# Statistical Analysis Plan

## Developmental Clinical Studies - Reversing endometrial glucocorticoid deficiency in heavy menstrual bleeding

## "Dexamethasone For Excessive Menstruation" **DexFEM Study III**

| | |
|---|---|
| **Version No** | 3.0 |
| **Date** | 29 June 2018 |
| **Author(s)** | Christopher Weir |
| **CI Name** | Professor Hilary Critchley |
| **CI Email address** | Hilary.Critchley@ed.ac.uk |
| **Co-sponsors** | University of Edinburgh & NHS Lothian<br>ACCORD<br>The Queen's Medical Research Institute<br>47 Little France Crescent<br>Edinburgh EH16 4TJ |
| **Funder** | Medical Research Council |
| **Funder Reference Number** | MR/J003611/1 |
| **REC Number** | 12/SS/0147 |
| **clinicaltrials.gov Number** | NCT01769820 |

| Signatures | |
|---|---|
| **Chief Investigator:**<br>Professor Hilary Critchley | **Date: 17/07/2018** |
| **Principal Co-Investigator:**<br>Dr Pamela Warner | **Date:** |
| **Trial Statistician:**<br>Professor Christopher Weir | **Date:** |

# Statistical Analysis Plan

## Developmental Clinical Studies - Reversing endometrial glucocorticoid deficiency in heavy menstrual bleeding

"Dexamethasone For Excessive Menstruation" **DexFEM Study III**

| | |
|---|---|
| **Version No** | 3.0 |
| **Date** | 29 June 2018 |
| **Author(s)** | Christopher Weir |
| **CI Name** | Professor Hilary Critchley |
| **CI Email address** | Hilary.Critchley@ed.ac.uk |
| **Co-sponsors** | University of Edinburgh & NHS Lothian<br>ACCORD<br>The Queen's Medical Research Institute<br>47 Little France Crescent<br>Edinburgh EH16 4TJ |
| **Funder** | Medical Research Council |
| **Funder Reference Number** | MR/J003611/1 |
| **REC Number** | 12/SS/0147 |
| **clinicaltrials.gov Number** | NCT01769820 |

| Signatures | | | |
|---|---|---|---|
| **Chief Investigator:** Professor Hilary Critchley | | **Date:** | |
| **Principal Co-Investigator:** Dr Pamela Warner | PE Warner | **Date:** | 03/07/2018 |
| **Trial Statistician:** Professor Christopher Weir | | **Date:** | 03/07/2018 |

| Document Control | | |
|---|---|---|
| **Version No** | **Date** | **Summary of Revisions** |
| 1.0 | 05Jul2016 | Initial Creation |
| 2.0 | 11Sep2017 | Addition of document history section. Updated reference to latest version of protocol, v11. |
| 3.0 | 29Jun2018 | Reformatted according to current ECTU SAP template as part of routine review prior to database lock. Updated reference to latest version of protocol, v12. Removal of GGT from list of blood results as not recorded during the trial. Clarification of the analysis population to be used for each analysis. Specification of how missing primary and secondary outcome values will be handled. Added summary of ovulation data based on cycle T3 urine samples. |

# Table of Contents

## List of Abbreviations

| Abbreviation | Full name |
|---|---|
| AE | Adverse event |
| CI | Confidence interval |
| CrI | Credible interval |
| Dex | Dexamethasone |
| HMB | Heavy menstrual bleeding |
| MBL | Menstrual blood loss |
| MCMC | Markov Chain Monte Carlo |
| MRI | Magnetic resonance imaging |
| NDLM | Normal dynamic linear model |
| S1 | Screening menstrual cycle 1 |
| S2 | Screening menstrual cycle 2 |
| SAE | Serious adverse event |
| SAP | Statistical analysis plan |
| SD | Standard deviation |
| SUSAR | Suspected unexpected serious adverse reaction |
| T1 | Treatment menstrual cycle 1 |
| T2 | Treatment menstrual cycle 2 |
| T3 | Treatment menstrual cycle 3 |

# 1. Introduction

**Overview of DexFEM research programme**

Dexamethasone For Excessive Menstruation (DexFEM) Study III is the final part of a programme of developmental clinical studies investigating the potential for Dexamethasone to reverse endometrial glucocorticoid deficiency in heavy menstrual bleeding. Study 1 was a mechanistic study exploring the relationships, with and without Dexamethasone treatment, between menstrual blood loss (MBL), endometrial MRI perfusion and endometrial biopsy data. Study 2 was a randomised, placebo-controlled crossover trial of Dexamethasone to gather preliminary safety and efficacy data in addition to investigating the relationship between patient diary data and measured MBL.

