STATISTICAL ANALYSIS PLAN

A randomised, controlled, assessor-blind, parallel groups, multicentre trial comparing the efficacy and safety of highly purified human chorionic gonadotropin (HP-hCG) and recombinant human chorionic gonadotropin (rhCG) for triggering of final follicular maturation in women undergoing controlled ovarian stimulation

FASHION
(Efficacy and safety of HP-hCG and rhCG for triggering of final follicular maturation)

Trial 000191

Investigational Product:	FE 999086, HP-hCG (CHORAPUR)
Indication:	For Assisted Reproductive Technology (ART) programme such as in vitro fertilisation: triggering of final follicular maturation and luteinisation after stimulation of follicle growth
Phase:	III
Author:	[redacted]
Date of issue:	11 FEB 2016
Version:	1.0

Change log

Version No.	Effective Date	Reason for the Change / Revision	Supersedes
1.0	11 FEB 2016	New document.	None

Signed[I] agreement on Statistical Analysis Plan

Author:

[redacted]
Global Biometrics, Clinical R&D
Ferring Pharmaceuticals A/S

Review:

[redacted]
Reproductive Health, Clinical R&D
Ferring Pharmaceuticals A/S

[redacted]
Global Pharmacovigilance, Clinical R&D
Ferring Pharmaceuticals A/S

[redacted] Clinical R&D
Ferring Pharmaceuticals A/S

[redacted]
Global Biometrics, Clinical R&D
Ferring Pharmaceuticals A/

[redacted]
Global Biometrics, Clinical R&D
Ferring Pharmaceuticals A/S

Approval:

[redacted]
Global Biometrics, Clinical R&D
Ferring Pharmaceuticals A/S

[I] Can be replaced by a list of reviewers/approvers if sign off can be performed in the REAL system

TABLE OF CONTENTS

1 Introduction

This statistical analysis plan (SAP) is based on the protocol for trial 000191 (Version 1.0, 30 January 2015) and describes the planned statistical analyses for the period from screening up to 5-6 weeks after transfer (end-of-trial).

A separate SAP will be prepared covering post-trial information (live birth and neonatal health) for subjects with a vital pregnancy.

1.1 Definitions/ Abbreviations

1.1.1 Definition of Terms

Terms	Definitions
Gestational age	Days between date of embryo transfer and date of birth + 17 days or days between date of blastocyst transfer and date of birth + 19 days
Primary infertility	No previous clinical pregnancy
Randomised	Subject randomised to trial treatment
Randomisation stratum	Number of follicles ≥12 mm at the end of stimulation: <10 or ≥10 based on the available data of the follicles (e.g. not necessarily the stratum that was used when the subject was randomised)
Screened	Subject who enters the screening phase
Screening failures	Screened subjects that are not randomised

1.1.2 Abbreviations

Abbreviations	Meaning of abbreviations in document
AE	adverse event
βhCG	beta unit of human chorionic gonadotropin
BMI	body mass index
DMC	Data Monitoring Committee
eCRF	electronic Case Report Form
FAS	full analysis set
GnRH	gonadotropin-releasing hormone
hCG	human chorionic gonadotropin
HP-hCG	highly purified human chorionic gonadotropin
ICSI	intracytoplasmic sperm injection
IM	intramuscular(ly)
ITT	intention-to-treat
IQ-range	inter quartile range
LLOQ	lower limit of quantification
MII	metaphase II
IMP	investigational medicinal product
OHSS	ovarian hyperstimulation syndrome
OR	oocyte retrieval
PN	pronuclei
PP	per-protocol
PT	preferred term
rhCG	recombinant hCG
SAE	serious adverse event
SC	subcutaneous(ly)
SOC	system organ class
ULOQ	upper limit of quantification

2 Trial Objectives and Endpoints

2.1 Objectives

Primary Objective

- To demonstrate non-inferiority of HP-hCG 5,000 IU compared with rhCG 250 µg for triggering of final follicular maturation with respect to number of oocytes retrieved in women undergoing controlled ovarian stimulation

Secondary Objectives

- To evaluate HP-hCG and rhCG with respect to oocyte fertilisation
- To evaluate HP-hCG and rhCG with respect to pregnancy
- To compare HP-hCG with rhCG with respect to endocrine profile
- To compare HP-hCG with rhCG with respect to safety profile

2.2 Endpoints

Primary Endpoint

- Number of oocytes retrieved

Secondary Endpoints

- Number of metaphase II oocytes (only applicable for insemination using ICSI)
- Number of fertilised (2PN) oocytes
- Fertilisation rate (the rate of fertilised oocytes to oocytes retrieved for subjects with oocytes retrieved and also the rate of fertilised oocytes to metaphase II oocytes for subjects with oocytes insemination using ICSI)
- Positive βhCG rate (positive serum βhCG test 13-15 days after transfer)
- Clinical pregnancy rate (at least one gestational sac 5-6 weeks after transfer)
- Vital pregnancy rate (at least one intrauterine gestational sac with fetal heart beat 5-6 weeks after transfer)
- Circulating concentrations of estradiol, progesterone and hCG at transfer and 13-15 days after transfer
- Frequency and intensity of adverse events
- Early and late OHSS rates
- Frequency and intensity of injection site reactions (redness, pain, itching, swelling and bruising) assessed by the subject after hCG administration

Post-trial Information

- Live birth rate and neonatal health at birth for subjects who have achieved a vital pregnancy

3 Trial Design

3.1 General Design Considerations

3.1.1 Overall Design and Control Methods

Trial Design

This is a randomised, assessor-blind, 3 parallel groups, multicentre phase III trial comparing efficacy and safety of HP-hCG (subcutaneous (SC) or intramuscular (IM) administration) and rhCG (SC) for triggering of final follicular maturation in women undergoing controlled ovarian stimulation.

