FERRING
PHARMACEUTICALS

Statistical Analysis Plan

Title of trial: A Prospective, Multi-Center, Non-Comparative Trial of the Clinical Safety of the Progesterone Vaginal Ring in Women Undergoing Assisted Reproductive Technology (ART) Procedures
Sponsor trial code: 000293
Date: 09 Aug 2019

STATISTICAL ANALYSIS PLAN

A Prospective, Multi-Center, Non-Comparative Trial of the Clinical Safety of the Progesterone Vaginal Ring in Women Undergoing Assisted Reproductive Technology (ART) Procedures

000293

Investigational Product:	Progesterone vaginal ring (PVR)
Indication:	Progesterone supplementation in women undergoing ART
Phase:	3b
Author:	[redacted]
Date of issue:	09 Aug 2019
Version:	4.0

Change log

Version No.	Effective Date	Reason for the Change / Revision	Supersedes
1.0	May 11, 2018	NA	None
2.0	Oct 08, 2018	• To make the SAP consistent with the protocol amendment. • More detailed analysis of embryo quality has been added. • More details have been added to the analysis of bleeding logs. • Typographical and formatting errors have been corrected.	1.0
3.0	Feb 26, 2019	• Responded to FDA feedback on statistical design by adjusting confidence intervals and assuring that endpoints align with the DR-PGN-302 study • Updated the document author to reflect the new statistician • Corrected formatting issues, typos, text relating to which drug was used for hCG triggering, and "Other Endpoint(s)" to "Other Assessment(s)" • Added tabulation of vaginal hemorrhage (was in protocol, but not SAP) • Clarified the use of Preferred Terms, how baseline is calculated for laboratory and vital signs variables, how shift tables are to be calculated, and the tabulation of bleeding log data for non-completers.	2.0

Change log

Version No.	Effective Date	Reason for the Change / Revision	Supersedes
4.0	Aug 09, 2019	• Responded to FDA feedback on statistical design by moving the frequentist sensitivity analysis to be the primary analysis and the Bayesian primary analysis to be the sensitivity analysis. • Updated the sample size justification to reflect the frequentist primary analysis. • Updated primary endpoint text throughout. • Updated Appendix 2 to reflect the revised sample size justification.	

TABLE OF CONTENTS

List of Tables

1 Introduction

This document describes the planned statistical analyses for trial 000293 based on the protocol Version 6.0 dated 9 August 2019.

1.1 Definitions/ Abbreviations

1.1.1 Definition of Terms

Terms	Definitions
Screened	Subject who enters the screening phase
ES cohort	The ES cohort consists of all subjects who were deemed eligible for trial participation based on inclusion and exclusion criteria at screening and treated with oral contraceptives, leuprolide acetate, and MENOPUR
mITT cohort	The mITT cohort consists of all subjects who had successful oocyte retrieval, received at least 1 dose of IMP, and had completed fresh embryo transfer.
Safety cohort	The safety cohort comprises all subjects treated with IMP.
Screening Phase	From the date of signing the informed consent through the day before the start of the oral contraceptive treatment.
Suppression Phase	From the start of oral contraceptive treatment until the day before the start of the stimulation phase (start of MENOPUR treatment).
Stimulation Phase	From the date of the first dose of MENOPUR through the day before the start of IMP.
Treatment Phase	From the date of the first dose of IMP (PVR) through the end of the trial.
Study Phase	From the date of the first dose of oral contraceptive through the end of the trial.

1.1.2 Abbreviations

Abbreviations	Meaning of abbreviations in document
AE	adverse event
ART	Assisted Reproductive Technology
ATC	Anatomical Therapeutic Chemical
eCRF	electronic case report form
EOT	End-of-trial
ES	eligible subjects
E2	Estradiol
FSH	follicle-stimulating hormone
hCG	human chorionic gonadotropin
IMP	Investigational Medicinal Product
IVF	in vitro fertilization
mITT	modified intention-to-treat
NIMP	Non-Investigational Medicinal Product
P	procedural visit
PVR	progesterone vaginal ring
SAE	serious adverse event
SD	Standard deviation
TVUS	transvaginal ultrasound

2 Trial Objectives and Endpoints

In this safety trial, the assessment of the trial will be on the totality of safety information collected. Hence, the purpose of categorizing the endpoints as primary and secondary is for prioritising the endpoints of interest rather than tying the trial's success to the outcome of the primary objective.

2.1 Objectives

Primary Objective

- To estimate the cumulative rate of any spontaneous abortion, including spontaneous clinically recognized pregnancy loss and blighted ovum during the trial (up to Week 12 following oocyte retrieval), in subjects treated with progesterone vaginal ring (PVR) following fresh embryo transfer.

Secondary Objectives

- To describe the cumulative rate of spontaneous abortions determined at 6 and 10 weeks post-oocyte retrieval, including spontaneous clinically recognized pregnancy loss and blighted ovum, in all subjects treated with PVR.
- To describe the cumulative rate of biochemical abortions determined at 6 and 10 weeks post-oocyte retrieval in all subjects treated with PVR.
- To describe the rate of ectopic and heterotopic pregnancy in all subjects treated with PVR following oocyte retrieval.
- To describe the safety of PVR through the collection of clinical laboratory tests and vital signs in all subjects treated with PVR.
- To assess the safety and tolerability of PVR in all subjects treated with PVR.
- To determine the positive β-hCG rate of subjects treated with PVR following oocyte retrieval.
- To determine the clinical pregnancy rate of subjects treated with PVR following oocyte retrieval.

2.2 Endpoints

Primary Endpoint

- Cumulative rate of any spontaneous abortion occurring on or before 12 weeks following oocyte retrieval in all subjects treated with PVR and undergoing fresh embryo transfer. *Note:* spontaneous abortion is defined as two positive β-hCG tests occurring at least two days apart on or after 2 weeks post-oocyte retrieval, but followed by observation of any empty intrauterine gestational sac (blighted ovum), intrauterine gestation without a fetal heartbeat, or absence of viable fetuses, as documented by transvaginal ultrasound (TVUS).

Secondary Endpoints

- Cumulative rate of spontaneous abortions determined at 6 and 10 weeks post-oocyte retrieval in all subjects treated with PVR.
- Cumulative rate of biochemical abortions determined at 6 and 10 weeks post-oocyte retrieval in all subjects treated with PVR. Biochemical abortion is defined as a positive β-hCG test at 2 weeks and 2 weeks + 3-4 days post-oocyte retrieval, but followed by no observed gestational sac on a later TVUS, or followed by a negative β-hCG test.
- Rate of ectopic and heterotopic pregnancies in all subjects treated with PVR following oocyte retrieval.
- Rate of abnormal findings in clinical laboratory tests and vital signs for all subjects treated with PVR.
- Frequency, intensity/grade, seriousness, and relatedness of adverse events (AEs) for all subjects treated with PVR.
- Frequency, intensity/grade of vaginal bleeding/spotting, vaginal hemorrhage, pain, vaginal infection, and vaginal irritation for all subjects treated with PVR.
- Frequency, intensity/grade, seriousness, and relatedness of AEs associated with vaginal and cervical abrasions and lesions and with vaginal adhesions for all subjects treated with PVR.
- Frequency and reason for PVR discontinuation.
- Positive β-hCG rate (positive serum β-hCG test) at 2 weeks and 2 weeks + 3-4 days post-oocyte retrieval in all subjects treated with PVR.
- Clinical pregnancy rate (TVUS showing at least 1 intrauterine gestational sac with fetal heartbeat) at 6 and 10 weeks post-oocyte retrieval in all subjects treated with PVR.

