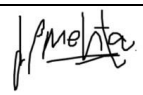
Letrozole or clomifene, with or without metformin, for ovulation induction in women with polycystic ovary syndrome: The LOCI Trial

Trial Registration: ISRCTN11828358

Statistical Analysis Plan

SAP Version Number	Protocol Version Number
1.0	6.8

Name of Author:	Versha Cheed	Role:	Trial Statistician	Affiliation:	BCTU University of Birmingham
Signature of Author:		Date:	12/12/2023		
Name of Reviewer:	Samir Mehta	Role:	Senior Statistician	Affiliation:	BCTU University of Birmingham
Signature of Author:		Date:	12/12/2023		
Name of Chief Investigator:	Professor Arri Coomarasamy	Role:	Chief Investigator	Affiliation:	Institute of Metabolism and Systems Research University of Birmingham, UK
Signature of Chief Investigator:		Date:	12-Dec-2023		
This Statistical Analysis Plan has been reviewed and approved by:					
Name of Approver:	Lee Middleton	Role:	Senior Statistician	Affiliation:	BCTU University of

Signature of Approver:		Date:			Birmingham

Statistical Analysis Plan (SAP) Amendments

SAP version number	SAP section number	Description of and reason for change	Timing of change with respect to interim analysis/ final analysis/ database lock	Blind Reviewer
1.0	n/a	First release	-	n/a

Abbreviations & Definitions	
Abbreviation / Acronym	**Meaning**
APGAR (score)	Appearance, Pulse, Grimace, Activity, and Respiration. (The score is a rapid method for evaluating neonates immediately after birth and in response to resuscitation)
BMI	Body Mass Index
BCTU	Birmingham Clinical Trials Unit
CI	Confidence Interval
CONSORT	Consolidated Standards of Reporting Trials
DMC	Data Monitoring Committee
EQ-5D-5L	European Quality of Life 5 Dimensions 5 Level Version
GnRH	Gonadotropin-releasing hormone
hCG	Human chorionic gonadotropin
HDU	High Dependency Unit
HTA	Health Technology Assessment
ISRCTN	International Standard Randomised Controlled Trial Number
IRR	Incidence Rate Ratio
ITT	Intention to Treat
ITU	Intensive Therapy Unit
IUGR	Intrauterine growth restriction
IVF	In-Vitro Fertilisation
mIU/ml	milli-international units per millilitre
ml	millilitre
MNAR	Missing Not At Random
NIHR	National Institute for Health and Care Research
OHSS	Ovarian Hyperstimulation Syndrome
OI	Ovulation Induction
PCOS	Polycystic Ovary Syndrome
RR	Relative Risk
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SUSAR	Suspected Unexpected Serious Adverse Reaction
TSC	Trial Steering Committee
Term	**Definition**
International Standard Randomised Controlled Trial Number	A clinical trial registry

Protocol	Document that details the rationale, objectives, design, methodology and statistical considerations of the study
Randomisation	The process of assigning trial subjects to intervention or control groups using an element of chance to determine the assignments in order to reduce bias
Statistical Analysis Plan	Pre-specified statistical methodology documented for the trial, either in the protocol or in a separate document

TABLE OF CONTENTS

1. Introduction 8
2. Background and rationale 8
3. Trial objectives 9
4. Trial methods 10
4.1. Trial design 10
4.2. Trial interventions 10
4.3. Primary outcome measure 11
4.4. Secondary outcome measures 11
4.5. Timing of outcome assessments 12
4.6. Randomisation 12
4.7. Sample size 13
4.8. Framework 14
4.9. Interim analyses and stopping guidance 14
4.10. Timing of final analysis 14
4.11. Timing of other analyses 14
4.12. Trial comparisons 15
5. Statistical Principles 15
5.1. Confidence intervals and p-values 15
5.2. Adjustments for multiplicity 15
5.3. Analysis populations 16
5.4. Definition of adherence 17
5.5. Handing protocol deviations 17
5.6. Unblinding 17
6. Trial population 18
6.1. Recruitment 18
6.2. Baseline characteristics 18
7. Intervention(s) 18
7.1. Description of the intervention(s) 18
7.2. Adherence to allocated intervention 18
8. Protocol deviations 19
9. Analysis methods 19
9.1. Covariate adjustment 19
9.2. Distributional assumptions and outlying responses 20
9.3. Handling missing data 20
9.4. Data manipulations 20
9.5. Analysis methods – primary outcome(s) 32
9.6. Analysis methods – secondary outcomes 33
9.7. Analysis methods – other outcomes 34
9.8. Safety data 34
9.9. Planned subgroup analyses 34
9.10. Sensitivity and supportive analyses 35
10. Analysis of sub-randomisations 37
11. Health economic analysis 37
12. Statistical software 37
13. References 37
Appendix A: Deviations from SAP 38
Appendix B: Trial schema 39
Appendix C: Schedule of assessments 40
Appendix D: Template report 41