Subjects who meet the criterion for triggering of final follicular maturation will be randomised in a 1:1:1 ratio to either HP hCG 5,000 IU IM, HP hCG 5,000 IU SC, or rhCG 250 µg SC.

Randomisation will be stratified by the number of follicles ≥12 mm at the end of stimulation (two strata: <10 and ≥10).

The end-of-trial assessments should take place at the subject's last scheduled trial visit or within 2 weeks after last scheduled visit in case of premature discontinuation. Subjects with a vital pregnancy at end-of-trial will be followed for post-trial information.

3.2 Determination of Sample Size

It is planned to randomise 258 subjects from 5 sites in Brazil.

The trial is dimensioned as a non-inferiority trial based on the primary endpoint, 'Number of oocytes retrieved' and a non-inferiority margin of -3.0 oocytes, see protocol section 3.5.1 for the rational of the non-inferiority margin.

Non-inferiority of HP-hCG compared with rhCG will be established by doing the comparisons in sequential order as follows:

1) HP-hCG IM vs. rhCG SC

2) HP-hCG SC vs. rhCG SC
This will only be evaluated if the first comparison establishes non-inferiority

In each step non-inferiority will be established if the respective 95% confidence interval lie entirely above -3.0 (the non-inferiority margin) for both the PP and the ITT analysis sets.

The sample size was determined for a 1:1:1 randomisation ratio based on a one-sided t-test at a 2.5% significance level. Assuming all treatment are equally effective and a standard deviation of 6.0 oocytes, a sample size of 68 subjects per treatment group would give 90% power to establish

non-inferiority for the first comparison. The number of subjects with major protocol deviations affecting the primary endpoint is assumed to be negligible. Therefore the sample size has not been adjusted to account for major protocol deviations affecting the primary endpoint.

The sample size needed to establish non-inferiority for the first comparison can be derived by solving the equation for the smallest integer n:

$$n = \frac{2s^2\left(t(2n-2)_{\alpha/2} + t(2n-2)_{\beta}\right)^2}{(\varepsilon-\delta)^2}$$

where n=sample size for one group

$t(df)_p$ is the p percentile of the t-distribution with df degrees of freedom

s=assumed standard deviation

δ=non-inferiority margin

ε=actual difference between treatments (assumed to be zero)

α/2=significance level (one-sided)

β=1-power.

The power to establish non-inferiority for both comparisons, assuming the ITT and PP analysis set to be equal in addition to the suggested sequential testing procedure, was approximated to 80% via simulation.

The assumption regarding the standard deviation underlying the sample size estimation will be evaluated by an Internal Data Monitoring Committee (DMC) in a blinded manner and the number of randomised subjects may be adjusted up to a maximum of 348 subjects. See protocol and the DMC working procedures for more details.

4 Subject Disposition

Screened subjects

All screened subjects will be accounted for. The total number of screened subjects will be summarised (n and % of total number of screened) by: Randomised, Reason for failure (Inclusion/exclusion criteria not met, Withdrawal by subject, Other) and All.

Unless otherwise specified the disposition tables will be produced overall and by stratum and will include the treatment groups and a total column.

Subject disposition

Subject disposition with respect to analysis sets (n and % of ITT): ITT Analysis Set, Full Analysis Set, PP Analysis Set and Safety Analysis Set.

Subject disposition by trial site

Subject disposition by trial site overall and by stratum (no treatment groups) (n and % of ITT with no percentage for the screened subjects): Screened, Analysis Sets (ITT, FAS, PP and Safety Analysis Set), Completed, Discontinued.

Subject completion/discontinuation - ITT

Subject completion by end-of-trial status (n and % of ITT): Completed, Discontinued (primary reason for premature discontinuation according to end-of-trial form).

Subject attendance at selected trial visits - ITT and PP

Subject attendance at selected visits (n and % of ITT or PP): Triggering visit, Oocyte retrieval visit, Transfer visit, βhCG visit, Clinical pregnancy visit and End-of-trial visit.

Subject attendance by trial procedure - ITT and PP

Subject attendance by selected trial procedures (Oocyte retrieval, Transfer, βhCG assessment, Clinical pregnancy and End-of-trial) including the detailed information of why the trial procedure was not performed.

Listings - ITT

Subject disposition with respect to analysis sets will be listed including information on trial completion. Subjects who discontinued from the trial will be listed including information on date and reason for discontinuation, the last visit attended (excluding the end-of-trial visit). Both listings will be sorted by centre, stratum and subject number.

5 Protocol Deviations

Protocol deviations will be rated as minor or major. Major protocol deviations impacting the primary endpoint and thereby affecting the conclusions will lead to exclusion of subjects from the PP analysis set. Subjects will not be excluded from the PP analysis set in case of only minor protocol deviations.

The list of major protocol deviations impacting the primary endpoint includes, but is not restricted to the following:

1) hCG administered but triggering criterion is not met
2) Wrong dose of hCG administered
3) Oocyte retrieval not performed 36h (±2h) after hCG administration

The sponsor's Medical Officer at Ferring Global Clinical R&D, the sponsor's Medical Officer in Brazil, and the project-responsible statistician at Ferring Global Biometrics will perform a blinded review of data before declaration of clean file and lock of database and rate protocol deviations as minor or major. The list of major protocol deviations will be detailed and documented in the clean file document prior to database release. Major protocol deviations impacting the primary endpoint will be tabulated and listed by subject for the ITT analysis set.

Major protocol deviations - ITT

Major protocol deviations leading to exclusion of subjects from the PP analysis set will be summarised overall (any major deviation) and by category (based on the coded terms) and produced overall and by stratum including the treatment groups and a total column (n and % of ITT).