3 Trial design

This is an open-label, single-arm, safety trial of PVR for luteal phase support in women undergoing IVF with fresh oocytes. Women between the ages of 18 and 34 years with tubal, idiopathic, male factor, ovulatory dysfunction, or endometriosis-linked infertility, who agree to be considered for inclusion in the trial, will be invited to be seen at the trial site

3.1 General Design Considerations

Once informed consent is obtained, eligible subjects will be started on an ovarian down-regulation / suppression protocol utilizing combined oral contraceptives supplied by the site for ≥14 days to ≤21 days (based upon the site standard of care), with leuprolide acetate at a dose of 0.1 mL (500 µg)/day beginning 4 days prior to the last birth control pill and for ≥ 10 days to ≤20 days; ovarian suppression is to begin in the cycle immediately prior to the ovarian stimulation cycle. After suppression, an ovarian stimulation protocol will begin on the second or third day after the start of menses with a reduction in leuprolide acetate dose to 0.05 mL (250 µg/day) followed by an individually determined ovarian stimulation protocol with highly purified, human menopausal gonadotropin (HP-hMG; MENOPUR) at an initial dose of 225 IU/day for 5 days according to label. Based on clinical monitoring, subsequent dosing should be adjusted according to individual subject response. Dose adjustments should not be made more frequently than once every two days and should not exceed 150 IU per adjustment. The maximum daily dose of MENOPUR should not exceed 450 IU and the minimum daily dose should not be lower than 75 IU. MENOPUR dosing should not continue for >20 days and coasting is not allowed. Subjects will be monitored to determine when to trigger ovulation with hCG. A TVUS will be performed to assess follicle size and serum estradiol levels will be determined. A subject must have at least 2 follicles ≥17 mm (mean of 2 dimensions) to receive the hCG trigger. Subjects who do not reach this threshold after 20 days of stimulation are to be discontinued from the trial. If the estradiol level is ≥5000 pg/mL, hCG should not be administered and the subject should be discontinued from the trial.

Oocyte retrieval will occur approximately 35-37 hours after hCG administration. On the day after oocyte retrieval, the subject will begin treatment with the investigational medicinal product (IMP), PVR. Following instruction on PVR insertion and removal, the subject will insert the first PVR (PVR1) at the trial site The PVR will remain in place a minimum of 23 hours/day. The subject will remove PVR1 before embryo transfer and reinsert it after the transfer. Fresh embryo transfer will occur 5 days after oocyte retrieval per the trial site's protocol. The number of embryos to transfer must be guided by the 2017 American Society for Reproductive Medicine (ASRM) and the Society for Assisted Reproductive Technology (SART) guidelines (SART/ASRM 2017): women with the expectation of one or more high-quality embryo(s) available for cryopreservation, and women with previous live birth after an IVF cycle, will have transfer of a single blastocyst. All other women will have transfer of no more than two blastocysts. Additional PVRs will be distributed for weekly insertion, with the second PVR (PVR2) to be inserted at home 7 days following the initial insertion of the first PVR (PVR1), and the third PVR (PVR3) to be inserted either at home or at the site 7 days following the insertion of PVR2.

Two weeks after oocyte retrieval, a blood draw will be performed to measure serum levels of progesterone and β-hCG. Subjects with a β-hCG level <5 mIU/mL will be discontinued from the trial. Those with a β-hCG level ≥5 mIU/mL will insert their next PVR and continue in the trial for up to a total of 10 weeks (up to a total of 10 PVRs). Additional serum pregnancy tests will be performed at 2 weeks + 3-4 days and at 3 weeks. Subjects who are no longer pregnant will be discontinued from the trial. Subjects with positive β-hCG results at 3 weeks will insert PVR4 and receive PVR5 and PVR6 to be inserted weekly.

A TVUS will be performed to document the presence of an intrauterine gestational sac 4 weeks + 3-4 days after oocyte retrieval. If it is determined that the subject has an ectopic pregnancy, the subject will be treated according to the site's standard protocol and will be withdrawn from the trial. Six weeks after oocyte retrieval, a TVUS will be performed to determine the presence and number of gestational sacs with/without fetal heartbeat. Pregnant subjects will insert PVR7 and receive additional PVRs to be inserted weekly (PVR8, PVR9, and PVR10). Ten weeks after oocyte retrieval (if the subject is pregnant), an ultrasound (either transvaginal or abdominal) will be performed to determine the number of gestational sacs present with and without fetal heartbeat and the estimated gestational age.

Safety will be monitored throughout the trial. Standardized criteria will be applied to AEs of special interest (Protocol Section 8.3) to establish grade (intensity). Patient-reported vaginal pain and/or irritation should be recorded as an AE if considered clinically significant by the physician. The physician can perform an unscheduled pelvic examination at his or her discretion based upon these AEs at any time. A scheduled pelvic examination will be performed at screening, on day 1 of PVR treatment, and 4 weeks + 3-4 days, 6, 10, and 12 weeks post-oocyte retrieval (Visits 4, 7P, 8, 9 and 10) to document the presence and grade (intensity) of any pain, irritation, abrasions, or lesions on the cervix or vagina and the presence and grade (intensity) of vaginal adhesions using standardized criteria. If any lesions or abrasions are found on the cervix or vagina or vaginal adhesions are noted, another examination will be performed 2-4 days later. Subjects will be followed until the lesions/abrasions resolve or until the final assessment at the End-of-trial visit, whichever comes first. Any clinically significant finding will be reported as an AE.

Post-trial procedures will be performed 2 weeks after the last exposure to investigational medicinal product (IMP) (12 weeks after oocyte retrieval).

3.2 Trial Design Diagram

Table 1 Trial Flow Chart

Screening	Suppression	Stimulation	Oocyte Retrieval	Treatment Period	Post Treatment Period
Consent and Screening	Down-regulation	Visits 1 and 2	Visit 3	Visits 4-9: Day after oocyte retrieval and ongoing pregnancy (~10 weeks) PVR1 through PVR10	Visit 10: End-of-trial visit (2 weeks after last exposure to Investigational Medicinal Product)

3.3 Determination of Sample Size

In this safety trial, the primary assessment of interest is on the number of subjects that experience a spontaneous abortion. Because this is a single-arm trial, all subjects that enter the mITT analysis population will have been treated with the investigational product, PVR. The primary analysis involves the construction of a two-sided 95% exact confidence interval surrounding the proportion of subjects in the mITT analysis population that experience a spontaneous abortion. The upper bound of said confidence interval is then compared with the threshold of a 15% spontaneous abortion rate.

