CONFIDENTIAL

STATISTICAL ANALYSIS PLAN

Protocol No. PHN-Udenafil-02

IND# 121648

FONTAN UDENAFIL EXERCISE LONGITUDINAL ASSESSMENT TRIAL (FUEL)

Mezzion Pharma Ltd. in partnership with Pediatric Heart Network

Version:

23 June 2016

CONFIDENTIAL

SIGNATURES OF AGREEMENT FOR STATISTICAL ANALYSIS PLAN

Name Date

Title

Affiliation

Name Date

Title

Affiliation

Name Date

Title

Affiliation

CONFIDENTIAL

TABLE OF CONTENTS

List of Abbreviations 5

1.0 Study Objectives 6

2.0 Overall Study Design 6

3.0 Sample Size 7

4.0 Drug Regimens 7

5.0 Study Procedures 7

6.0 Statistics 9

6.1 Analysis Populations 9

6.2 Accountability and Protocol Deviations 10

6.3 Demographic and Baseline Characteristics 10

6.4 Efficacy Analyses 10

6.4.1 Primary Efficacy Endpoint (Aerobic Exercise Performance) 10

6.4.2 Secondary Efficacy Endpoints 11

6.4.3 Missing Data 12

6.4.4 Subpopulations 13

6.4.5 Adjustments for Multiple Comparisons 13

6.5 Safety Evaluations 13

6.5.1 Extent of Exposure 13

6.5.2 Adverse Events 14

6.5.3 Clinical Laboratory Tests 15

6.5.4 Vital Signs 16

6.6 Prior and Concomitant Medications 16

6.7 Telephone Interview 17

6.8 Interim Analyses 17

7.0 Quality Control and Software 17

8.0 Changes from the Protocol..... 18

List of Abbreviations

ANCOVA	analysis of covariance
ATC	Anatomical-Therapeutic-Chemical classification
BNP	brain-type natriuretic peptide
HAES	Habitual Activity Estimation Scale
ICH	International Conference on Harmonisation
ITT	intent-to-treat
MedDRA	Medical Dictionary for Regulatory Activities
MPI	myocardial performance index
PAT	pulse amplitude tonometry
PCQLI	Pediatric Cardiac Quality of Life Inventory
PDE5	phosphodiesterase type 5
PedsQL	Pediatric Quality of Life Inventory
SAE	serious adverse event
TEAE	treatment-emergent adverse event
VAT	ventilator anaerobic threshold
VE/VCO_2	ventilatory equivalents of carbon dioxide
VO_2	minute oxygen consumption

1.0 Study Objectives

The purpose of this study will be to evaluate the clinical efficacy and safety of udenafil, an orally administered, potent and selective inhibitor of phosphodiesterase type 5 (PDE5), versus placebo for the treatment of adolescent subjects who have undergone the Fontan procedure

2.0 Overall Study Design

This study is a randomized, double-blind, 26-week, multicenter trial of udenafil (87.5 mg, twice daily) versus placebo on aerobic exercise performance as measured by change in minute oxygen consumption (VO_2) at maximum exercise effort (maximal VO_2) from baseline to six months in adolescent survivors of the Fontan procedure.

Eligible subjects will be randomly assigned to one of two treatment groups (udenafil or placebo) by a web-based system at the data coordinating center. Randomization will occur within strata defined by ventricular morphology (single left versus single right or mixed).

The first dose of randomized study drug will be administered at the clinic (Visit 1) and outpatient dosing will be used for the remainder of the 26-week study. Subjects will return to the clinic for assessments at 2 weeks (Visit 2), 13 weeks (Visit 3), and 26 weeks (Visit 4) or study completion after initiation of study drug dosing.

Maximal VO_2 will be assessed at Visits 1 and 4 (or study completion) using a braked cycle ergometer following a ramp protocol. Additional assessments will include ventricular performance as measured by echocardiogram, endothelial function assessment with EndoPat®, and serum brain-type natriuretic peptide (BNP) level. Quality of life will be assessed with the Pediatric Quality of Life Inventory (PedsQL), Pediatric Cardiac Quality of Life Inventory (PCQLI), and the Habitual Activity Estimation Scale (HAES).