Listings - ITT

The listing will include the planned (randomised) and actual treatment and details about the deviation.

6 Analysis Sets

Four different analysis sets are defined in the protocol: ITT, FAS, PP and the safety analysis set. The ITT and PP analysis sets are the main analysis sets used for the primary efficacy evaluation. In addition, if FAS is different from ITT, (e.g. some subjects are randomised but not exposed to IMP) the FAS may also be used to describe the primary and key secondary efficacy endpoints. The safety analysis set will be used for the safety evaluation.

6.1 Intention-To-Treat Analysis Set

The intention-to-treat (ITT) analysis set is defined as all randomised subjects. Subjects will be analysed according to planned (randomised) treatment.

6.2 Full-Analysis Set

The full analysis set (FAS) is defined as all subjects randomised and exposed to IMP. Subjects will be analysed according to planned (randomised) treatment received.

6.3 Per Protocol Analysis Set

The PP analysis set is defined as all subjects randomised and exposed to IMP except those excluded as a result of major protocol deviations as described in protocol section .9.3 Subjects will be analysed according to actual treatment received.

6.4 Safety Analysis Set

The safety analysis set is defined as all subjects randomised and exposed to IMP. Subjects will be analysed according to actual treatment received.

6.5 Additional Data Presentations

Efficacy (based on ITT) and safety (based on safety analysis set) will in some cases be reported using the following subgroups reflecting different aspects of the trial:

- Subjects with oocyte retrieved (>=1 oocytes)
- Fertilisation method (IVF, ICSI or IVF and ICSI)
- Subjects with ICSI only
- Subjects with transfer
- Number of embryos/blastocysts transferred (single transfer or double transfer)
- Quality of embryos/blastocysts transferred (good quality or not good quality)
- Transfer Day (Day 3 or Day 5)
- Subjects with positive βhCG

7 Trial Population

Categorical data will be summarised using numbers and percentages (n and % of all observed) in addition to the sum 'All'. The percentages are based on the total number of subjects with a corresponding assessment in the analysis set. Continuous data will be presented, using the number of subjects (n), mean and standard deviation, median, inter-quartile range (IQ-range), minimum and maximum.

Missing data will not be imputed.

Summary tables will include the treatment groups and a total column. The purpose of these tabulations is to characterise the treatment groups and assess the degree of similarity achieved by randomisation. Tabulations will be produced overall and by stratum for the analysis set (ITT only or both ITT and PP as specified below). In case the ITT and PP analysis sets are identical duplicate tables will not be produced.

The data used to describe the trial population will be listed for the ITT analysis set.

7.1 Demographics and Other Baseline Characteristics

Demographics by stratum - ITT and PP

Demographics captured at screening: Age (years), Age grouped (<35, 35-37, >=38, All), Race, Ethnicity and Country.

Body measurements by stratum - ITT and PP

Body measurements captured at screening: Height (m), Weight (kg), BMI (kg/m^2) and BMI grouped (<18.5, 18.5-<25.0, 25.0-<30.0,>=30.0, All).

7.2 Transvaginal Ultrasound

Transvaginal ultrasound is performed to assess the ovarian at stimulation Day 1, the endometrium at stimulation Day 1 and end of stimulation and the follicular development at stimulation Day 1, 6, subsequent stimulation days, and at end of stimulation. The data captured in the stimulation period will be presented as part of baseline characteristics.

7.2.1 Ovarian Volume

The average ovarian volume (cm^3) will be calculated using the equation for an ellipsoid as:

$$\frac{\pi}{6}\cdot\left(\frac{l_r\cdot w_r\cdot d_r}{2}+\frac{l_l\cdot w_l\cdot d_l}{2}\right)\cdot\frac{1}{1000}$$

where l_x, w_x, and d_x are the length, width and depth in mm of the right (x=r) and left (x=l) ovary respectively. The total ovarian volume is then derived as twice the average volume. The total and average ovarian volume will only be reported if observations are available for both ovaries.

Ovarian volume before stimulation by stratum - ITT and PP

Ovarian volumes stimulation Day 1: Total ovarian volume (cm^3) and Average ovarian volume (cm^3).

7.2.2 Endometrial Status

Endometrial status by stratum and visit - ITT and PP

Endometrial status by visit (Day 1 and end of stimulation): Endometrial thickness (mm), Endometrial thickness (categories) (<=6 mm, 7-9 mm, >=10 mm, All) and Endometrial echogenicity pattern (Hypoechogenic, Isoechogenic, Hyperechogenic, Not possible to evaluate, All).

7.2.3 Follicular development

Follicular development will combine the data from both the left and the right ovary and summaries will be made for stimulation Day 1, Day 6 and end of stimulation.

Follicles by stratum and visit - ITT and PP

By visit (Day 1, 6 and end of stimulation on subject level): Total number of follicles and Average follicle size (mm).

7.3 Endocrine Parameters

Values below the lower limit of quantification (LLOQ) will be included as LLOQ/2. Values above the upper limit of quantification (ULOQ) will be included as ULOQ.

Endocrine parameters by stratum - ITT and PP

The endocrine profile at baseline: Estradiol (ρmol/L), Progesterone (nmol/L) and hCG (IU/L).

7.4 Infertility History, Menstrual History and Reproductive History

Infertility history by stratum - ITT and PP

Infertility history: Primary infertility (Yes, No, All), Duration of infertility (months) and Primary reason for infertility (Unexplained infertility, Tubal infertility, Mild male factor, Moderate male factor, Severe male factor, Anovulatory Infertility WHO Group I, Anovulatory Infertility WHO Group II, Endometriosis (Stage I/II), Endometriosis (Stage III/IV), Other, All).

Menstrual history by stratum - ITT and PP

Menstrual history: Average duration of menstrual cycle (days).