Study 3 is a double-blind, parallel group, placebo-controlled, response-adaptive randomised trial of six doses of Dexamethasone. Total daily doses of placebo, 0.4, 0.8, 1.0, 1.2, 1.5 or 1.8mg of Dexamethasone are administered via twice daily oral doses for 5 days during the luteal phase of menstruation. 108 women will be randomised with the aim of achieving 100 who complete the study. The Dexamethasone randomisation probabilities will be adapted, based on analysis of the MBL primary outcome data gathered to date, after 16, 32, 50, 66 and 84 patients have been randomised.

**General Approach**

This statistical analysis plan (SAP) covers the statistical analysis of DexFEM Study 3 only. It is written with reference to the current version of the protocol, version 12, dated 02 October 2017.

**Primary Objective**

To identify the optimal dose of oral Dexamethasone (Dex) for amelioration of HMB in women with objectively verified HMB.

**Secondary Objectives**

1. To gather safety data for Dexamethasone in women with objectively verified HMB.
2. To collect information to enable development of local delivery system(s) for Dexamethasone.
3. To establish that an adaptive trial design is an efficient method of screening doses to identify which should be taken forward to a Phase III trial against standard treatment.

Secondary objective 2 will be addressed via analysis of mechanistic data from Study 1. Secondary objective 3 was established in the simulation studies used to develop the Study 3 adaptive design.

# 2. Statistical Methods section from the protocol

*9.2.3 Study 3*

*9.2.3.1 Baseline characteristics*

*Summary data will be used to compare baseline characteristics between women randomised to placebo and each of the Dexamethasone doses: mean, standard deviation, median, minimum, maximum for continuous variables, and number (percentage) of individuals for categorical variables.*

*9.2.3.2 Primary endpoint*
*The dose-response curve for change in MBL between baseline and cycles during randomised treatment will be analysed using a Normal Dynamic Linear Model (NDLM) which is flexible and requires few assumptions about the shape of the underlying dose-response curve. The NDLM requires a Bayesian analysis framework, including the specification of prior distributions for all parameters. Suitable prior distributions will be confirmed by modelling and simulations during the work-up phase. The NDLM analysis will determine which of the doses studied is optimal to take forward for further study (in terms of posterior probability of efficacy, efficacy being defined as a 25% reduction in MBL versus baseline). For each Dexamethasone dose, a 95% credible interval will be calculated for the mean difference in MBL change versus placebo. Mean baseline MBL will be included as a covariate in the NDLM.*

*9.2.3.3 Secondary endpoints*
*The menstrual diary score for menstrual period volume will be analysed using the method given for the primary endpoint in section 9.2.3.2. Binary or ordinal secondary endpoints will be analysed using a generalised dynamic linear model. For each Dexamethasone dose, a 95% credible interval will be calculated for the odds ratio versus placebo.*

## 3. Overall Statistical Principles

Throughout, an overall two-sided significance level of 5% will be used in frequentist analysis. Where Bayesian analysis is conducted, 95% credible intervals (CrI) will be presented.

Categorical variables will be summarised using frequencies and percentages; continuous variables will be summarised using the mean, standard deviation (SD), median, lower quartile, upper quartile, minimum and maximum values.

The planned analyses will be performed using the SAS and WinBUGS statistical software.

**Analysis populations**
Randomised set
All patients who were enrolled in the study and randomised.

Full analysis set
The full analysis set will be defined as all participants who were randomised and had at least one on-treatment MBL measurement recorded.

Safety analysis set
All participants who were randomised and received at least one dose of trial medication will be included in the safety analysis set. Results will be presented according to treatment received, rather than treatment allocated at randomisation, if these differ.

**Outcomes**
Primary outcome
- Change in mean MBL between screening cycles S1, S2 and cycles T2, T3 during randomised (Dexamethasone/placebo) treatment. Note that MBL is only measured in the second and third cycles (T2, T3) during randomised treatment.

Secondary outcomes

- Menstrual diary score for volume of menstrual period
- Satisfaction with treatment
- Intolerable side-effects
- Period pain

## 4. List of Analyses

### 4.1 Recruitment and retention

The total number assessed for eligibility, randomised, receiving treatment, followed up and analysed will be reported by randomised group in accordance to the CONSORT 2010 statement [1].