Blood samples for pharmacokinetic assessments of udenafil and its primary metabolite will be obtained after the first dose of randomized study drug and at Visit 4 (or study completion). Safety will be monitored by adverse event reports, resting vital signs, and clinical laboratory tests.

When an individual subject completes the study, the subject's primary cardiologist will be notified, and the study drug will be stopped; there will be no need to wean subjects off the study drug. Subjects will not be told which drug they received until all recruitment and follow-up is completed per-protocol. All subjects will be invited to participate in an open-label, 12-month safety extension trial of udenafil.

The blind for the study will be broken when the later of the following 2 events occur and all database queries have been resolved:

- Visit 4 is completed for the last subject who enters the safety extension trial, or
- The 90-day follow-up is completed for the last subject who does not enter the safety extension trial.

3.0 Sample Size

A sample size of 198 subjects per group will provide 90% power to detect a mean treatment group difference in change from baseline to Week 26 in maximal VO_2 scores of 10%; that is, an improvement of 2.8 mL/kg/min in the udenafil group compared to zero change in the control group, assuming a Type I error of 0.05 and standard deviation of 7.235. A difference of 2.8 represents approximately 0.4 standard deviations, a difference equivalent to a 10% increase from an expected baseline of 28 mL/kg/min.

4.0 Drug Regimens

Subjects will be randomized to either oral udenafil or matching placebo. Subjects will take udenafil 87.5 mg or matching placebo twice a day. Duration of study drug administration will be six months. Study drug may be temporarily discontinued for up to 2 weeks for side effects. Study drug may be resumed on resolution of symptoms or side effects if considered safe in the judgment of the investigators and the subject's primary physician and with consent of subject.

Study drug will be permanently discontinued if deemed necessary by the study investigator or primary physician either due to side-effects or due to need for open-label drug administration of a PDE5 inhibitor.

5.0 Study Procedures

Study procedures will be performed as outlined in **Error! Reference source not found.**. In-clinic procedures associated with the first dose of study drug at Visit 1 are outlined in Table 1.

CONFIDENTIAL

Table 1 Schedule of Measurements/Events

Visit Number	1	TC1	TC2	2	TC3-5	3	TC6-7	4	TC8-10
Time point	Screening/ Baseline Day 0	Day 1	Week 1	Week2	Weeks 3, 4, 8	Week 13	Week 17, 21	Week 26	Week 30, 34, 39
Visit Window(s)			± 3 days	± 3 days	± 3 days	±10 days	± 3 days	±10 days	± 10 days
Type of Visit(s)	In-Person	Call	Call	In-Person	Call	In-Person	Call	In-Person	Call
Informed Consent/Assent	X								
Subject ID# assigned	X								
Inclusion/Exclusion Criteria	X								
Physical Measurements	X								
Medical history	X								
Demographics/expanded demo	X								
Prior/Concomitant Meds	X		X	X	X	X	X	X	
EndoPAT®	X							X	
Serum or urine pregnancy test[a]	X			X		X		X	
Clinical Laboratory tests (creatinine, ALT, AST)	X							X	
Biomarker (BNP) sample	X							X	
Genetic repository sample (optional)	X[a]							X[b]	
Echocardiogram	X							X	
Exercise test	X[c]							X[c]	
Randomization	X								
Dispense Study Drug	X			X		X			
Resting BP and HR	X			X		X		X	
Administer first dose of study drug (see Additional Procedures and Measurements in Table 2)	X								
Peds QL generic and cardiac modules; PCQLI; HAES	X							X	
Drug Accountability	X		X	X	X	X	X	X	
Pregnancy counseling (if applicable)	X		X	X	X	X	X	X	
Adverse events assessment	X	X	X	X	X	X	X	X	X[a]
Morisky Scale (MMAS)			X		X		X		
Schedule/Confirm next visit	X	X	X	X	X	X	X	X	

[a]as applicable

[b]as applicable if missed at baseline visit

[c]Baseline visit: If failed 1st exercise test, do NOT randomize; At baseline and Week 26 visits, at Site Investigator's discretion, exercise test may be repeated w/in 14 days -- If a second exercise test needed, you MUST perform urine (or blood) pregnancy test for eligibility (if applicable).