Obstetric history (natural conception) - ITT and PP

Obstetric history (all parameters should be treated as categorical): Previous clinical pregnancy (Yes, No, All), Number of previous clinical pregnancies (0, 1, 2, 3, 4, All), Previous live birth (Yes, No, All), Number of previous live births (0, 1, 2, 3, All), Number of fetuses (pregnancy level) (1, 2,

All) and Outcome (fetus level) (Live birth, Still birth, Ectopic pregnancy, Miscarriage 1st trimester, Miscarriage 2nd trimester, Miscarriage 3rd trimester, Termination, All).

Previous fertility treatment by stratum - ITT and PP

Previous fertility treatment (all parameters should be treated as categorical): Previous fertility treatment (Yes, No, All), Number of previous fertility treatments (0, 1, 2, 3, 4, 5, 6, >6, All), Previous clinical pregnancy (Yes, No, All), Number of clinical pregnancies (0, 1, 2, 3, All), Previous live births (Yes, No, All), Number of previous live births (0, 1, 2, All), Number of fetuses (pregnancy level) (1, 2, All) and Outcome (fetus level) (Live birth, Still birth, Ectopic pregnancy, Miscarriage 1st trimester, Miscarriage 2nd trimester, Miscarriage 3rd trimester, Termination, All).

7.5 Physical and Gynaecological Examination

Physical and gynaecological examination will be summarised by the reported body system with responses 'Normal', 'Abnormal, NCS' and 'Abnormal, CS' (no total column for categories within body systems. These summary tables will only be produced overall (i.e. not by stratum) for the ITT analysis set.

Physical examination - ITT

Gynaecological examination - ITT

7.6 Medical History

Medical history recorded at screening visit will be coded using MedDRA. Medical history will be reported as ongoing if the diagnosis is present at the randomisation. The summaries will be made by SOC (alphabetically) and preferred term (in decreasing order of frequency based on the total column) including number of subjects with observation (n) and % of ITT. These summary tables will only be produced overall (i.e. not by stratum) for the ITT analysis set.

Incidence of ongoing medical history events by MedDRA system organ class and preferred term - ITT

The number of subjects with "Any ongoing medical history" will be included.

Incidence of past medical history events by MedDRA system organ class and preferred term - ITT

The number of subjects with "Any past medical history" will be included.

7.7 Concomitant Medication

Concomitant medication include two categories; "concomitant fertility medication / non-investigational medicinal products" and "other concomitant medication".

7.7.1 Concomitant Fertility Medication

Concomitant fertility medication provided as non-investigational medicinal products are: gonadotropin, GnRH antagonist and progesterone.

Duration of a treatment (days) is defined as the number of days from first exposure to the day of last exposure (both inclusive). If a subject misses an intermediate dose she will still be considered as being under treatment.

Gonadotropin by stratum - ITT and PP

Duration of gonadotropin treatment (days), Total gonadotropin dose (IU), Dose on stimulation Day 1 (IU/day), Dose on stimulation Day 6 (IU/day), Dose on last stimulation day (IU/day) and Average daily dose (IU/day).

GnRH antagonist by stratum - ITT and PP

Duration of GnRH antagonist treatment (days) and Total GnRH antagonist dose (mg).

Progesterone by stratum - ITT and PP

Progesterone by stratum - Subjects with positive beta-hCG - ITT and PP

Progesterone by stratum - Subjects with negative beta-hCG - ITT and PP

Duration of progesterone treatment (days) and Total progesterone dose (mg)

7.7.2 Other Medications

Concomitant medication (except for the treatments that are given as part of the protocol, such as the IMP, gonadotropin, GnRH antagonist and progesterone) will be summarised by ATC classification 1st level (alphabetically) and ATC classification 2nd level (in decreasing order of frequency based on the total column). The tables are currently only planned for the ITT analysis set and without stratum.

The medications will be tabulated separately for:

1) Prior medication; i.e. medication taken exclusively prior to treatment (i.e. with stop date before date of IMP administration);
2) Concomitant medication, i.e. medication taken during the treatment period (i.e. medication that was not stopped before date of first IMP-administration and not started after the end-of-trial visits.

If the timing of the dose of a concomitant medication cannot be established in relation to the administration of IMP, it will be considered as concomitant medication.

Concomitant medication by ATC level 1 and ATC level 2 - ITT

The number of subjects with “Any concomitant medication” will be included.

Prior medication by ATC level 1 and ATC level 2 - ITT

The number of subjects with “Any prior medication” will be included.

8 Exposure and Treatment Compliance

The exposure data will be listed for the ITT analysis set.

8.1.1 Extent of Exposure

Triggering of final follicular maturation by stratum - ITT

Triggering criteria met at last stimulation day (Yes, No, All), Triggering drug (the actual drug used HP-hCG, rhCG, Not triggered, All), Administration route (SC, IM) and dose (HP-hCG (IU), rhCG (µg)).

8.1.2 Treatment Compliance

No summary tables are planned on compliance, since non-compliance is expected to be limited due to the fact that the IMP is administered as a single injection by the trial staff (or at the attendance of the trial staff for the SC administration).

9 Efficacy

9.1 General Considerations

The efficacy endpoints will be listed for the ITT analysis set. The listings will be sorted by centre, stratum, subject number and visit.

In case the ITT and PP analysis sets are identical duplicate tables will not be produced.

If the FAS is different from the ITT analysis set, (e.g. some subjects are randomised but not exposed to IMP) tables for primary and key secondary efficacy endpoints will also be produced for the FAS.