1. Introduction

This document is the Statistical Analysis Plan (SAP) for the LOCI trial, and should be read in conjunction with the current trial protocol. This SAP details the proposed analyses and presentation of the data for the main paper(s) reporting the results of this trial.

The results reported in these papers will follow the strategy set out here. Subsequent analyses of a more exploratory nature will not be bound by this strategy, though they are expected to follow the broad principles laid down here. The principles are not intended to curtail exploratory analysis (e.g. to decide cut-points for categorisation of continuous variables), nor to prohibit accepted practices (e.g. transformation of data prior to analysis), but they are intended to establish rules that will be followed, as closely as possible, when analysing and reporting data.

Any deviations from this SAP will be described and justified in the final report or publication of the trial (using a table as shown in Appendix A). The analyses will be carried out by an appropriately qualified statistician, who will ensure integrity of the data during their data validation processes.

2. Background and rationale

The background and rationale for the trial are outlined in detail in the protocol. In brief, infertility affects one in six couples, with 25% of infertility being due to anovulation (not releasing eggs from the ovaries). Polycystic ovary syndrome (PCOS) is also very common (approximately 10% of women of reproductive age in the UK have this condition), and it is responsible for 85% of anovulation. Invasive treatments, such as an operation called ovarian diathermy or in-vitro fertilisation (IVF), may overcome anovulation from PCOS, but are associated with significant risks and costs. An effective oral tablet treatment would reduce the need for invasive, risky and costly fertility treatments for women with PCOS, and may improve patient experience.

Eligibility:

Inclusion criteria:

- Women diagnosed with PCOS (according to Rotterdam criteria) and evidence of anovulation (anovulation is defined as irregular cycles lasting <21 or >35 days, OR fewer than 8 periods per year, OR absence of raised serum progesterone >20 nmol/l seven days prior to a period);
- Presentation with infertility or wishing to conceive;
- Male partner with normal sperm count (≥15 million per millilitre (ml)) and progressive

motility ≥32% OR total motility ≥40% in the last 3 years;

- Willing and able to give informed consent;
- Age ≥18 to ≤42 years at randomisation;
- Body Mass Index ≤35 kg/m^2.

Exclusion:

- More than six previous OI (ovulation induction) treatments (cycles) with either letrozole or clomifene in the previous 12 months;
- Intention to continue current use of metformin treatment for ovulation induction or for other indications;
- Metformin use in the previous 14 days;
- Women opting for alternative methods of ovulation induction or treatment (ovarian stimulation with pituitary suppression using gonadotropins with or without pituitary suppression, e.g. with gonadotropin-releasing hormone [GnRH] agonists, antagonists, or progestogens), or performing intrauterine or intracervical insemination*
- Contraindications to letrozole, clomifene, metformin use and/or pregnancy (see section 7.3 in protocol for full details on contraindications);
- Previous participation in the LOCI trial.

* hCG (Human Chorionic Gonadotropin) trigger injections were listed in the original exclusion criteria but have been removed because the use of ovulation trigger is often considered routine care in participating centres and there is no interaction with the OI treatment.

3. Trial objectives

The primary objectives are:

To test the hypotheses that in women with PCOS and infertility, letrozole versus clomifene, metformin versus placebo, and letrozole plus metformin versus clomifene plus metformin increases the live birth rate (≥ 34 weeks of gestation) by at least 10% (absolute difference). This will involve the following three primary comparisons:

Letrozole vs clomifene

To compare the effectiveness of letrozole vs clomifene in women with anovulatory PCOS and infertility on live birth rate (≥34 weeks of gestation).

Metformin vs placebo
To compare the effectiveness of metformin vs placebo in women with anovulatory PCOS and infertility undergoing ovulation induction on live birth rate (≥34 weeks of gestation).

Letrozole plus metformin vs clomifene plus metformin
To compare the effectiveness of letrozole plus metformin vs clomifene plus metformin in women with anovulatory PCOS and infertility on live birth rate (≥34 weeks of gestation).