Table 1 Schedule of Assessments at Visit 1

Procedure	Time From Administration of First Dose of Study Drug			
	< 1 Hour Prior to Dose	**0 Hour**	**2 Hours (± 30 Minutes)**	**4Hours (± 30 Minutes)**
Resting vital signs[a]	X			X[b]
Administer first dose of study drug		X		
Vital signs immediately before and following a 6-minute self-limited walk			X	
Pharmacokinetic blood sample immediately after 6-minute walk			X	
Record adverse events		X	X	X
Dispense study drug for home administration				X
Discharge from clinic				X

a Includes blood pressure and heart rate.

b Before discharge from the clinic.

6.0 Statistics

All statistical tests will be at the 0.05 alpha level, 2-tailed, unless stated otherwise. Probabilities will be rounded to 4 decimal places before assigning statistical significance.

Descriptive statistics for categorical endpoints will include the number and percent of subjects in each treatment group and category. Quantitative endpoints will be summarized for each treatment group with the mean, median, standard deviation, minimum value, and maximum value.

Individual subject data will be presented in data listings, which will include normal reference ranges when appropriate.

6.1 Analysis Populations

Three analysis populations will be defined:

- Intent-to-treat (ITT) Population: The population will include all randomized subjects and will be the primary population for efficacy analyses. Treatment assignments will be analyzed according to the randomized treatment assignment.

- Safety Population: The population will include all subjects who take at least one dose of randomized study drug and will be the primary population for safety analyses. If the wrong study drug is administered (e.g., subject is randomized to undenafil but receives placebo), treatment assignments will be analyzed according to the actual study drug received.

- Per Protocol Population: The population will include all subjects in the Safety Population who meet all entry criteria, or if criteria were not met, were granted a waiver by the Sponsor. Subjects with major protocol deviations will be excluded

(including subjects who receive the wrong study drug). The Per Protocol Population will be used for sensitivity analyses of the primary efficacy endpoint.

6.2 Accountability and Protocol Deviations

Screening failures (i.e., subjects who were not randomized) will be summarized by the number and percentage of failures for each primary reason.

The number and percent of subjects who are randomized, treated with randomized study drug, prematurely discontinue, and complete the study will be summarized by treatment group. The number and percent of subjects will be summarized by treatment group for each reason for premature discontinuation. For subjects who complete the study, the number of subjects who continue into the extension study will be summarized.

The number and percent of enrolled subjects will be summarized by study site.

The number of subjects included in each analysis population will be summarized by treatment group.

Major protocol deviations will be identified prior to breaking the blind. Deviations may include, but are not limited to, departure from inclusion/exclusion criteria, received the wrong study drug; failure to perform the required assessments at specified time points; and scheduling of visits not in accordance with specifications.

6.3 Demographic and Baseline Characteristics

Demographic (e.g., age, sex, and race) and baseline characteristics will be summarized descriptively for the Safety and ITT Populations. Summaries will be provided for all subjects and for each treatment group.

Medical and surgical history/physical findings will be classified by system organ class and preferred term using the Medical Dictionary for Regulatory Activities (MedDRA). The number and percentage of subjects with abnormalities in medical and surgical histories in each system organ class and preferred term will be summarized for all subjects and by treatment group for the Safety Population.

6.4 Efficacy Analyses

6.4.1 Primary Efficacy Endpoint (Aerobic Exercise Performance)

The treatment group difference for change in aerobic exercise performance (as measured by VO_2 at maximum exercise effort [maximal VO_2]) from baseline to Visit 4 will be assessed with an analysis of covariance (ANCOVA) with fixed factors for ventricular morphology (single left versus single right or mixed) and treatment group with a continuous covariate of baseline maximal VO_2.

Sensitivity analyses will include:

- Primary efficacy analysis performed for the Per Protocol Population
- ANCOVA model to assess treatment group differences in ranked changes of maximal VO_2 for the ITT and Per Protocol Populations.