Statistical simulations have been performed to understand the operating characteristics under varying assumptions for both the sample size of the mITT analysis population and the "true" proportion of spontaneous abortions. For each simulation repetition, a sample of subjects was simulated using the binomial distribution and a given pair of sample size and "true" proportion of spontaneous abortions assumptions. A two-sided 95% exact confidence interval was then created based on the simulated data and the upper bound of this confidence interval was compared to the threshold of 0.15 (corresponding to a spontaneous abortion rate of 15%). If the upper bound of the two-sided 95% exact confidence interval was less than 0.15, the simulation repetition was declared a "success". If, however, the upper bound of the two-sided 95% exact confidence interval was greater than or equal to 0.15, the simulation repetition was declared a "failure". This process was repeated 10,000 times for each combination of assumed sample size and "true" proportion of spontaneous abortions. The proportion of simulation repetition successes to the total number of simulation repetitions then described the power of the proposed trial given the pair of assumptions.

The sample size necessary to achieve 80% power to exclude the possibility of a 15% spontaneous abortion rate after the use of PVR is described in Table 2. If the "true" spontaneous abortion rate is 8%, 9%, or 10%, the sample size required to achieve 80% power to exclude the possibility of a 15% spontaneous abortion rate after the use of PVR is 180, 240, or 365, respectively.

Table 2 Power Achieved over Varying Sample Sizes and Assumed Spontaneous Abortion Rates

Required Sample Size (mITT Analysis Population)	Assumed "True" Spontaneous Abortion Rate		
	8%	9%	10%
180	**80.3%**	64.8%	46.1%
240	93.1%	**81.8%**	63.2%
365	98.9%	94.2%	**80.9%**

Taking the middle of this range, a trial with 240 subjects in the mITT analysis population is considered as being adequately powered under a relevant assumption in the target population. Full simulation code is provided in Appendix 2.

4 Subject Disposition

The number of subjects that discontinue the trial during the stimulation phase and after start of IMP treatment (treatment phase) will be summarized by reason for discontinuation. The subjects screened and not treated with both the IMP and the non-investigational medicinal product (NIMP), MENOPUR, will be presented in a separate data listing.

5 Protocol Deviations

Major protocol deviations are a subset of the important protocol deviations that may significantly impact the primary endpoint or efficacy parameters of the trial. Major protocol deviations include the following criteria:

- significant deviation in the treatment regimens of IMP and NIMP,
- use of prohibited medication, and
- non-compliance with the blastocyst transfer policy

All protocol deviations will be listed however, only major protocol deviations will be summarized in a table.

6 Analysis sets

6.1 Modified Intention-To-Treat (mITT) Cohort

The mITT cohort consists of all subjects who had successful oocyte retrieval, received at least 1 dose of IMP, and had completed fresh embryo transfer.

6.2 Safety Cohort

The safety cohort comprises all subjects treated with IMP.

6.3 Eligible Subjects (ES) Cohort

The ES cohort consists of all subjects who were deemed eligible for trial participation based on inclusion and exclusion criteria at screening and treated with oral contraceptives, leuprolide acetate, and MENOPUR

7 Trial population

7.1 Demographics and Other Baseline Characteristics

Descriptive statistics for demographics and other baseline characteristics will be presented for the subjects in the mITT, safety, and ES cohorts.

Categorical data will be summarised using numbers and percentages. The percentages are based on the total number of subjects with a corresponding assessment. Continuous data will be presented, for example, using the number of subjects (N), mean and standard deviation, median, minimum and maximum. All baseline characteristics will be listed. Associated demographic listings will also be produced.

7.1.1 Demographics

Descriptive statistics of screening demographics variables (e.g., age, race and ethnic origin, height, weight and BMI) will be summarized.

7.1.2 Endocrine Profile

The screening AMH, FSH and E2 assessments will be summarized for the mITT and ES cohorts. Furthermore, the AMH level will be summarized by categories < 5 ng/mL and ≥ 5 ng/mL. Values below the lower limit of quantification (LLOQ) will be included as LLOQ/2.

For stimulation day 1, E2 will be assessed by the sites. This assessment will be summarized. Values below the lower limit of quantification (LLOQ) will be included as LLOQ/2.

7.1.3 Gynaecological History and Infertility Diagnosis

Gynaecological history, infertility diagnosis and menstrual history collected at screening will be summarized for the mITT and ES cohorts.

7.1.4 Smoking and Alcohol Habits

Smoking and Alcohol consumption history and current habits at screening will be summarised for the mITT and ES cohorts.

7.1.5 Transvaginal Ultrasound

Visibility of the Ovaries

The visualization of the left and right ovaries (No/Yes) will be summarized.

Antral Follicles

For summary of antral follicles please see stimulation day 1 analysis in Section 9.5.1

7.2 Medical History

Medical history recorded at screening visit will be coded using MedDRA version 21 or later.

7.3 Prior and Concomitant Medication

Prior and concomitant medication will be summarised by ATC classification 1st level (alphabetically), ATC classification 2nd level (in decreasing order of frequency) and treatment group. These medications will be tabulated separately for:

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

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. Prior-medication will be summarized for the mITT and ES cohorts while concomitant medication will be summarized for the mITT cohort.

7.4 Physical Examination

Subjects with abnormalities at any screening, baseline, or post-baseline visit will be listed with all physical examination evaluations.

8 Exposure and Treatment Compliance

8.1 Extent of Exposure

All NIMP summary will be based on the safety and ES cohorts. In contrast, the IMP analysis will be based on the safety and mITT cohorts.

8.1.1 PVR

The total number of PVRs used and the duration of treatment will be summarized. Furthermore, the drug compliance rate will also be summarized. Compliance is defined as follows:

(number of PVRs used) / (number of PVRs expected to be used).

The number of PVRs expected to be used is defined to be the number of weeks from oocyte retrieval to either confirmation of no pregnancy via β-hCG test or TVUS assessment or in the case of a subject where there is no such confirmation, the end of trial visit.

The number of times a subject was unable to keep a PVR in for at least 23 hours and the number of times a subject recorded a damaged PVR will be summarized. A detailed associated listing will be produced.

8.1.2 Oral Contraceptives

In this trial, all subjects will be prescribed the same contraceptives. The total dose and duration of treatment will be summarized. Associated listing will also be generated.

8.1.3 Leuprolide Acetate

For Leuprolide, a summary of the total dose and duration of treatment will be generated. Furthermore, the start and end date of Leuprolide Acetate from last day of oral contraceptive use will be summarized. Associated listing will be generated.

8.1.4 MENOPUR Administration

Exposure to MENOPUR will be summarized by the total dose administered and duration of treatment (days). The maximum MENOPUR dose, the dose on stimulation day 1, the final dose, and the number of dose adjustments will also be summarized. Associated listing will be generated.

8.1.5 hCG

Intramuscular injection of 10,000 IU hCG (NOVAREL) will be administered to induce final follicular maturation. A subject must have at least 2 follicles ≥17 mm to receive the hCG trigger. The following will be summarized: number of subjects who met the hCG criteria, number of subjects that had hCG administered, number of subjects who met the hCG criteria and had hCG administered, and the number of days from when the hCG criteria was met (the date of final stimulation where at least 2 follicles ≥17 mm is observed) to when the hCG is administered. Associated listing will be generated.

9 Efficacy

9.1 General Considerations

All confidence intervals will be two-sided 95% confidence intervals.

9.2 Primary Endpoint

The primary endpoint, spontaneous abortion, is defined as two positive β-hCG tests occurring at least two days apart on or after 2 weeks post-oocyte retrieval, but followed by observation of any empty intrauterine gestational sac (blighted ovum), intrauterine gestation without a fetal heartbeat, or absence of viable fetuses, as documented by transvaginal ultrasound (TVUS).