4. Trial methods

4.1. Trial design

LOCI is a 2x2 factorial randomised, double-blind, placebo-controlled multicentre superiority trial of investigational medicinal products, with health economic evaluation and a six-month internal pilot to ensure ability to recruit and randomise. See Appendix B for trial schema.

Participants will be recruited from gynaecology departments and and/or fertility centres in the United Kingdom.

Participants, investigators, research midwives/nurses, laboratory outcome assessors and other attending clinicians will remain blind to the trial treatment allocation throughout the duration of the trial.

4.2. Trial interventions

Letrozole for 5 days starting on day 2 or 3 of the menstrual cycle, plus metformin or placebo daily. Initial letrozole dose will be 2.5 mg daily and may be increased to a maximum dose of 7.5 mg daily until ovulation is confirmed, for a maximum of 6 treatment cycles. The comparison will be clomifene for 5 days starting on day 2 or 3 of the menstrual cycle, plus metformin or placebo daily. Initial clomifene dose will be 50 mg daily and may be increased to a maximum dose of 150 mg in a similar way to letrozole, for a maximum of 6 treatment cycles.

The trial supports and documents the local clinical prescription recommended for each participant. In addition, the use of a human chorionic gonadotrophin (hCG) trigger injection following ovulation induction will be permitted and appropriately documented for participants requiring it. The maximum dose of metformin will be 1500 mg daily and continued until 14 completed weeks of pregnancy or until the end of the 6 treatment cycles.

4.3. Primary outcome measure

The primary outcome is live birth at and beyond 34 completed weeks of gestation, as a proportion of all women randomised.

4.4. Secondary outcome measures

The secondary outcomes are as follows:

Secondary outcomes:

- Treatment outcomes: Ovulation rate, time from randomisation to pregnancy, number of ovulation induction cycles required for pregnancy, number of ovulation induction cycles to live birth ≥24 weeks.
- Pregnancy end outcomes: Ongoing pregnancy at 12 weeks (range 11^{+0} to 14^{+0} weeks) of gestation, pregnancy loss (defined as pregnancy loss <24 weeks of gestation), termination, stillbirth, molar pregnancy, pregnancy of unknown location, ectopic pregnancy, multiple live births, gestational age at live birth.
- Where live birth ≥24 weeks: time from conception to delivery (gestational age at live birth), gestational age <28/<32/<37 weeks, singleton live births at and beyond 34 completed weeks of gestation, live births at and beyond 37 completed weeks of gestation, mode of birth (unassisted vaginal, instrumental vaginal, elective caesarean section, emergency caesarean section, vaginal breech birth, other), birth weight, APGAR score (at 1, 5 and 10 minutes) and APGAR score <7 out of 10 (at 1, 5 and 10 minutes).
- Antenatal outcomes: antepartum haemorrhage, pregnancy-induced hypertension, pre-eclampsia, intrahepatic cholestasis of pregnancy, preterm (<37 weeks) pre-labour rupture of membranes, gestational diabetes (other complications will be tabulated but not formally analysed)
- Intrapartum outcomes: chorioamnionitis, fetal growth restriction, macrosomia (other complications will be tabulated but not formally analysed)
- Post-partum outcomes: haemorrhage (other complications will be tabulated but not formally analysed)
- Maternal outcomes: admission to high dependency unit (HDU), admission to intensive therapy unit (ITU) (other complications will be tabulated but not formally analysed)
- Neonatal outcomes: discharge to hospital, early infection, retinopathy of prematurity, necrotising enterocolitis, intraventricular haemorrhage, respiratory distress syndrome, ventilation or oxygen support (other complications will be tabulated but not formally analysed)

- Survival at 28 days of neonatal life
- Health economic evaluation: Hospital resource use and EQ-5D-5L questionnaire (this analysis is documented the health economics analysis plan and not performed as part of this SAP).

Safety outcomes:

- Neonatal congenital or chromosomal abnormalities
- Maternal adverse events (tabulated but not formally analysed)
- Multiple pregnancies
- Ovarian hyperstimulation syndrome (OHSS)
- Serious adverse events

Side effects attributed to clomifene/letrozole and metformin/placebo (nausea or vomiting/abdominal pain/headache/joint pain/fatigue/mood changes/breast discomfort/visual symptoms/hot flushes) will also be tabulated during the ovulation induction and pregnancy phases.

4.5. Timing of outcome assessments

The schedule of trial procedures and outcome assessments are given in Appendix C.