9.1.1 Primary and Secondary Endpoints

The result of the analysis of the primary endpoint (number of oocytes retrieved) is essential for the non-inferiority claim. The number of MII oocytes (only applicable for subjects where all oocytes are intended to be inseminated using ICSI) and the number of 2PN oocytes are considered the key secondary endpoints supportive of the primary endpoint. However, non-inferiority does not need to be established for these secondary supportive endpoints. The analyses of the remaining secondary efficacy endpoints are intended to provide additional characterisation of the treatment effect.

9.1.2 Analysis and Presentation of Primary and Secondary Endpoints

Summary tables for the primary endpoint and the secondary efficacy endpoints will be presented overall and by stratum for both the ITT and the PP analysis sets.

All statistical tests will be performed using a two-sided test at a 5% significance level. Treatment differences will, where appropriate, be presented with two-sided 95% confidence intervals and p-values corresponding to the hypothesis of “equal effect” against the alternative “different effect”.

9.1.3 Multiplicity

Multiplicity needs to be addressed for the primary endpoints since it involves 2 hypotheses and each hypotheses needs to be demonstrated for 2 analysis sets. No adjustment for the multiple hypotheses is required, since there is only one primary endpoint and the evaluation is performed in a sequential manner as described in Section 9.2.1. No adjustment is needed for the 2 analysis sets since non-inferiority has to be established for both.

Concerning the superiority claim for the primary endpoint and the secondary endpoints, no formal adjustment for multiplicity will be utilised. Statistically significant results among these secondary endpoints will be interpreted cautiously.

9.1.4 Missing Data

Missing observations for the primary endpoint (number of oocytes retrieved) will be imputed as zero irrespective of the reason why data was not recorded.

Missing observations for the number of 2PN oocytes will be imputed as zero. This also includes subjects with no oocytes retrieved; these will have the number of 2PN oocytes set to zero.

For subjects where all oocytes are intended to be inseminated using ICSI, missing observations for the number of MII oocytes will be imputed as zero. For subjects who do not have any oocytes inseminated using ICSI, missing observations for the number of MII oocytes will not have a value imputed.

Missing observations for the βhCG assessment will be imputed as negative, unless a positive result is observed at the clinical pregnancy visit. For example, if the outcome of βhCG is missing but clinical pregnancy is "positive", βhCG will be imputed as "positive".

Missing observations for the clinical pregnancy assessment and vital pregnancy assessment will be imputed as "negative" irrespective of why data were not recorded.

9.2 Primary Endpoint

9.2.1 Primary Analysis

The primary endpoint 'Number of oocytes retrieved' captured 36h±2h after triggering will be analysed using an ANOVA model with treatment (HP-hCG 5.000 IU IM, HP-hCG 5.000 IU SC and rhCG 250 µg SC), stratum (<10 and >=10) and trial site (5 sites are planned) as fixed factors. The model will be fitted for the PP and the ITT analysis sets. Since this is a non-inferiority trial, the PP and the ITT analysis sets have equal importance and should lead to similar conclusions for a robust interpretation. The 2-sided 95% confidence limits for the mean treatment differences (HP-hCG IM - rhCG SC and HP-hCG SC - rhCG SC) will be calculated based on the fitted model for each analysis set.

Non-inferiority of HP-hCG compared with rhCG will be established by doing the comparisons in sequential order as follows:

1) HP-hCG IM vs. rhCG SC

2) HP-hCG SC vs. rhCG SC
   This will only be evaluated if the first comparison establishes non-inferiority

In each step non-inferiority will be established if the respective 95% confidence interval lie entirely above -3.0 (the non-inferiority margin) for both the PP and the ITT analysis sets. If the 95% confidence interval for the mean treatment difference not only lies above the non-inferiority limit but also above zero for the ITT analysis set, there is evidence of superiority in terms of statistical

significance at the two-sided 5% level. In this case, the p-value from the test for superiority will be reported.

Formally the model can be written as follows:

$OR_i = centre_{j(i)} + stratum_{k(i)} + treatment_{l(i)} + \varepsilon_i$,

where:

OR_i denotes the number of oocytes retrieved for subject i

$centre_{j(i)}$ denotes the centre for subject i

$stratum_{k(i)}$ denotes the stratification factor for subject i

$treatment_{l(i)}$ denotes the treatment for subject i

ε_i denotes an error term assumed to be normally distributed with mean zero and a common standard deviation across subjects.

The hypotheses to be tested in a in sequential order are:

H_{10}: $OR_{HP\text{-}hCG\ IM} - OR_{rhCG} \leq -3.0$

H_{1A}: $OR_{HP\text{-}hCG\ IM} - OR_{rhCG} > -3.0$

H_{20}: $OR_{HP\text{-}hCG\ SC} - OR_{rhCG} \leq -3.0$

H_{2A}: $OR_{HP\text{-}hCG\ SC} - OR_{rhCG} > -3.0$

where $OR_{HP\text{-}hCG\ IM}$, $OR_{HP\text{-}hCG\ SC}$ and OR_{rhCG} denotes the number of oocytes retrieved with HP-hCG (IM or SC) and rhCG, respectively.

9.2.2 Additional Analyses

The amount of imputed data (missing observations for the primary endpoint imputed as zero) are expected to be limited due to the trial design. Subjects with missing observations for the primary endpoint are likely to be excluded from the PP analysis set and hence the ITT analysis set can be used to assess the implication of the imputation method.

The fixed effects and the corresponding p-values will be presented for the following models:

1 the primary analysis model
2 the model that also includes the treatment-by-stratum interaction
3 the model excluding stratum
4 the model fitted to each stratum separately (subgroup analyses).

In addition, more variables may be added to the primary ANOVA model in order to adjust for baseline differences or to investigate for a potential treatment interaction.

9.2.3 Summary Tables

Oocytes retrieved by stratum - ITT/PP

Number of oocytes retrieved.