6.4.2 Secondary Efficacy Endpoints

Secondary efficacy endpoints include change from baseline to Visit 4 (Week 26) in:

- Exercise capacity
    - Minute oxygen consumption at the ventilator anaerobic threshold (VAT)
    - Ventilatory equivalents of carbon dioxide (VE/VCO_2) at VAT
    - Respiratory rate and minute ventilation at peak exercise
    - Maximal work rate
    - Work rate at the VAT
- Ventricular performance
    - Myocardial performance index (MPI) determined by velocities obtained from blood pool Doppler assessment of the inflow and outflow tract of the dominant ventricle. It is the primary performance endpoint of the echocardiogram.
    - Ventricular cavity size, eccentricity, and mass (based on echocardiogram)
    - Systolic function estimate using mean dP/dt during isovolumetric contraction (dP/dt_{ic}) and peak systolic annular velocity (S') on tissue Doppler
    - Qualitative and quantitative estimate of atrioventricular valve insufficiency
- Endothelial function: log-transformed reactive hyperemia index (lnRHI), as measured by pulse amplitude tonometry (PAT) testing using the EndoPAT® device
- Natural logarithm transformation of BNP
- Functional health status as measured by the full scale PedsQL
    - Physical functioning score

  - Psychosocial functioning score
  - Cardiac-specific quality-of-life score
- Pediatric Cardiac Quality of Life Inventory (PCQLI) score

Each secondary efficacy endpoint will be summarized descriptively by treatment group.

Treatment group differences for endpoints associated with the PedsQL and PCQLI will be assessed with Friedman's test using ventricular morphology (single left versus single right or mixed) as the stratification factor. The treatment group difference for the qualitative estimate of atrioventricular valve insufficiency will be assessed with a Cochran-Mantel-Haenszel test using ventricular morphology (single left versus single right or mixed) as the stratification factor.

The remaining secondary efficacy endpoints will be analyzed with the ANCOVA model with fixed factors for ventricular morphology (single left versus single right or mixed) and treatment group with a covariate for the baseline value.

6.4.3 Missing Data

The following imputation method will be used for the primary analysis of the primary outcome:

- Subjects who have died or have dropped out of the study with unknown vital status will be assigned a maximal VO_2 of zero at Visit 4.
- Subjects who are known to be alive, but who discontinued from the study (and are missing maximal VO_2 at Visit 4) will be assigned the latest value available (e.g., value from end of study visit if available, and from baseline visit otherwise).
- Subjects who complete Visit 4, but are physically unable to reach maximum effort in cardiopulmonary exercise testing after two attempts, will be assigned their baseline value (i.e., zero change).

Two additional analyses to assess the impact of missing data will be conducted as follows:

- Subjects with missing maximal VO_2 at Visit 4 will be assigned a maximal VO_2 of zero
- Subjects with missing maximal VO_2 at Visit 4 will be excluded from analysis (i.e., observed cases analysis).

Missing item scores for rating scales will be imputed according to published methodology for the scale. Otherwise, missing data will not be estimated for secondary efficacy endpoints.

6.4.4 Subpopulations

Treatment group differences for the primary efficacy endpoint will be summarized descriptively for the following subpopulations based on the following baseline characteristics:

- Gender (female, male)
- Race (Asian, African-American, Caucasian, other)
- Ethnicity (Hispanic, non-Hispanic)
- Ventricular morphology (single left, single right or mixed)
- Age at Fontan surgery (< 3 years, ≥ 3 years)
- Baseline serum BNP level (< median, ≥ median)
- Percent of predicted maximal VO_2 at baseline (< 75%, ≥ 75%)

A 95% confidence interval of the difference in treatment group means will be provided within each subpopulation.

6.4.5 Adjustments for Multiple Comparisons

Formal hypothesis testing will be performed only for the primary efficacy endpoint. Thus, no adjustment for multiple comparisons will be necessary.

6.5 Safety Evaluations

Safety assessments will be summarized for the Safety Population. No formal hypothesis testing will be performed.