Withdrawals from the trial by randomised group will be listed, giving the timing of withdrawal and, where available, the reason for withdrawal. Corresponding data will be listed for withdrawals from treatment who were still retained in follow-up. The number of participants in whom the blinding of randomised treatment allocation was broken early will be tabulated.

The numbers of participants receiving any treatment, reaching the end of each study treatment cycle, and receiving treatment in each of the three treatment cycles (T1, T2, T3) will be reported by randomised group.

A listing of protocol violations will be provided.

### 4.2 Baseline characteristics

The following categorical baseline characteristics will be reported for the randomised analysis set by randomised group:

- Ethnicity; number of births; number of miscarriages; number of terminations; current breastfeeding (Y/N); has painful periods (Y/N); takes illegal substances (Y/N); smoking history (current, previous or never); blood results of clinical concern (Y/N)
- Medical history (normal/abnormal): respiratory, cardiovascular, gastrointestinal, endocrine/metabolic, central nervous system, liver, renal, genitourinary, breast, musculoskeletal, haematological, dermatology; ear, eye, nose and throat; allergies; cancer (uterus, cervix, ovarian or breast); psychiatric
- Blood laboratory result: eGFR ≥60 mL/min or <60mL/min

The following continuous baseline characteristics will be reported for the randomised analysis set by randomised group:

- Age (years); years since last pregnancy (if applicable); age at menarche (years); systolic blood pressure (mmHg); diastolic blood pressure (mmHg); weight (kg); minimum length of menstrual cycle in previous 3 months (days); maximum length of menstrual cycle in previous 3 months (days); minimum length of menses in days in previous 3 months (days); maximum length of menses in days in previous 3 months (days); number of years that HMB has been a problem; units of alcohol consumed per week; cigarettes smoked per day; screening cycle S1 menstrual blood loss (mL); screening cycle S2 menstrual blood loss (mL); mean screening cycle menstrual blood loss (mL)
- Blood laboratory results: sodium (mmol/L); potassium (mmol/L); urea (mmol/L); creatinine (mmol/L); ALT (U/L); Alk Phos (U/L); bilirubin (umol/L); HbA1c (mmol/L); plasma glucose (mmol/L);

haemoglobin (g/dL); white cell count ($10^9$/L); neutrophils ($10^9$/L); lymphocytes ($10^9$/L); platelets ($10^9$/L)

## 4.3 Primary outcome

Analyses of the primary outcome will be performed using the full analysis set.

Descriptive statistics
The observed MBL data will be listed for cycles S1, S2, T2 and T3. We will also list the mean screening MBL, S = (S1+S2)/2, the mean treatment MBL, T = (T2+T3)/2, and the primary outcome: change in mean MBL, Y = (T-S), during randomised treatment. In the event of one MBL measurement being missing, the above calculations will be performed using the observed data only. Summary statistics for the primary outcome will be provided by randomised group.

Primary analysis
The main analysis of the primary outcome will use a Bayesian second order normal dynamic linear model (NDLM) [2]. The NDLM is specified in terms of observation and evolution equations.

The observation equation

$$Y_i = \theta_j + \beta Ybl_i + v_i \qquad j = 1,...,J;\ v_i \sim i.i.d.\ N(0, \sigma_v^2)$$

models $Y_i$, the primary outcome in subject $i$, as a function of $\theta_j$, the treatment effect at dose level $j$ and includes the mean-centred mean screening MBL, $Ybl_i$, as a covariate. $\beta$ is the regression coefficient for $Ybl_i$ and $v_i$ is the observational error term.

The evolution equation

$$\theta_j = \theta_{j-1} + \delta_{j-1} + \omega_j$$
$$\delta_j = \delta_{j-1} + \varepsilon_j \qquad j = 1,...,J;\ \omega_j, \varepsilon_j \sim i.i.d.\ N(0, \sigma_\omega^2)$$

equates the treatment effect at dose level $j$ to the treatment effect at the previous dose level plus a systematic deviation, $\delta_{j-1}$, and an evolution error, $\omega_j$. The change in treatment effect from dose level $j-1$ to dose level $j$, $\delta_{j-1}$, is modelled as a random walk with step size $N(\delta_{j-1}, \sigma_\omega^2)$. The model assumes linear changes in the treatment effect from one dose level to the next but does not restrict the overall curve to follow a particular parametric form. $v_i$, $\omega_j$ and $\varepsilon_j$ are mutually and internally independent. $\omega_j$ and $\varepsilon_j$ are assumed to have the same variance. The observation and evolution error variances are assumed constant for all subjects and dose levels.