9.3 Secondary Endpoints

9.3.1 Metaphase II Oocytes and Fertilised Oocytes

An ANOVA model, similar to that used for the primary endpoint, will be fitted for the following endpoints:

- Number of metaphase II oocytes (only applicable for insemination using ICSI)
- Number of fertilised (2PN) oocytes

The model will be fitted for both the ITT and the PP analysis sets. The 2-sided 95% confidence limits for the mean treatment differences (HP-hCG IM - rhCG SC and HP-hCG SC - rhCG SC) and the corresponding p-value for no treatment difference will be presented for each comparison.

Subgroup analysis for stratum will be made by fitting the ANOVA to each stratum separately.

Summary Tables

Metaphase II oocytes(subject level) by insemination method (ICSI only) and stratum – ITT/PP

Number of metaphase II oocytes, Metaphase II oocytes to oocytes retrieved (%) and Metaphase II oocytes continued.

Fertilised oocytes (subject level) by stratum – ITT/PP

Fertilised oocytes (subject level) by stratum - Subjects with oocytes retrieved – ITT/PP

Fertilised oocytes (subject level) by insemination method – ITT/PP

Number of fertilised oocytes and Oocytes continued.

9.3.2 Fertilisation Rate

The fertilisation rate is defined both based on oocytes retrieved and MII oocytes and expressed as percentages. The fertilisation rate will be summarised using descriptive statistics.

Fertilisation rate:

- 100*(Number of fertilised oocytes/Number of oocytes retrieved) for subjects with oocytes retrieved

- 100*(Number of fertilised oocytes/Number of metaphase II oocytes) for subjects with all oocytes inseminated using ICSI

Summary Tables

Insemination Day 0 (embryo level) by stratum – ITT

Insemination Day 0 (embryo level) by insemination method – ITT

Oocyte maturity stage (Germinal vesicle (GV), Metaphase I (MI), Metaphase II (MII), Degenerated, Other and All), Destiny (Continue, Out of trial, All), Oocyte inseminated (Yes, No, All), Method of insemination (IVF; ICSI, All).

Fertilisation Day 1 (embryo level) by stratum – ITT

Fertilisation Day 1 (embryo level) by insemination method – ITT

Pronuclei at Day 1 (>2 pn, 2 pn, 1 pn, 0 pn, Damaged, All), Destiny (Continue, Out of trial, All).

Fertilisation (subject level) by stratum - Subjects with oocytes retrieved - ITT

Fertilisation (subject level) by insemination method - ITT

Fertilisation rate relative to oocytes retrieved and Fertilisation rate relative to MII oocytes.

9.3.3 Number and Quality of Embryos Day 3

The number of embryos and the number of good quality embryos (≥6 blastomeres and ≤20% fragmentation) available on day 3 will be tabulated.

Embryo assessments Day 3 (embryo level) by stratum – ITT

Embryo assessments Day 3 (embryo level) by insemination method – ITT

Embryo stage (Cleavage stage, Compacting/Compaction, Degenerated, All), Degree of fragmentation (<=20%, >20%, All), Visual sign of multi-nucleation (Yes, No, All), Good-quality embryo (Yes, No, All), Destiny (Continue, Transferred, Cryopreserved, Out of trial, All).

Embryo quality Day 3 (subject level) by stratum – ITT

Embryo quality Day 3 (subject level) by day of transfer – ITT

Embryo quality Day 3 (subject level) by insemination method – ITT

Number of embryos, At least one embryo on Day 3 (Yes, No, All), Number of good-quality embryos, At least one good-quality embryo on Day 3 (Yes, No, All), Number of embryos transferred (0, 1, 2, All).

9.3.4 Number and Quality of Blastocysts Day 5

The number of blastocysts and the number of good quality blastocysts (grade 3BB or above) available on day 5 will be tabulated.

Blastocyst assessments Day 5 (embryo level) by stratum – ITT

Blastocyst assessments Day 5 (embryo level) by insemination method – ITT

Embryo stage (Blastocysts, Morula, Degenerated, Cleavage stage, All), Blastocyst expansion and hatching status (1 to 6, as defined on the eCRF), Inner cell mass grading (A: Tightly packed, many cells, B: Loosely grouped, several cells, C: Very few cells, All), Trophectoderm grading (A: Many cells forming a cohesive epithelium, B: Few cells forming a loose epithelium, C: Very few, large cell, All), Quality of blastocyst (XAB combo), Good-quality blastocyst (Yes, No, All), Destiny (Transferred, Cryopreserved, Out of trial, All), Artificial shrinking (Yes, No, All).

Blastocyst quality Day 5 (subject level) by stratum – ITT

Blastocyst quality Day 5 (subject level) by insemination method – ITT

Number of blastocysts, At least one blastocyst on Day 5 (Yes, No, All), Number of good-quality blastocysts, At least one good-quality blastocyst on Day 5 (Yes, No, All).

Blastocyst destiny at day 5 (subject level) by stratum - ITT

Blastocyst destiny at day 5 (subject level) by insemination method – ITT

Number of blastocyst transferred (0, 1, 2, All), Blastocysts cryopreserved (0, 1, 2, >=3, All).

9.3.5 Pregnancy Rates

The following rates are defined as the proportion of subjects with the event of interest:

- Positive βhCG rate (positive serum βhCG test 13-15 days after transfer)
- Clinical pregnancy rate (at least one gestational sac 5-6 weeks after transfer)
- Vital pregnancy rate (at least one intrauterine gestational sac with fetal heart beat 5-6 weeks after transfer)

The proportions with the different denominators (all subjects, subjects with oocytes retrieved and subjects with transfer) will be calculated as crude rates (without adjustment of stratum and centre) and presented for the different denominators (ITT and PP).

In addition, the pregnancy rates will be summarised by day of transfer (Day 3 or Day 5), number of embryos/blastocysts transferred (single transfer or double transfer) and quality of embryos/blastocysts transferred (good quality or not good quality).