6.5.1 Extent of Exposure

The number of days from first to last dose will be summarized descriptively for each treatment group. Percent compliance for the entire study will be calculated as follows:

Number tablets missed = (number returned – (2 × number of days off study))

Number tablets expected = 2 × number of days on study

Percent Compliance = (1 - (number missed / number expected)) × 100.

A subject will be considered on-study from Visit 1 until study completion or premature discontinuation of study drug. Percent compliance will be summarized descriptively for each treatment group.

6.5.2 Adverse Events

The number and proportion of subjects who report treatment-emergent adverse events (TEAEs) will be summarized for each treatment group. A TEAE is an event that begins after receipt of randomized study drug. All adverse events will be classified by MedDRA with respect to system organ class and preferred term.

The adverse event profile will be characterized with intensity and relationship to study drug (related, unrelated, unknown). Related adverse events will be defined as events considered possibly or probably related to treatment by the investigator. Events with unknown severity or relationship will be counted as unknown. The incidence of serious adverse events will be summarized similarly.

Subjects who report the same preferred term on multiple occasions will be counted once for the preferred term: under the highest severity when summarized by severity and under the closest relationship to study drug when summarized by relationship. If a subject reports multiple preferred terms for a system organ class, the subject will be counted only once for that system organ class.

Proportions for adverse events that are gender-specific (e.g., dysmenorrhea) will be based on the number of subjects from that gender.

The number and percentage of subjects who experience TEAEs will be summarized by treatment group for the following:

- By system organ class and preferred term
- By intensity (mild, moderate, or severe/life threatening/death), system organ class, and preferred term
- By relationship to study drug (not related or related), system organ class, and preferred term
- By weight tertile, system organ class, and preferred term
- By age tertile, system organ class, and preferred term
- Serious adverse events (SAEs) by system organ class and preferred term
- SAEs by relationship to study drug (not related or related), system organ class, and preferred term
- SAEs by weight tertile, system organ class, and preferred term
- SAEs by age tertile, system organ class, and preferred term

- Adverse events resulting in discontinuation of study drug by system organ class and preferred term

An adverse event cluster including occurrences of hypotension (e.g., blood pressure decreased, blood pressure diastolic decreased, blood pressure systolic decreased, hypotension, procedural hypotension), loss of consciousness, dizziness (e.g., dizziness, dizziness exertional, procedural dizziness), presyncope, or syncope (syncope, syncope vasovagal) will be summarized as follows:

- The number and percent of subjects reporting at least 1 TEAE from the above cluster will be summarized by treatment group.
- The number and percent of patients reporting each of the preferred terms will also be summarized by treatment group.
- The 2 summaries will also be repeated by
    - age tertiles
    - weight tertiles.

An adverse event cluster including death, hospitalization for heart failure, and transplant will be summarized as follows:

- The number and percent of subjects reporting at least 1 TEAE from the above cluster will be summarized by treatment group.
- The number and percent of patients reporting each of the preferred terms will also be summarized by treatment group.
- The 2 summaries will also be repeated by
    - age tertiles
    - weight tertiles.

6.5.3 Clinical Laboratory Tests

Laboratory values will be converted to the project-defined unit of measurement before analysis.

Clinical laboratory variables will be presented in 2 ways. First, change from baseline to Visit 4 (study completion) will be summarized descriptively for each treatment group. The baseline value will be defined as the last assessment on or before dosing at Visit 1.

Second, the number and proportion of subjects with treatment-emergent abnormal laboratory values will be tabulated and the subjects identified. Treatment-emergent abnormal laboratory tests are those in which the baseline value is normal (within the laboratory normal reference range) and post-baseline value is abnormal (i.e., meets Grade III or Grade IV toxicity criteria from the National Cancer Institute Common Terminology Criteria. All laboratory values obtained after Visit 1 will be included in the analysis.

6.5.4 Vital Signs

Change from baseline (last value before dosing at Visit 1) to each scheduled assessment will be summarized descriptively by treatment group for each vital sign variable.

Vital signs that are potentially clinically significant will be identified with the criteria in Table 2 and summarized descriptively by treatment group. Baseline will be defined as the last vital sign value obtained before the first dose of study drug.