The derivation of this endpoint will be as follows:

- Any subject that has two positive β-hCG tests at least two days apart, and
- meeting at least one of the following conditions at visit 7P (4 weeks and 3-4 days post oocyte retrieval) or at a later visit:
  1. the TVUS assessment at a later visit shows that there is no presence of gestational sac,
  2. there is a presence of empty gestational sac,
  3. there is a gestational sac that has a fetus without a heartbeat, or
  4. any AE of pregnancy loss that meets the definition of spontaneous abortion.

In order to clearly identify AEs that are spontaneous abortions, listing of all AEs related to pregnancy loss, will be generated and provided to the medical team together with the reported and coded terms. Appendix 3 lists the preferred terms associated with pregnancy loss that will be used for generating this listing. This listing will be provided periodically throughout the trial. The medical team will use these listings to identify AEs that meet the spontaneous abortion definition. The identification process will be completed prior to data base lock. Since spontaneous abortion can occur outside of scheduled visits, for accurate accounting of all spontaneous abortions, the AE eCRF page needs to be included.

9.3 Primary Variable Analysis

The primary endpoint will be analysed for the mITT cohort.

9.3.1 Analysis of the Primary Endpoint

The number and rate of any spontaneous abortions (occurring on or before 12 weeks following oocyte retrieval) for subjects in the mITT cohort will be summarized and the associated two-sided exact 95% confidence interval will be generated. Upon generating the two-sided exact 95% confidence interval, the upper bound of the interval will be assessed for whether it excludes or fails to exclude a spontaneous abortion rate of 15%.

9.3.2 Hierarchical Modelling Details for the Sensitivity Analyses

Define n_1 as the number of subjects in the current trial and Y_{1i} as the indicator of spontaneous abortion for the i^{th} subject in the current trial where $Y_1 \sim Bin(n_1, \pi_1)$. Let π_1 be the true rate of spontaneous abortions in the current trial.
The previous trial (Study DR-PGN-302) reported a spontaneous rate of abortion of 55/549. This prior information about abortion rate is incorporated into the modelling framework through a hierarchical model. Dynamic borrowing between this trial's results and the current trial's data allows borrowing to occur to the extent indicated by these two data sources. More borrowing occurs when the two rates are similar while less borrowing occurs when they differ. This "dynamic" borrowing property is distinct from other approaches which use a fixed informative prior or apriori assume an amount of borrowing between these two data sources. Let $Y_0 \sim Bin(n_0=549, \pi_0)$ where π_0 is the true rate of spontaneous abortions from the previous trial.

The parameter is transformed to the logit scale for modelling purposes. Let $\theta_1 = \log(\pi_1/(1-\pi_1))$ and $\theta_0 = \log(\pi_0/(1-\pi_0))$. The hierarchical model assumes that θ_0 and θ_1 have an across studies distribution:

$$\theta_1 , \theta_0 \sim N(\mu, \tau^2)$$

The across trial mean μ and variance τ^2 are unknown and hence have a prior distribution which is combined with the data to produce estimates of μ and τ^2:

$$\mu \sim N(-2.2, 2.25)$$

$$\tau^2 \sim IG(0.125, 0.005)$$

The mean parameter μ assumes a relatively vague prior on the original scale and centered at 10%. The variance component τ^2 controls the degree of borrowing among studies. Small values of τ^2 result in a greater degree of borrowing while large values of τ^2 correspond to less borrowing. The parameter τ^2 is estimated using the data, so the observed between studies variation is a key component of the model behavior. The prior specification on τ^2 corresponds to an inverse-gamma distribution with mean 0.2 and weight 0.25.

At the end of this trial, the Bayesian hierarchical model specified above and the spontaneous abortion rate in the current trial together with the spontaneous abortion rate in the previous trial (i.e. 55 spontaneous abortion among 549 subjects) will be used to estimate the posterior distribution of the spontaneous abortion rates in the current trial. This posterior distribution will be used to generate a 95% credible interval for spontaneous abortion rate in the current trial. Since this posterior distribution is intractable, MCMC method via Openbugs in R (e.g., R2OpenBUGS) will be used to generate samples from the posterior distribution and these samples will be used to generate the 95% credible interval. Finally, whether or not the 95% credible interval lies entirely below 15% will be assessed.

9.4 Secondary Endpoints

9.4.1 Spontaneous Abortion within 6 and 10 Weeks of Oocyte Retrieval

The frequency and proportion of subjects that have any spontaneous abortion during the trial within 6 and 10 weeks (i.e., up to 42 and 70 days) post-oocyte retrieval will be derived for subjects in the mITT and safety cohorts for subjects that had oocyte(s) retrieved. Associated 95% confidence intervals will also be presented.

9.4.2 Biochemical Abortions within 6 and 10 Weeks of Oocyte Retrieval

Biochemical abortion is defined as a positive β-hCG test (at 2 weeks and 2 weeks + 3-4 days post-oocyte retrieval) but followed by no observed gestational sac on a later TVUS, or followed by a negative β-hCG test. Additionally, a biochemical abortion can be defined by an AE that meets the definition of biochemical abortion (see Appendix 3). To allow for assessment that might occur outside the visit window, a subject will be considered to be biochemically pregnant if there are two positive assessments at least two days apart on or after 2 weeks post-oocyte retrieval.

The frequency and proportion of subjects that have biochemical abortion within 6 and 10 weeks (i.e., up to 42 and 70 days) post-oocyte retrieval will be derived for subjects in the mITT and safety cohorts for subjects that had oocyte(s) retrieved. Associated 95% confidence intervals will also be presented.

9.4.3 Ectopic and Heterotopic Pregnancies

At time of TVUS examination, ectopic and heterotopic pregnancies may be detected and recorded in the TVUS eCRF. These events may also occur outside such assessment visits and may be recorded in the AE CRF. The preferred terms associated with ectopic pregnancy are presented in Appendix 3 and will be used to select and check for such terms. The frequency and proportion of subjects with ectopic and heterotopic pregnancies will be pooled from the two sources and summarized.

9.4.4 Positive β-hCG

Positive β-hCG is defined as a positive serum β-hCG test at 2 weeks and 2 weeks + 3-4 days post-oocyte retrieval. To allow for assessment that might occur outside the visit window, a subject will be considered to be biochemically pregnant if there are two positive assessment at least two days apart on or after 2 weeks post-oocyte retrieval. The positive β-hCG rate will be derived for subjects in the mITT and safety cohorts for subjects that had oocyte(s) retrieved. Associated 95% confidence intervals for these event rate estimates will also be presented.

Subjects who do not have two β-hCG tests due to missing data, early withdrawal, or any other reason will be counted as not having positive β-hCG unless a positive result is observed at a later pregnancy assessment. For example, if the outcome of these β-hCG tests is missing and a TVUS confirms clinical pregnancy at a later date then β-hCG will be imputed as ‘positive’.

9.4.5 Clinical Pregnancy

Clinical pregnancy is defined as a TVUS showing at least 1 intrauterine gestational sac with fetal heartbeat at 6 and 10 weeks (i.e., up to 42 and 70 days) post-oocyte retrieval. The frequency and proportion of subjects that are clinically pregnant at 6 and 10 weeks post-oocyte retrieval will be derived for subjects in the mITT and safety cohorts for subjects that had oocyte(s) retrieved. Associated 95% confidence intervals for these event rate estimates will also be presented.