The pregnancy rates for each treatment (HP-hCG IM, HP-hCG SC and rhCG SC) will be presented together with the 2-sided 95% Clopper-Pearson confidence intervals based on the ITT analysis set. No formal treatment comparison are planned for the pregnancy rates since the trial is not powered for these secondary endpoints.

Summary Tables

Beta-hCG by stratum – ITT/PP

Beta-hCG by stratum - Subjects with oocytes retrieved – ITT/PP

Beta-hCG by stratum - Subjects with transfer – ITT/PP

Beta-hCG (Positive, Negative, All).

Clinical pregnancy by stratum – ITT/PP

Clinical pregnancy by stratum - Subjects with oocytes retrieved – ITT/PP

Clinical pregnancy by stratum - Subjects with transfer – ITT/PP

Clinical pregnancy (Positive, Negative, All), Intrauterine pregnancy (Yes, No, All), Ectopic pregnancy (Yes, No, All).

Vital pregnancy by stratum – ITT/PP

Vital pregnancy by stratum - Subjects with oocytes retrieved – ITT/PP

Vital pregnancy by stratum - Subjects with transfer – ITT/PP

Vital pregnancy (Positive, Negative, All), Intrauterine gestational sacs with heart beat (1, 2, 3, All) and Intrauterine fetal heart beats (1, 2, 3, All).

9.3.6 Endocrine Parameters

The endocrine parameters are measured at randomisation (end of stimulation), at the transfer visit and at the visit 13-15 days after transfer (pregnancy monitoring). Data will be evaluated based on the ITT analysis without any data imputation.

The endocrine parameters will be evaluated by fitting a separate ANCOVA model to the log(relative change) to the transfer visit and the pregnancy monitoring visit. The models will include treatment, centre, stratum and pregnancy status (positive or negative βhCG at the βhCG visit) as fixed factors and log(baseline) as continuous covariate.

Subgroup analyses for stratum and positive/negative βhCG at the βhCG visit will be made by fitting the ANCOVA to each stratum separately.

The endpoints are:

- Circulating concentrations of estradiol, progesterone and hCG at transfer and 13-15 days after transfer

The 2-sided 95% confidence limits for the mean treatment ratios (HP-hCG IM / rhCG SC and HP-hCG SC / rhCG SC) and the corresponding p-value for no treatment difference (ratio=1) will be presented for each visit and comparison, overall and by stratum.

Exploratory analysis will be made to investigate if the ovarian response (number of oocytes retrieved) impact the circulating levels the endocrine parameters.

Summary Statistics and Graphs

The actual measurements and the change from baseline will be summarised overall and by stratum. In addition Box plots will be used to display the actual levels and the change from baseline over time.

9.4 Other Assessments

9.4.1 Insemination

Insemination (subject level) by stratum – ITT

Method of insemination (n, %) (None, IVF, ICSI, IVF and ICSI and All).

9.4.2 Embryo or Blastocyst Transfer Procedure

Embryo or blastocyst transfer procedure (subject level) by stratum – ITT

Day of transfer (3, 5), Number of embryo/blastocyst transferred (0, 1, 2 or All), Method of transfer (Ultrasound guided, Clinical touch or All), Difficult transfer (Yes, No, All) and Any eventuality (Yes, No, All).

10 Safety

10.1 General Considerations

The safety endpoints will be listed for the safety analysis set. The listings will be sorted by centre, stratum, subject number, visit or onset as applicable.

Safety parameters will be evaluated for the safety analysis set, i.e. data will be reported according to actual treatment received. The safety endpoints are:

- Frequency and intensity of AEs
- Early and late OHSS rates
- Frequency and intensity of injection site reactions (redness, pain, itching, swelling and bruising) assessed by the subject after hCG administration

Post-trial Information

- Live birth rate and neonatal health at birth for subjects who have achieved a vital pregnancy

In addition the following safety parameters are collected:

- Physical examination at screening and end-of-trial
- Gynaecological examination at screening and end-of-trial
- Body weight at screening and end-of-trial

Blood collection is performed for potential analysis of anti hCG antibodies at end-of-stimulation (baseline), transfer and 13-15 days after transfer.

OHSS and injection site reactions are both recorded on specific eCRFs designed to include the relevant details for each event. OHSS is also reported on the standard AE form whereas injection site reactions are only reported on the standard AE form if they require active management.

The AE/event tabulations will be prepared overall, i.e. not by stratum. The summaries will include the number of subjects reporting an AE/event, the percentage of subjects (%) with an AE/event, and the total number of AEs/events (E).

10.2 Adverse Events

AEs will be recorded from signed informed consent until the end-of-trial visit. In addition, serious adverse events (SAEs) reported as part of the pregnancy outcome and neonatal health will be recorded during the post-trial follow-up period.

If a subject suffers from the same AE more than once and the subject recovers in between the events, the AEs will be reported separately.

AEs are grouped according to exposure as follows:

- Pre-treatment AE, i.e. any AE occurring after signed informed consent and before administration of the randomised treatment, or a pre-existing medical condition that worsens in intensity after signed informed consent but before administration of randomised treatment.
- Treatment-emergent AE, i.e. any AE occurring after administration of randomised treatment and before the end-of-trial visit, or a pre-treatment adverse event or pre-existing medical condition that worsens in intensity after administration of randomised treatment and before the end-of-trial visit.

If the timing of an AE cannot be established in relation to administration of randomised treatment, it will be considered as a treatment emergent AE.

Pre-treatment AEs will be listed only.

AEs are classified according to the MedDRA. The system organ class (SOC) and the preferred term (PT) will be used to categorise events. The version of MedDRA will be documented.