Table 2 Criteria for Potentially Clinically Significant Vital Signs

Vital Sign	Criterion	Definition of Potentially Clinically Significant Value
Systolic blood pressure	Low	Value ≤ 80 mmHg and decreased ≥ 20 mmHg from value before dosing
	High	Value ≥ 140 mmHg and increased ≥ 20 mmHg from value before dosing
Diastolic blood pressure	Low	Value ≤ 50 mmHg and decreased ≥ 15 mmHg from value before dosing
	High	Value ≥ 100 mmHg and increased ≥ 15 mmHg from value before dosing
Pulse	Low	Value ≤ 45 bpm and decreased ≥ 15 bpm from value before dosing
	High	Value ≥ 130 mmHg and increased ≥ 15 mmHg from value before dosing

bpm = beats per minute

6.6 Prior and Concomitant Medications

The Concomitant Medications World Health Organization drug dictionary will be used to classify all medications with respect to the Anatomical-Therapeutic-Chemical classification (ATC system) and preferred drug name. Prior and concomitant drug usage will be summarized by ATC level 3 and preferred drug name. The summary will provide

the number and percent of subjects in each treatment group who receive at least 1 non-study medication.

Medications with a start date before the first dose of study drug will be considered prior medications. Medications with a stop date after the first dose of study drug or ongoing at study completion/discontinuation will be considered concomitant medications. Therefore, medications that start before the study and continue into the study will be counted as both prior and concomitant medications.

For the Safety Population, separate summaries will be provided for prior and concomitant medications.

6.7 Telephone Interview

Patient responses to questions during the telephone interview will be summarized descriptively. Responses to these questions have not been validated as indicators of treatment compliance.

6.8 Interim Analyses

A Data Safety Monitoring Board will meet at least twice per year to review safety data. Formal stopping boundaries are not proposed for the 26-week trial. However, the Data Safety Monitoring Board may recommend stopping the study for other reasons, such as safety findings from this trial and other studies, or concerns about study conduct.

In addition, premature termination of this study may occur due to the impact of results released from other studies, due to failure to enroll, or due to withdrawal of study approval by clinical site Institutional Review Boards. In addition, the National Heart, Lung and Blood Institute retains the right to discontinue the study prior to the inclusion of the intended number of subjects, but intends to exercise these rights only for valid scientific or administrative reasons.

After primary outcome data is obtained for approximately half of the originally planned sample size (approximately N=200 subjects) the variance of the primary outcome will be estimated, using a blinded method with lumped variance. If this estimated variance is higher than the one used for the sample size calculations, then the sample size will be recalculated and increased correspondingly. Otherwise, the sample size doesn't change and the trial proceeds as planned. This approach is based on blinded re-estimation of nuisance parameters and doesn't lead to inflation of type I error. It is consistent with the regulatory guidance for industry in adaptive design of clinical trials.

7.0 Quality Control and Software

All data displays and analyses will adhere to the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use

Harmonized Tripar tite Guideline: Structure and Content of Clinical Study Reports (ICH Topic E3).

All analyses will be performed using SAS® Version 9 (or later) software. The responsible statistical contract research organization will follow the company's standard operating procedures in the creation and quality control of all tables, listings, figures and analyses. The Sponsor or its designee will review all tables, listings, and figures for accuracy.

8.0 Changes from the Protocol

Key changes to statistical analyses between protocol finalization and this Statistical Analysis Plan are described below.

Analysis Populations

The protocol defined 2 non-ITT analysis populations based on compliance with use of PDE5 inhibitors (udenafil or other medication) and compliance with study drug dosing, respectively. These populations were replaced by the more traditional Per Protocol Population based on major protocol deviations.

Subpopulation Analyses

The protocol stated that potential differential impact of continuous age on treatment outcomes would be assessed and that a treatment by subpopulation interaction test would be conducted to identify differential treatment effects. Tests for treatment by subpopulation will not be performed due to the low power of the tests.

The protocol also stated that treatment group differences would be assessed within subpopulations if the interaction p-value were < 0.10. These tests will not be performed due to their potentially low power and bias due to lack of randomization within each subpopulation.

Missing Values

Minor clarifications were made.