We define the evolution error variance as a multiple of the observational error variance, i.e. $\sigma_\omega^2 = W\sigma_v^2$. A uniform $U(0.001, 100)$ prior distribution is used for $W$ and a vague half-normal prior distribution $N(0, 100)$, defined only on the non-negative range of the scale, is used for $\sigma_v^2$. We also specify vague $N(0, 10000)$ prior distributions for $\theta_0$, $\delta_0$ and $\beta$.

WinBUGS [3] will be used to fit this model using Markov Chain Monte Carlo (MCMC) methods. Model parameter estimates will be based on 10,000 simulated draws from the marginal posterior distributions, having discarded the first 5000 burn-in iterations from the sampler.

For each Dexamethasone dose, a 95% CrI will be calculated for the mean difference (Dexamethasone minus placebo) in the primary outcome. As a further indication of the degree of efficacy, the posterior probability that the best performing dose gives more than a 10mL reduction in mean MBL compared to placebo will be reported. For completeness, the posterior probability that the best performing dose gives any mean reduction in MBL relative to placebo will also be reported.

It has been advocated that one should stratify by adaptation stage in the final analysis of an adaptive trial to ensure a contemporaneous randomised comparison versus placebo within each stage [4]. We have opted not to follow this approach in DexFEM as the numbers of women studied within each adaptation stratum (between 16 and 18) will be too small to enable reliable fitting of the complex NDLM across 7 randomised groups.

Model checking
Convergence of the MCMC sampler will be monitored by sampling from two chains simultaneously using over-dispersed initial values and calculating the BGR diagnostic [5]. This assessment will be supported by visual inspection of the sampling histories.

Sensitivity analyses
The sensitivity of the results to the choice of vague priors for $W$, $\sigma_v^2$, $\theta_0$, $\delta_0$ and $\beta$ will be explored by substituting alternative variance values in the normally distributed priors and using different upper and lower limits in the uniform prior for $W$.

Secondary analysis
As a further indicator of which of the six Dexamethasone doses is best, the posterior probability of efficacy and its 95% CrI will be calculated for each dose. Here, efficacy will be defined as a 25% reduction from the mean baseline MBL (calculated as the average across all participants in the full analysis set).

### 4.4 Secondary outcomes

Secondary outcomes analysis will be performed on an analysis set equivalent to the full analysis set: all randomised patients with at least one on-treatment measurement of the secondary outcome recorded.

The menstrual diary score for menstrual period volume will be also be analysed using a normal dynamic linear model similar to that described above for the main analysis of the primary outcome. The menstrual diary score outcome will be defined as the change from baseline, calculated as ((T1+T2+T3)/3)-((S1+S2)/2). In the event of one or more menstrual diary scores being missing, the above calculations will be performed using the observed data only. Mean screening menstrual diary score will be included as a mean-centred covariate.

Each of the binary or dichotomised ordinal secondary endpoints (questions from menstrual diary questionnaire Q2, Q3a, Q3b, Q3c, Q3d, Q3e and treatment review questionnaire Q4a, Q4b, Q4c, Q5, Q6a, Q6b) will be analysed in a Bayesian comparison between the proportion for each Dexamethasone dose versus that for placebo. For each Dexamethasone dose, the difference in proportion (Dexamethasone dose minus placebo) and its 95% CrI will be calculated. A vague Uniform (0,1) prior will be placed on the proportion for each Dexamethasone dose and for placebo. The primary analysis for each diary question will use data from cycle T3; supporting analyses will use the same method to analyse data for each of cycles T1 and T2.

Menstrual diary questionnaire Q2 will be dichotomised as "light" or "medium" versus "heavy" or "very heavy". Menstrual diary questionnaire Q3a, Q3c, Q3e will be dichotomised as "much less" or "a bit less" versus "about same", "a bit more" or "much more". Menstrual diary questionnaire Q3b, Q3d will be dichotomised as "a bit more" or "much more" versus "about same", "a bit less" or "much less".

Treatment review questionnaire Q4a will be dichotomised as "much lighter bleeding" or "lighter bleeding" versus "about the same", "a bit worse" or "bleeding much worse".

Q4b will be dichotomised as "much less pain" or "less severe pain" versus "about the same", "a bit worse" or "pain much worse".