For the clinical pregnancy rate derivation at 6 weeks post-oocyte retrieval, if a subject has missing TVUS assessment and no other TVUS assessment at a later date, the subject will be counted as not having a clinical pregnancy. In contrast, if there is a later assessment with at least 1 intrauterine gestational sac with fetal heartbeat, the subject will be imputed as having a clinical pregnancy at 6 weeks post-oocyte retrieval.

9.5 Other Assessments

9.5.1 Number and Size of Follicles during Stimulation

For each subject, the number of follicles on stimulation day 1 and last day of stimulation will be summarized for the mITT and ES cohorts. The summary will include the following:

1. the number of subjects that have follicles of size <10 mm, 10 mm, 11 mm, 12 mm, 13 mm, 14 mm, 15 mm, 16 mm, and ≥17 mm,

2. a summary of the number of follicles ≥17 mm (the count for each subject will be summarized across subjects),

3. the number of subjects that have at least two folliciles that are ≥17 mm, and

4. a summary of the percentage of follicles ≥17 mm, ≥15 mm, and ≥12 mm (for each subject the percentage will be first derived and this subject level sammary will be summarized across subject using mean, SD, etc.).

9.5.2 E2 and Progesterone Profiles

The central laboratory E2 and progesterone profiles will be summarized for mITT and ES cohorts, using descriptive statistics by scheduled visit, as well as for the change from screening for post-screening visits.

For patients in the mITT population, the local laboratory E2 assessments will be summarized by visit.

9.5.3 Oocytes

For each subject, the number of oocytes retrieved, the number of germinal vesicle, metaphase I, metaphase II and Degenerated oocytes will be summarized by frequency distribution and by descriptive statistics for the mITT and ES cohorts.

9.5.4 Characterization of Fertilized Oocytes one Day (Day 1) after Oocyte Retrieval

The status of fertlized oocytes will be summarized 1 day after oocyte retrieval for the mITT and ES cohorts. The summary will be both at the subject level and at the oocyte level.

For the summary at the subject level, for each subject, the percentage of fertilized oocytes with pronuclei 2 pn will be derived. The percentage will be derived over the number of oocytes retrieved. These percentage will further be summarized using descriptive statastistics (i.e., mean, SD, etc.). The percentage of fertilized oocytes with continue destiny will be similarly analyzed.

For the fertlized oocyte level summary, the number and percentage of fertilized oocytes falling in each category of pronuclei will be presented (i.e., > 2 pn, 2 pn, 1 pn, 0 pn and damaged). The percentage derivation will be over the total number of oocytes classified.

These summaries will be obtained for mITT and ES cohorts.

9.5.5 Quality of Blastocysts 5 Days (Day 5) after Oocyte Retrieval

The quality of blastocysts 5 days after oocyte retrieval will be assessed at both the blastocyst level and at the subject level. The evaluation of fertilized oocytes at Day 5 will consist of assessment of embryo stage and classification of blastocysts according to blastocyst expansion and hatching status, blastocyst inner cell mass grading and trophectoderm grading. Furthermore, the destiny of the embryo will be recorded.

Embryo stage is classified as blastocyst, morula, degenerated or cleavage stage. For embryos still at the cleavage stage the number of blastomeres will be recorded.

Destiny at Day 5 is either transferred, cryopreserved or out of trial.

The scoring of blastocysts is based on the classification system by Gardner & Schoolcraft[3].

Blastocyst expansion and hatching status will be assessed as one of the following:

1. An early blastocyst, blastocoel being less than half volume of that of the embryo
2. A blastocyst with a blastocoel whose volume is half of, or greater than half of, that of the embryo
3. A blastocyst with a blastocoel completely filling the embryo
4. An expanded blastocyst with a blastocoel volume larger than that of the early embryo, with a thinning zona
5. A hatching blastocyst with the trophectoderm starting to herniate through the zona
6. A hatched blastocyst, in which the blastocyst has completely escaped from the zona

For blastocysts with expansion and hatching status on Day 5, blastocyst inner cell mass grading and trophectoderm grading will be evaluated.

Blastocyst inner cell mass grading will be assessed as one of the following:

A. Tightly packed, many cells
B. Loosely grouped, several cells
C. Very few cells

Trophectoderm grading will be assessed as one of the following:

A. Many cells forming a cohesive epithelium
B. Few cells forming a loose epithelium
C. Very few, large cells

Based on the blastocyst expansion and hatching status, blastocyst inner cell mass grading and trophectoderm grading, the embryo will be classified as a good-quality blastocyst if the grade is 3BB or above as illustrated in Table 3.

Table 3 Good-Quality Blastocysts

3AA	4AA	5AA	6AA
3AB	4AB	5AB	6AB
3AC	4AC	5AC	6AC
3BA	4BA	5BA	6BA
3BB	4BB	5BB	6BB
	4BC	5BC	6BC
	4CA	5CA	6CA
	4CB	5CB	6CB
	4CC	5CC	6CC

The embryos on Day 5 will be summarised on both the embryo level and at the subject level.

Embryo Level

At the embryo level all embryos evaluated will be included in the tables when reporting embryo stage and destiny of the embryo. For summary of blastocyst expansion and hatching status, blastocyst inner cell mass grading and trophectoderm grading all available embryo evaluations will be included. Frequency tables will be produced for the embryo stage, destiny, blastocyst expansion and hatching status, blastocyst inner cell mass grading and trophectoderm grading. These summaries will be obtained for the subjects in the mITT and ES cohorts.

Subject Level

At the subject level the following will be derived for all subjects in the mITT and ES cohorts for the subjects that had oocytes retrieved:

- Number of blastocysts at Day 5
- Number of good-quality blastocysts at Day 5

These counts will be summarized using descriptive statistics of mean, median, SD, minimum and maximum.

9.5.6 Endometrial Thickness

Endometrial thickness at the time of blastocyst transfer will be summarized for the mITT and ES cohorts.

10 Safety

10.1 General Considerations

Safety parameters will be evaluated for the ES and safety cohorts. Safety summaries will be presented by the 5 phases of the trial. These phases are defined as follows:

- Screening Phase: From the date of signing the informed consent through the day before the start of the oral contraceptive treatment.
- Suppression Phase: From the start of oral contraceptive treatment until the day before the start of the stimulation phase (start of MENOPUR treatment).
- Stimulation Phase: From the date of the first dose of MENOPUR through the day before the start of IMP.
- Treatment Phase: From the date of the first dose of IMP (PVR) through the end of the trial.
- Study Phase: From the date of the first dose of oral contraceptive through the end of the trial.

10.2 Adverse Events

Adverse events (AEs) are classified according to the latest available version of Medical Dictionary for Regulatory Activities (MedDRA).

Written narratives will be issued for all serious AEs (including deaths). Treatment-emergent AEs are AEs that occur during the ***treatment phase*** or a pre-existing medical condition that worsened in intensity during the ***treatment phase***.

