Statistical Analysis Plan

What is the effect of intravenous iron supplementation on exercise capacity and quality of life in patients with IPAH and iron deficiency?

Version: 4

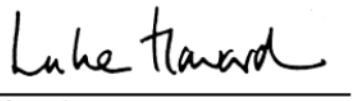
Chief Investigator: Dr Luke Howard DPhil, FRCP

Document Version History

Version Number	Reason for Update	Updated By:	Date
1	Creation of new SAP	Les Huson	16 Oct 2017
2	Adding more detail	Les Huson	30 Oct 2017
3	Adding more detail	Les Huson/Luke Howard	6 Nov 2017
4	Remove multiplicity adjustment	Luke Howard	13 Nov 2017

Signatures

Chief Investigator
Dr. Luke Howard



Signature

13 December 2017
Date

Trial Statistician
Dr Les Huson

Signature

13th December 2017
Date

1. Introduction

This Statistical Analysis Plan (SAP) describes the analyses that will be performed on the primary and secondary efficacy endpoints recorded in two Phase II proof-of-concept studies of the use of intravenous iron infusion to treat iron deficient patients with IPAH. The SAP is based on version 9.0 of the UK/EU study protocol, which is dated 24th October 2016, and on version 1.0 of the China study, which is dates 8th August 2014.

Both studies are double-blind, placebo controlled crossover trials, one of which recruits patients in various centres in the UK and EU, and the second of which recruits patients in China.

The two studies are very similar in design, the major difference being that the UK/EU study uses Ferinject® for the treatment of iron deficiency, whereas the China study uses Cosmofer®. There are also some differences between the endpoints recorded in the two studies (documented below).

In both studies patients are allocated at random to receive either active treatment or placebo in the first of two treatment periods, with the other treatment being administered, following a washout phase, in the second treatment period.

2. Sample size

The original version of both protocols contained the following statement on sample size:

"The null hypothesis is that iron replacement has no effect on patients with iron deficiency and IPAH, with a two-sided alternative that iron replacement changes PVR in patients with iron deficiency and IPAH. Assuming a standard deviation of 250 dynes/s/cm and a drop-out rate of 10%, 60 patients randomised 1:1 would give this study 80% power to detect a 194 dynes/s/cm reduction in PVR with iron treatment at a significance level of p=0.05 using two-sample t-test."

However, a blinded conditional power analysis of the PVR endpoint, conducted in April 2015 for the UK/EU study only, indicated that the study had conditional power (under the current trend hypothesis) of less than 60% to detect a statistically significant effect.

The current version of the UK/EU protocol contains an amended primary endpoint, applicable to the UK/EU study, and makes the following statement on power and sample size, which is based on a planned recruitment of 60 patients chosen on the basis of safety and feasibility considerations:

"For the cardiopulmonary exercise tests, (with) a 12-weekly drop-out rate of 10% and standard deviation of 5 ml/min/kg for peak VO2 we would have 80% power to detect a mean change of 1.94 ml/min/kg at a significance level of alpha=0.05 using a paired analysis"

3. Efficacy Endpoints

For both studies, the original protocol specified the following primary endpoint:

"Change in resting Pulmonary Vascular Resistance (PVR) between baseline and 12 weeks, measured by cardiac catheterisation".

In June 2015, following the blinded conditional power calculation, a protocol amendment was made for the UK/EU study which changed the primary endpoint from PVR to Endurance Time and cardiac catheterisation was dropped from the protocol. The China study continues to have PVR as the primary endpoint. Apart from the cardiac catheterisation procedure, the other main difference between the UK/EU study and the China study is that the former uses both incremental and endurance exercise tests, whereas the latter uses incremental exercise tests only.