Q4c will be dichotomised as "much better" versus "a little better", "about the same", "a little worse" or "much worse".

Q5 will be dichotomised as "excellent", "very good" or "good" versus "fair" or "poor".

Q6a and Q6b will be dichotomised as "yes" versus "not sure" or "no".

The following posterior probabilities will also be calculated from treatment review questionnaire data:
• Q4a: probability ("lighter bleeding" or "much lighter bleeding") for best performing dose is at least 15% greater than for placebo
• Q4b: probability ("less severe pain" or "much less pain") for best performing dose is at least 15% greater than for placebo
• Q4c: probability ("felt much better") for best performing dose is at least 15% greater than for placebo
• Q6a: probability that Q6a="yes" for the best performing dose is at least 10% greater than for placebo
• Q6b: probability that Q6b="yes" for the best performing dose is at least 10% greater than for placebo

### 4.5 Exploratory analysis

Exploratory analyses will be performed using observed data from the randomised analysis set.

The following menstrual cycle data derived from urine sampling will be summarised, by randomised group and overall, according to study cycle and across the whole study:

- Menstrual cycle length (for S1, S2, T1, T2, T3)
- Timing of end of treatment in a study cycle, relative to start date of next period (T1, T2, T3)

Categorical data on whether cycle T3 was ovulatory will be summarised, by randomised group and overall, for two variables:

- Ovulatory cycle, judged using a wider luteal phase time window (starting from day [menstrual cycle length-12] )
- Ovulatory cycle, judged using a more specific luteal phase time window (starting from day [menstrual cycle length-7] )

### 4.6 Safety

Safety data will be reported using the safety analysis set.

#### Treatment completeness

Treatment completeness, defined as the treatment being taken in 10 consecutive mornings and evenings, will be summarised by randomised group and overall for (a) the proportion of participants taking complete treatment in each cycle and (b) the proportion of participants taking complete treatment in all study cycles. (a) and (b) will consider only cycles prior to withdrawal, for participants who withdrew from the study early. Where data are missing on whether treatment was taken on a particular day, it will be assumed not to have been taken unless the number of returned tablets indicates otherwise.

Blood laboratory results
The categorical and continuous blood laboratory results listed in the Baseline Characteristics section will be summarised at follow up (end of treatment cycle T3) by treatment received. A pooled summary across all Dexamethasone doses will be provided. Equivalent summaries for change in blood laboratory results (follow-up minus baseline) will also be provided.

Adverse events
A full listing of all adverse events (AEs) including serious adverse events (SAEs) will include subject ID, AE description, treatment received, treatment cycle start date, AE onset date, causality, severity, seriousness, expectedness and outcome.

Adverse events from non-randomised patients who only took part in screening will be reported separately.

A corresponding listing will provide details for the subset of AEs classified as serious.

A further listing of suspected unexpected serious adverse reactions (SUSARs) will be provided.

The number of AEs occurring will be summarised by treatment received, as will the number of SAEs occurring. The number and percentage of subjects experiencing at least one AE will be summarised, as will the number and percentage of subjects experiencing at least one SAE.

### 4.7 Description of changes from protocol statistical methods section

The trial protocol proposes that binary or ordinal secondary outcomes would be analysing using a generalised dynamic linear model. This analysis plan considers such an approach unrealistic, given the difficulty of gaining good estimates of the evolution variance and other variance parameters for the dynamic linear model when there is a relatively small number of participants being studied at each Dexamethasone dose in this trial. In this SAP we propose instead that ordinal secondary outcomes are dichotomised, and that these and the other binary secondary outcomes are compared to placebo for each Dexamethasone dose using a simpler Bayesian analysis of the difference in proportions.

## 5. Validation and QC

The outputs from the primary analysis of the primary outcome and the derivation of the primary outcome and the menstrual diary score secondary outcome will be verified by a second statistician.

## 6. References

1. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.

2. West MJ, Harrison PJ. Bayesian forecasting and dynamic models. New York: Springer-Verlag, 1997.

3. Spiegelhalter DJ, Thomas A, Best NG, et al. WinBUGS version 1.4 user manual. MRC Biostatistics Unit, Cambridge, available from www.mrc-bsu.cam.ac.uk/bugs , 2003.

4. Bailey R. Designs for dose-escalation trials with quantitative responses. Stat Med 2009; 28: 3721–3728.

5. Brooks SP, Gelman A. Alternative methods for monitoring convergence of iterative simulations. J Comput Graph Stat 1998; 17: 434–455.