10.2.1 Overview of Treatment-Emergent Adverse Events

A TEAE overview summary table will be prepared including the number of subjects reporting a TEAE, the percentage of subjects (%) with a TEAE, and the number of events (E) reported, for the following categories:

- Treatment-emergent adverse events
- Treatment-emergent Deaths
- Treatment-emergent Serious adverse events
- Treatment-emergent adverse events leading to withdrawal
- Treatment-emergent severe adverse events
- Treatment-emergent Adverse drug reactions

10.2.2 Incidence of Adverse Events

Treatment-emergent adverse events will be summarised in a Table by SOC and PT of MedDRA using the latest available version. The Table will display the total number of subjects reporting an AE, the percentage of subjects (%) with an AE, and the number of events (E) reported. AEs will be

presented by system organ class (SOC) sorted alphabetically and preferred term (PT) sorted in decreasing frequency of occurrence.

Summary tables will be prepared for:

- All Treatment-emergent adverse events
- Treatment-emergent adverse events with an incidence [≥ 5%] of subjects in any treatment group
- Treatment-emergent adverse events by causality (related/unrelated)
- Treatment-emergent adverse events leading to death
- Treatment-emergent adverse events by intensity
- Treatment-emergent Serious adverse events
- Treatment-emergent Adverse events leading to withdrawal
- Non-serious treatment-emergent AEs with an incidence of at least 5%

Furthermore, the overall AE summary table by system organ class and preferred term will also be produced for the study phase, suppression phase, and the stimulation phase for the ES and safety cohorts.

Supporting data listings will be provided for:

1. All treatment-emergent adverse events sorted by centre and subject number
2. All treatment-emergent adverse events sorted by MedDRA System Organ Class and Preferred Term
3. Serious adverse events
4. Adverse events leading to death
5. Adverse events leading to withdrawal.

For listings specified in 3-5, there will be a variable that will indicate which phase of the trial each event occurred.

A separate AE listing will be provided for the screening phase, suppression phase, stimulation phase, and treatment phase of the trial.

10.2.3 Treatment-Emergent AEs of Special Interest

Standardized criteria will be applied to some AEs of special interest to establish intensity/grade. For AEs associated with vaginal bleeding/spotting, vaginal pain and irritation, and vaginal or cervical abrasions and lesions, standardized criteria will be used by the sites to record the intensity of these AEs (Protocol Section 8). For these AEs and AEs associated with vaginal hemorrhage, vaginal infection and vaginal adhesions, a summary table of the treatment-emergent AEs of special interest will be produced by preferred term, grade (intensity), and seriousness for subjects in the safety cohort for the treatment phase and the study phase. These AEs will be identified using preferred terms.

10.3 Safety Laboratory Variables

Baseline for all laboratory analyses will be the values obtained at visit 4. If visit 4 is missing, then the last assessment prior to visit 4 will be used. Treatment-emergent laboratory data will include tests completed after the first dose of IMP through the residual time of drug effect. End of trial will include the last post-baseline observation during the trial.

Laboratory variables will be grouped under "Hematology", "Clinical Chemistry" or "Urinalysis".

10.3.1 Summary Statistics

For treatment phase of the trial, for the safety cohort, mean change and mean percentage (%) change from baseline at end of trial will be presented for each laboratory variable. In addition, descriptive statistics, i.e., the number of subjects with data, mean (standard deviation), median, minimum, and maximum values, will be presented for observed values and change from baseline at each time-point for each laboratory variable. The screening value will also be summarized for both the safety and the ES cohorts.

10.3.2 Laboratory Variable Shift Tables

Changes relative to normal ranges are presented with shift tables with total number of subjects, and number and percent of subjects who experienced a shift from baseline to their end of trial visit. The following categories for shift tables are defined:

- Low: Values which are below the lower reference range limit;
- Normal: Values which are within the lower and upper reference range;
- High: Values which are above the upper reference range limit.

For all hematology and clinical chemistry variables, shift tables will be prepared to show all shifts. More specifically, for hematology and clinical chemistry, tables presenting the changes from *Low* or *Normal* to *High* and from *High* or *Normal* to *Low* will be provided.

10.3.3 Data Listings

Data listings will be prepared for all subjects with any abnormal laboratory value at any time-point (including screening, baseline).

10.3.4 Urinalysis

Urinalysis parameters will be summarized by visit for the ES and safety cohorts. For continuous variables, mean, standard deviation, median, minimum and maximum will be used to summarize the assessment. In contrast, for categorical parameters, percentages and frequencies will be used. Furthermore, for each urinalysis parameter the number of subjects with abnormal laboratory values will be summarized by visit for the ES and safety cohorts.

10.3.5 Vital Signs

Baseline for all vital signs analyses will be calculated in the same manner as the safety laboratory variables in Section 10.3. Treatment-emergent vital signs data will include tests completed after the

first dose of IMP through the end of the trial. End of trial assessment is defined to be the last post-baseline observation during the treatment phase of the trial.

10.3.5.1 Summary Statistics

For the treatment phase of the trial, for the safety cohort, mean change and mean percentage (%) change from baseline to the end of trial will be presented for each vital signs variable. In addition, descriptive statistics, i.e., the number of subjects with data, mean (standard deviation), median, minimum, and maximum values, will be presented for observed values and change from baseline at each time-point for each vital signs variable.

10.3.5.2 Notable Changes in Vital Signs

Summary tables will be prepared displaying the number and percentage of subjects with normal baselines who had one or more pre-specified notable treatment-emergent values, according to the definition in Appendix 1.

10.3.5.3 Data Listings

Data listings will be prepared by centre for all subjects with any notable vital signs value at any time-point (including screening, baseline).

10.4 Other Safety Variables

10.4.1 Bleeding Log Assessment

At the start of PVR, each subject will receive a bleeding log. For each incidence of sanitary napkin use for bleeding/spotting, subjects will be instructed to record the date, time of day, and the severity of the bleeding. For each subject, for each day: the grade (intensity) of bleeding associated with each sanitary napkin use will be summed. The sum of the daily grades (intensity) of bleeding will be averaged by visit window (see Table 4) where the visit window will cover the subject's Bleeding Log recording period (i.e., from the start of the PVR to the End-of-trial visit). Finally, this subject data will be summarized across all subjects for the safety cohort by visit window. The summary will include mean, median, standard deviation, minimum and maximum.

The analysis outlined above will be repeated removing the date of fresh transfer, periods associated with AEs reported as vaginal hemorrhage, the primary endpoint of spontaneous abortion and information collected after the cessation of the use of PVR. For vaginal hemorrhaging, spontaneous abortion and biochemical abortion, this will be achieved by removing bleeding log assessment collected starting from two days prior to the start of such an event up to two days after the end of such an event. Furthermore, for subjects with any negative β-hCG result, the bleeding log assessment will be summarized.

For subjects that discontinue from PVR early, the bleeding logs will be calculated normally (as per Table 4) until PVR is discontinued. Bleeding log entries after the visit preceding the Early Withdrawal Visit and bleeding log entries after PVR is withdrawn will be assigned to Visit 9. At the end of study visit and beyond, bleeding log entries will be assigned to Visit 10.

