatten dadurch einen stärkeren Bewegungsdrang als sonst?	☐ 0	☐ 1	☐ 2	☐ 3
i. Gedanken, dass Sie lieber tot wären oder sich Leid zufügen möchten	☐ 0	☐ 1	☐ 2	☐ 3

Gesamtwert _____ = Addition ____ + ____ + ____
der Spaltensummen