The following efficacy endpoints are therefore specified in the UK/EU study protocol:

Primary:

• Endurance time at the end of endurance bicycle cardiopulmonary exercise testing at 80% peak work rate determined from the incremental exercise test

Secondary:

• Incremental bicycle cardiopulmonary exercise testing - peak VO2 (ml/min/kg), VO2 at metabolic threshold, VE/VCO2 slope, VO2/WR slope and O2 pulse.
• VO2 at the end of endurance bicycle cardiopulmonary exercise testing at 80% peak work rate determined from the incremental exercise test
• Gas exchange at 3 minutes after the start of the work phase of an endurance exercise test at 80% peak work rate determined from the incremental exercise test
• Iron indices – serum iron, transferrin saturations, ferritin, soluble transferrin receptor (sTfR), unsaturated iron binding capacity (UIBC), red cell distribution width (RDW) and erythropoietin (EPO) levels
• 6 minute walk distance and Borg dyspnoea scale
• NYHA WHO functional class
• NT-pro-BNP
• Quality of life (CAMPHOR questionnaire) and the self-reported Patient Global Assessment
• Cardiac MRI - right ventricular volumes, mass, ejection fraction, stroke volume and diastolic function (at sites where facility is available).

The following efficacy endpoints are specified in the China protocol:

Primary:

The change in resting pulmonary vascular resistance (PVR) between baseline and 12 weeks, measured by cardiac catheterisation.

Secondary:

- Incremental bicycle cardiopulmonary exercise testing - peak VO2 (ml/min/kg), anaerobic threshold, VE/VCO2 slope, VO2/WR slope, O2 pulse and tissue oxygenation index.
- Resting cardiopulmonary haemodynamics – right atrial pressure, pulmonary arterial pressure, pulmonary capillary wedge pressure, cardiac output and stroke volume
- Iron indices – serum iron, transferrin saturations, ferritin, unsaturated iron binding capacity (UIBC), red cell distribution width (RDW).
- 6 minute walk distance and Borg dyspnoea scale
- NYHA WHO functional class
- NT-pro-BNP
- Safety - the occurrence of adverse events
- Cardiac MRI - right ventricular volumes, mass, ejection fraction, stroke volume and diastolic function

4. Statistical Analyses

The principal objective of all of the statistical analyses planned for these studies is to test for a treatment effect (i.e. a difference between control and test treatments) and to produce suitable interval estimates of the effect, for each of the efficacy endpoints.

The overall strategy for the statistical analysis is that each of the two studies will first be analysed separately, using the methods outlined below, and results will be reported separately for the two studies. Following these separate analyses, and subject to satisfactory evidence of homogeneity, the individual results (treatment effect estimates and/or associated p-values) will be combined using meta-analysis techniques, in order to produce overall summaries of the effects of iron supplementation. If there is evidence that the results from the two studies cannot validly be combined, the results from individual studies only will be reported.

In order to facilitate the planned meta-analysis of the two sets of study results, each of the endpoints specified above, whether in the UK/EU study or the China study, will be analysed for both studies (if the relevant data are available), so that the individual study results can be combined. This means that some endpoints will be analysed in each study that are not explicitly specified in the protocol for that study.

All primary and secondary efficacy endpoints will be analysed using linear mixed models appropriate for a crossover design. The linear models will include administration sequence, period and treatment as fixed effects and subject as a random effect, and all models will include relevant and statistically significant baseline covariates. If the assumption of normal errors appears inappropriate for any of these linear models, (as evidenced by initial exploratory analyses), the linear models will use a suitable alternative error structure.

From the linear mixed models estimates of the treatment effect (difference in mean outcomes between the two treatment arms) will derived. These will summarised using adjusted (least-squares) means and mean differences, together with 95% confidence intervals and associated p-values. P-values testing for sequence and period effects will also be derived for each of the linear mixed models.

In any case in which the use of a linear mixed model appears, in initial exploratory analyses, to be inappropriate for the data structure, suitable nonparametric statistical procedures will be employed. In particular, for cardiac MRI outcomes, where the recorded endpoint will be multivariate in nature, the multi-response permutation procedure (MRPP) will be used to test for a treatment effect.

The main analyses of the efficacy endpoints will be performed using the Intention-to-treat (ITT) analysis set, with imputation of missing data, if required, carried out using multiple imputation (MCMC) techniques.

In addition to the ITT analyses, exploratory analyses of other patient subsets, including centre-specific subsets, may be performed.

For the meta-analysis of the two studies, treatment effect estimates will be derived by combination of the individual study effect estimates and/or by direct pooling of individual patient data, and p-value combination methodologies (i.e. combinations of the individual study p-values) will be used to derive an overall meta-analytic p-value from the combined results.