Table 4 Visits numbers, nominal visit days from oocyte retrieval date and associated visit windows

Visit	Nominal visit date from date of oocyte retrieval (Day 1= date of oocyte retrieval)	Visit window (Day 1= date of oocyte retrieval)
Visit 3	Day 1	
Visit 4	Day 2	Start of IMP - Day 3
Visit 5	Day 5	Day 4 – Day 9
Visit 6	Day 14±1	Day 10 – Day 15
Visit 6P	Day 17±1	Day 16 – Day 19
Visit 7	Day 21±1	Day 20 – Day 26
Visit 7P	Day 31 ±1	Day 27 – Day 36
Visit 8	Day 42±1	Day 37 – Day 56
Visit 9	Day 70±1	Day 57 – Day 77
Visit 10	Day 84	Day 78 – Day 90

10.4.2 Vaginal Hemorrhage

Vaginal hemorrhage is defined in Section 8.3.6 of the protocol and the assessment of vaginal hemorrhage will be collected in a specific eCRF module. Based on the eCRF module, the number of subjects with vaginal hemorrhage will be summarized for the Safety cohort.

10.4.3 Pelvic Examination

A pelvic examination will be performed at screening, Visits 4, 7P, 8, 9, and 10 to document the presence and grade (intensity) of any pain, irritation, abrasions, or lesions on the cervix or vagina using standardized criteria, and the presence and grade (intensity) of any vaginal adhesions using standardized criteria. If any lesions or abrasions are found on the cervix or vagina or vaginal adhesions are noted, another examination will be performed 2-4 days later. The grade (intensity) for each type of assessment will be summarized by visit for the safety cohort. Furthermore, a shift table will be presented summarizing the change in grade (intensity) from Visit 4 to the End-of-trial, where the End-of-trial is defined to be the last assessment in the treatment phase.

An associated listing will also be produced.

11 Interim Analyses

No interim analysis is planned for this trial.

12 Deviations from Protocol

The following changes and additions from the planned displays and analyses described in the previous version of the clinical trial protocol have been implemented in the statistical analysis plan:

- All 90% confidence intervals have been replaced with 95% confidence intervals based on FDA input on the statistical analysis plan.
- The embryo quality analysis that is described in this SAP was not mentioned in the clinical trial protocol.
- The definition of the Eligible Subjects cohort is taken to be a subject that has received at least one dose of oral contraceptives, leuprolide acetate, and MENOPUR. This aligns with the analogous cohort from the DR-PGN-302 study. In the clinical trial protocol, an “or” is used instead of an “and”.

These revisions have been implemented in the clinical trial protocol version 6.0.

13 References

[1] Center for Disease Control 2005 Assisted Reproductive Technology Success Rates. U.S. Department of Health and Human Services; http://www.cdc.gov/ART/ART2005/508PDF/2005ART508.pdf.

[2] Practice Committee of the Society for Assisted Reproductive Technology and Practice Committee of the American Society for Reproductive Medicine. Guidance on the limits to the number of embryos to transfer: a committee opinion. Fertil Steril 2017; 107:901–3.

[3] Gardner DK, Schoolcraft WB. In vitro culture of human blastocysts. In: Towards reproductive certainty (Eds Jansen R & Mortimer D). The plenary proceedings of the 11th world congress on in vitro fertilization and human reproductive genetics. The Parthenon Publishing Group. 1999. Pp 378-388

14 APPENDICES

Appendix 1 Notable Vital Signs

Table 5 Notable Criteria for Vital Signs*

Variable	Criterion Value	Change from Baseline
Systolic blood pressure	≥ 180 mmHg ≤ 90 mmHg	Increase of ≥ 20 mmHg Decrease of ≥ 20 mmHg
Diastolic blood pressure	≥ 105 mmHg ≤ 50 mmHg	Increase of ≥ 15 mmHg Decrease of ≥ 15 mmHg
Pulse rate	≥ 120 bpm ≤ 50 bpm	Increase of ≥ 15 bpm Decrease of ≥ 15 bpm
Body weight	None	Increase of ≥ 7% Decrease of ≥ 7%
Body temperature**	≥ 38.3° C	

* To be identified as notable abnormal, a treatment value must meet the criterion value and also the specified change from baseline.

** The notable criteria is met for body temperature whenever this criteria is met.

Appendix 2 Primary Analysis Simulation Code and the Sensitivity Analysis Operating Characteristics

Primary Analysis Simulation Code

Below is the R code that runs the simulation described in the sample size justification.

```
#################
### LIBRARIES ###
#################

library(ggplot2)
library(tidyverse)

#############
### SETUP ###
#############

### SET THE SEED ###
set.seed(7639216)

### THE NUMBER OF SIMULATION REPS ###
n.sim <- 10000

### THE NUMBER OF SUBJECTS IN THE MODIFIED ITT ANALYSIS POPULATION ###
n.mitt <- seq(from = 100, to = 400, by = 5)

### THE TRUE PR(SPONTANEOUS ABORTION | IN THE MITT ANALYSIS POPULATION)
###
p.sa <- seq(from = 0.08, to = 0.1, by = 0.01)

### A PLACE TO STORE THE RESULTS ###
sim.res <- matrix(NA, nrow = length(n.mitt), ncol = length(p.sa))
rownames(sim.res) <- n.mitt
colnames(sim.res) <- paste0(p.sa * 100, "%")

##################
### SIMULATION ###
##################

pb <- txtProgressBar(min = 1, max = length(n.mitt), style = 3)

for(a in 1:length(n.mitt)){

  setTxtProgressBar(pb, value = a)
```

```
  for(b in 1:length(p.sa)){

    sim.sara <- rbinom(n = n.sim, size = n.mitt[a], prob = p.sa[b])

    sara.res <- unlist(lapply(X = sim.sara, FUN = function(x, n = n.mitt[a]){

      ifelse(binom.test(x = x, n = n, p = 0.15, alternative = c("two.sided"), conf.level =
0.95)$conf.int[2] < 0.15, 1, 0)

    }))

    sim.res[a, b] <- mean(sara.res) * 100

  }

}

########################
### PLOT THE RESULTS ###
########################

### MOLD THE RESULTS FOR PLOTTING ###
sim.res.long <- data.frame(N = as.numeric(rownames(sim.res)), SA8PCT = sim.res[, 1], SA9PCT =
sim.res[, 2], SA10PCT = sim.res[, 3]) %>%
  gather(key = "SA", value = "POWER", 2:4)

p <- ggplot(sim.res.long, aes(x = N, y = POWER, group = SA)) +
  geom_hline(yintercept = 80, linetype = "dashed", color = "grey", size = 2) +
  geom_line(aes(color = SA), size = 1.5) +
  scale_x_continuous(name = "Sample Size (in MITT Analysis Population)", breaks = seq(from =
100, to = 400, by = 20)) +
  scale_y_continuous(name = "Power", limits = c(0, 100)) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_color_discrete(name = "Spontaneous\nAbortion\nRate",
                       breaks = c("SA8PCT", "SA9PCT", "SA10PCT"),
                       labels = c("8%", "9%", "10%"))

########################
### TABULAR OUTPUT ###
########################

### FIND WHICH RESULTS ARE ABOVE 80% POWER ###
sim.res.80 <- sim.res >= 80

### FIND THE INDICES ###
```

```
sim.res.ind <- apply(X = sim.res.80, MARGIN = 2, FUN = function(x){

  TEMP <- length(x)
  if(any(x)){

    TEMP <- min(which(x))

  }
  TEMP

})

### REPORT THE POWER FOR THESE INDICES

sim.table <- sim.res[sim.res.ind, ]
```

Hierarchical Modelling Details

Define n_1 as the number of subjects in the current trial and Y_{1i} as the indicator of spontaneous abortion for the i^{th} subject in the current trial where $Y_1 \sim \text{Bin}(n_1, \pi_1)$. Let π_1 be the true rate of spontaneous abortions in the current trial.