Missing values will be treated as missing, except for causality, intensity, seriousness, and outcome of adverse events where a worst-case approach will be used:

- if causality is missing, the AE will be regarded as related to the randomised treatment with a reasonable possibility
- if the intensity of an AE is missing, the AE will be regarded as severe
- if seriousness is missing, the AE will be regarded as serious
- if outcome is missing and no date of outcome is present, the outcome is regarded as ‘Not yet recovered’

Written narratives will be issued for all serious AEs (including deaths) and AEs leading to discontinuation.

10.2.1 Overall Summary of Treatment-Emergent Adverse Events

A summary table of treatment-emergent AEs will include the following categories:

- Adverse events
- Severe adverse events
- Adverse drug reactions (adverse events with causality to randomised treatment judged as a reasonable possibility)

- Adverse events leading to discontinuation
- Serious adverse events
- Adverse events leading to death

10.2.2 Incidence of Treatment-emergent Adverse Events

Treatment-emergent adverse events will be summarised overall ("Any adverse event"), by SOC (alphabetically) and PT (in decreasing frequency of occurrence) for:

- Adverse events
- Adverse drug reactions
- Adverse events leading to death
- Adverse events by intensity (mild, moderate or severe separately)
- Adverse drug reactions by intensity (mild, moderate or severe separately)
- Serious adverse events
- Adverse events leading to discontinuation
- Pregnancy loss

Treatment-emergent adverse events will be summarised overall ("Any adverse event") and by PT (in decreasing frequency of occurrence) for:

- Adverse events with an incidence of at least 5% (in any treatment group)
- Non-serious adverse events with and incidence of at least 5% (in any treatment group)

Treatment-emergent adverse events will be summarised overall ("Any adverse event"), by SOC (alphabetically) and by High Level Group Term (HLGT) (in decreasing frequency of occurrence) for:

- Adverse events

Supporting data listings sorted by SOC (alphabetically), PT (in decreasing frequency of occurrence), centre, stratum, treatment, subjects and onset will be provided for:

- All adverse events
- All adverse events sorted by PT (not SOC)
- Adverse drug reactions
- Adverse events leading to death
- Severe adverse events
- Serious adverse events

- Adverse events leading to discontinuation
- Pregnancy loss

10.3 Treatment-emergent OHSS

Early OHSS is defined as OHSS with onset ≤9 days after triggering of final follicular maturation and late OHSS as OHSS with onset >9 days after triggering. OHSS symptoms will be classified using Golan's system (see protocol sec 8.3.1)

Treatment-emergent OHSS by grade

A summary table (n, %) of treatment-emergent OHSS by severity classification and grade (Any Grade, Mild OHSS: Grade 1 and 2, Moderate OHSS: Grade 3, Severe OHSS: Grade 4 and 5 and Moderate/Severe) will be prepared for:

- OHSS
- Early OHSS
- Late OHSS

10.4 Early Pregnancy Loss

An early pregnancy loss is defined by a positive βhCG testnot followed by a vital pregnancy. Number of subjects with early pregnancy loss will be summarised (n and % of patients with a positive βhCG test).

10.5 Injection-site Reactions

Overall summary of injection-site reactions

An overall summary table of subjects with injection-site reaction (n, %) will be prepared for the following categories:

- At least one injection-site reaction
- At least one moderate/sever injection-site reaction

Injection-site reactions by symptom, time and severity

A summary table of "injection-site reaction by symptom, time and severity (n, %, E) will be prepared by symptom (redness, pain, itching, swelling and bruising), by time (All time points, Immediately after, 30 minutes, 24 hours) and intensity (None, Mild, Moderate, Severe and Moderate/severe).

10.6 Physical and Gynaecological Examinations

Shift tables (baseline versus end-of-trial visit) will be made for physical examination and gynaecological examinations per body system/test. The tables will include subjects with non-missing data at both visits for a body system/test.

Physical examination: change in evaluation at end-of-trial visit

Gynaecological examination: change in evaluation at end-of-trial visit

A summary table per body system/test, baseline versus end-of-trial visit with N_1, n and % (N_1= number of subjects with non-missing value in available baseline category and at least one non-missing value recorded at end-of-trial visit, n=number of subjects with non-missing value in each category for baseline and end-of-trial visit, % = percentage of subjects).

All data for the subjects with one or more abnormal findings (at baseline or end-of-trial for any of the body systems/tests) will be listed by body system/test and time point.

10.7 Body Weight

Body weight is measured at screening and end-of-trial. The actual measurements and the change from baseline will be summarised.

10.8 Anti-drug Antibodies

Summaries will be made as applicable if the blood samples are analysed for anti hCG antibodies.

11 Interim Analyses

No interim analysis intended to compare treatment groups with respect to efficacy or safety is planned.

The assumption underlying the sample size estimation (assuming a standard deviation of 6.0 oocytes) will be evaluated by the internal DMC. The sample size reassessment will be done without breaking the blind and without inflating the type I error of the trial[1], in line with the current regulatory guidelines[2, 3]. The details are explained in the working procedures for the internal DMC.

12 Deviations from Protocol Analysis

There are no changes to the analyses outlined in the protocol.

13 Tables, Listings and Figures

The planned summary tables, figures and listings (TLF) for the end-of-text are included in this SAP. No separate document will be prepared with shells.

14 References

[1] Kieser M, Friede T. Simple procedures for blinded sample size adjustemtn that do not affect the type error rate.. Statist. Med. 2003; 22:3571-3581.

[2] U.S. Department of Health and Human Services Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Draft guidance February 2010, Adaptive Design Clinical Trials for Drugs and Biologics.

[3] Committee for Medicinal Products for Human Use (CHMP), Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. Adopted by CHMP 18 October 2007, CHMP/EWP/2459/02.