4.6. Randomisation

Participants will be randomised online via a secure internet facility at the level of the individual in a 1:1 ratio to either letrozole or clomifene and at the same time randomised to metformin or placebo.

A minimisation algorithm will be used within the online randomisation system to ensure balance in the treatment allocations (both randomisations will occur simultaneously, effectively resulting in four allocation groups) over the following variables:

- maternal age (<35 and ≥35 years);
- body mass index (<30 and ≥30 kg/m^2);
- any previous pregnancy (yes and no);
- previous exposure to either clomifene or letrozole;
- any menstrual periods in the preceding 6 months (yes and no);
- randomising centre

A 'random element' will be included in the minimisation algorithm, so that each participant has

a probability (unspecified here) of being randomised to the opposite treatment that they would have otherwise received.

Full details of the randomisation specification will be stored in a confidential document at BCTU.

4.7. Sample size

Original sample size

Just under two thousand participants (n = 1,992) will provide greater than 99% power (at p=0.05) to detect differences in the letrozole vs clomifene comparison (assuming rates of 39% vs 29% respectively for rate of live birth ≥34 weeks) and the same power to detect differences in the metformin vs placebo comparison (also assuming rates of 39% vs 29%, respectively). Adjusting for a worst-case scenario of 5% attrition, the total number required will be 2,100 participants. The high rate of power (>99%) has been chosen to ensure we will have 90% power (p=0.05 and 1,050 participants) to answer the further question of letrozole plus metformin vs clomifene plus metformin (the latter comparison assuming rates of 44% and 34%, respectively).

The basis for the proportions used in these calculations is provided in the next paragraph.

The live birth rates used in the sample size calculations were taken from our systematic review (see protocol Table 8). We used the estimate in the clomifene alone group (24%) as the base estimate for these calculations as this was by far the largest group, involving 36 trial groups and 2,299 women; it was also the most conservative estimate (lowest rate), which was important as information was only available on the rate of pregnancy and not live birth rate. A 10% absolute increase was identified in our clinician survey as minimally important, so we chose this as the difference we wanted to detect when comparing clomifene with letrozole or with the addition of metformin (i.e., both increased to 34%). We assumed using both letrozole and metformin would have an additive effect, i.e. increased to 44%. Overall, this amounted to an assumption of 29% versus 39% in the two main comparisons (letrozole vs clomifene and metformin vs placebo) when we take into account the factorial design (i.e. the overall rates at the margins). No interaction of the effect is assumed in these calculations as the biological mechanism of these agents is considered sufficiently different and, to our knowledge, is unlikely to have a pronounced effect on outcome.

Revised sample size

Following review by the funder (NIHR HTA) and due to the impact of the COVID-19 pandemic on participant accrual, the sample size target was revised on 3 February 2023 to aim for at least 80% power within the letrozole plus metformin vs clomifene plus metformin comparison, rather than the 90% power originally described. 800 participants (with primary outcome data

gained on 760) within this comparison would provide 81% power (p=0.05) to detect a difference of 10% between groups (the same assumptions of 44% and 34% in the two groups as previously stated).

Due to the factorial design, this agreement would mean a sample of 1600 participants (with primary outcome data gained on 1520 assuming 5% loss to follow-up) within the two main comparisons (letrozole vs clomifene and metformin vs placebo). This number of participants would provide 98.5% power (at p=0.05) to detect a difference of 10% between groups (the same assumption of 39% vs 29% in both comparisons as previously stated).

4.8. Framework

The objective of the trial is to test the superiority of one intervention to another in the three trial comparisons (see section 4.12 for list of these comparisons).

The null hypothesis is that there is no difference in the proportion of women randomised who experience a live birth at and beyond 34 weeks of gestation between the intervention groups. The alternative hypothesis is that there is a difference between the groups.

4.9. Interim analyses and stopping guidance

A separate Data Monitoring Committee (DMC) reporting template was drafted and agreed by the DMC including an agreement on which outcomes were to be reported at interim analyses. The statistical methods stated in this SAP was followed for the outcomes included in the DMC report, where possible.

Interim presentation of safety and efficacy data to the independent DMC took place during the study. The committee met prior to study commencement to agree the manner and timing of such analyses including the analysis of the primary and major secondary outcomes and full assessment of safety (SAEs) at least at annual intervals. Criteria for stopping or modifying the study based on this information was ratified by the DMC and therefore no formal stopping rules were applied.