The previous trial (Study DR-PGN-302) reported a spontaneous rate of abortion of 55/549. This prior information about abortion rate is incorporated into the modelling framework through a hierarchical model. Dynamic borrowing between this trial's results and the current trial's data allows borrowing to occur to the extent indicated by these two data sources. More borrowing occurs when the two rates are similar while less borrowing occurs when they differ. This "dynamic" borrowing property is distinct from other approaches which use a fixed informative prior or apriori assume an amount of borrowing between these two data sources. Let $Y_0 \sim \text{Bin}(n_0=549, \pi_0)$ where π_0 is the true rate of spontaneous abortions from the previous trial.

We transform to the logit scale for modelling purposes. Let $\theta_1 = \log(\pi_1/(1-\pi_1))$ and $\theta_0 = \log(\pi_0/(1-\pi_0))$. The hierarchical model assumes that θ_0 and θ_1 have an across studies distribution:

$$\theta_1 , \theta_0 \sim N(\mu, \tau^2)$$

The across trial mean μ and variance τ^2 are unknown and hence have a prior distribution which is combined with the data to produce estimates of μ and τ^2:

$$\mu \sim N(-2.2, 2.25)$$

$$\tau^2 \sim IG(0.125, 0.005)$$

The mean parameter μ assumes a relatively vague prior on the original scale and centered at 10%. The variance component τ^2 controls the degree of borrowing among studies. Small values of τ^2 result in a greater degree of borrowing while large values of τ^2 correspond to less borrowing. The parameter τ^2 is estimated using the data, so the observed between studies variation is a key

component of the model behavior. The prior specification on τ^2 corresponds to an inverse-gamma distribution with mean 0.2 and weight 0.25.

Operating Characteristics

The planned sample size is based only on subjects that undergo fresh embryo transfer. No further drop-outs are assumed. The analysis is evaluated across a range of true effects for the current trial's spontaneous abortion rate. For the null scenario where in truth the PVR spontaneous abortion rate is 15%, 100,000 sets of trials were simulated. For all other assumed scenarios, 5,000 sets of trials were simulated.

Overall Success

We evaluated the overall probability that the 95% credible interval is entirely below 15% for each scenario of the true PVR rate. If PVR truly maintains an abortion rate below 15%, then this represents the power for achieving a credible interval entirely below 15%. If the true rate of abortions for PVR is 15%, then obtaining a credible interval entirely below 15% is the type I error. Results are provided in the table below.

True PVR Rate	8%	9%	10%	11%	12%	13%	14%	15%
Trial Success	0.979	0.933	0.830	0.670	0.485	0.297	0.167	0.084

When in truth the PVR spontaneous abortion rate is 9%, the trial has 93.3% power of achieving a 95% credible interval entirely below 15%. Alternatively, when in truth the spontaneous abortion rate is 15%, the trial has a 8.4% rate of success.

Inference Associated with Various Observed Rates

For various observed rates of spontaneous abortions, the corresponding primary analysis based on the hierarchical model is described below. The posterior mean, 95% credible interval, effective number of borrowed subjects, and the borrowing percentage (out of the n=549 subjects from the previous trial) is provided. The effective number of borrowed subjects calculation [$E(\pi_1)*(1-E(\pi_1))/Var(\pi_1)–1–240$] is somewhat approximate as it is based on the binomial likelihood. Results are provided in the below table.

Observed Data	Estimated Rate (%) (95% CI)	Effective Number of Borrowed Subjects	Borrowing Percentage
19/240 (7.92%)	8.6% (5.7, 11.7)	103	18.9%
24/240 (10.00%)	10.0% (7.2, 13.3)	130	23.7%
36/240 (15.00%)	13.8% (10.1, 18.5)	9	1.7%

Borrowing based on the hierarchical modelling is dynamic, with higher borrowing achieved when the current trial's results are similar to the previous data and minimal borrowing when the current trial's results are at 15%. This is evidenced by both the decreasing number of borrowed subjects and the increasing width of the credible intervals. Also, based on 240 subjects being included in the

analysis, the maximal number of observed responses which maintains a 95% credible interval entirely below 15% is 28.

R2OpenBUGS Codes for Analysing the Data at the End of the Trial

```
for (i in 1:N) {
                    y[i]  ~ dbin(theta[i], n[i])
                     logit(theta[i]) <- logit.theta[i]
                     logit.theta[i] ~ dnorm( mu, tau)
                       }
          mu ~ dnorm(-2.2, 0.4444444)
          tau ~ dgamma(0.125, 0.005)
```

Appendix 3 Preferred Terms Associated With Spontaneous Abortion and Ectopic pregnancy

All the preferred terms associated with Spontaneous abortion are listed below. This is based on MedDRA version 21.0. Prior to dry-run and DBL, the dictionary will be checked to make sure there is no change of terms with dictionary update.

Spontaneous abortion associated preferred terms
Abortion
Abortion complete
Abortion complete complicated
Abortion complicated
Abortion early
Abortion incomplete
Abortion incomplete complicated
Abortion induced
Abortion induced complete
Abortion induced complete complicated
Abortion induced complicated
Abortion induced incomplete
Abortion induced incomplete complicated
Abortion infected
Abortion late
Abortion missed
Abortion of ectopic pregnancy
Abortion spontaneous
Abortion spontaneous complete
Abortion spontaneous complete complicated
Abortion spontaneous complicated
Abortion spontaneous incomplete
Abortion spontaneous incomplete complicated

Abortion threatened

The relevant preferred terms associated with biochemical abortion are listed below. This is based on MedDRA version 21.0. Prior to dry-run and DBL, the dictionary will be checked to make sure there is no change of terms with dictionary update.

Biochemical abortion associated preferred terms
Biochemical pregnancy

The relevant preferred terms associated with ectopic pregnancy are listed below. This is based on MedDRA version 21.0. Prior to dry-run and DBL, the dictionary will be checked to make sure there is no change of terms with dictionary update.

Ectopic pregnancy associated preferred terms
Abortion of ectopic pregnancy
Ectopic pregnancy
Ectopic pregnancy termination
Ectopic pregnancy under hormonal contraception
Ectopic pregnancy with contraceptive device
Ruptured ectopic pregnancy

The relevant preferred terms associated with adverse events of special interest (Appendices 2, 3, and 4 of the Protocol) are listed below. This is based on MedDRA version 21.0. Prior to dry-run and DBL, the dictionary will be checked to make sure there is no change of terms with dictionary update.

Preferred terms associated with adverse events of special interest
Vulvovaginal pain
Dyspareunia
Vulvovaginal discomfort

Vulvovaginal pruritus
Vulval oedema
Vulvovaginal erythema
Vulvovaginal rash
Bartholin's cyst
Urethral cyst
Vulval disorder
Vaginal oedema
Vulvovaginal dryness
Vaginal discharge
Vaginal erosion
Vaginal lesion
Cervix oedema
Cervical friability
Cervix erythema
Cervical discharge
Cervix disorder
Vulvovaginal adhesion
Vaginal haemorrhage
Vaginal infection
Uterine cervical erosion