4.10. Timing of final analysis

The final analysis for the study will occur once all participants have completed all assessments and corresponding outcome data has been entered onto the study database and validated as being ready for analysis.
This is provided that the trial has not been stopped early for any reason (e.g. DMC advice or funding body request).

4.11. Timing of other analyses

Not applicable.

4.12. Trial comparisons

The three primary comparisons will be composed of:

A) those randomised to letrozole versus those randomised to clomifene;

B) those randomised to metformin versus those randomised to placebo; and

C) those randomised to letrozole plus metformin versus those randomised to clomifene plus metformin.

Analyses will be conducted separately for each comparison. Unless stated otherwise, the methods detailed here will be applicable to each comparison.

In addition, there will be a further two exploratory comparisons:

D) letrozole plus metformin versus letrozole
E) clomifene plus metformin versus clomifene

The five comparisons are considered distinct questions and will be reported separately; hence no adjustment for multiple testing across comparisons is proposed.

5. Statistical Principles

5.1. Confidence intervals and p-values

For all primary and secondary outcomes, estimates of differences between groups will be presented with measures of uncertainty (95% confidence intervals), unless otherwise stated. For the primary outcome, a p-value will be produced, with statistical significance considered at the 5% level. A 95% confidence interval will be utilised. P-values will be presented for the safety outcomes (see section 5.2).

5.2. Adjustments for multiplicity

As detailed in section 4.12, the trial will have three main separate comparisons plus two exploratory comparisons. These comparisons are considered distinct questions and will be reported separately. Therefore, no adjustments for multiple testing are proposed from this perspective.

Safety outcomes will be subject to statistical testing without adjustment for multiple testing as adjustment for multiplicity is counterproductive for considerations of safety[1].

5.3. Analysis populations

Participants will be analysed in the treatment group to which they were randomised, regardless of treatment compliance. Analysis populations will be as follows:

- Primary outcome, live birth at or beyond 34 weeks: all randomised participants. Analyses will also be produced using women who achieved pregnancy, defined as a positive pregnancy test (urine or serum beta-hCG test >10 mIU/ml), as the denominator but this will be treated as supportive evidence only)
- Secondary outcomes:
    - (Treatment outcomes) Ovulation rate, time from randomisation to pregnancy: all randomised participants
    - (Treatment outcomes) Number of ovulation induction cycles required for pregnancy: women who achieved pregnancy (defined as a positive pregnancy test (urine or serum beta-hCG test >10 mIU/ml))
    - (Pregnancy end outcomes) Ongoing pregnancy at 12 weeks (range 11^{+0} to 14^{+0} weeks) of gestation, pregnancy loss (defined as pregnancy loss <24 weeks of gestation), termination, stillbirth, molar pregnancy, pregnancy of unknown location, ectopic pregnancy, multiple live births: all randomised participants (analyses will also be produced using women who achieved pregnancy, defined as a positive urine or serum beta-hCG test >10 mIU/ml, as the denominator but this will be treated as supportive evidence only)
    - Time from conception to delivery (gestational age at live birth), gestational age <28/<32/<37 weeks, singleton live births at and beyond 34 completed weeks of gestation, live births at and beyond 37 completed weeks of gestation, mode of birth (unassisted vaginal, instrumental vaginal, elective caesarean section, emergency caesarean section, vaginal breech birth, other), number of ovulation induction cycles to live birth ≥24 weeks : women with live birth ≥24 weeks
    - Birth weight, APGAR score (at 1, 5 and 10 minutes) and APGAR score <7 out of 10 (at 1, 5 and 10 minutes): babies born alive ≥24 weeks
    - Antenatal outcomes: women achieving pregnancy
    - Intrapartum outcomes: women with live birth≥24 weeks
    - Post-partum outcomes: women with live birth≥24 weeks
    - Maternal outcomes: all randomised participants

- Neonatal outcomes: babies born alive ≥24 weeks.
- Survival at 28 days of neonatal life: babies born alive ≥24 weeks

- Safety outcomes:
  - Neonatal congenital or chromosomal abnormalities: babies born alive ≥24 weeks
  - Maternal adverse events (tabulated but not formally analysed), multiple pregnancies, ovarian hyperstimulation syndrome (OHSS), serious adverse events (SAEs): all randomised participants
  - Side effects attributed to clomifene/letrozole and metformin/placebo: all randomised participants

5.4. Definition of adherence

Adherence will be monitored by tablet counting in the first instance with the participants' honest assessment of their drug compliance if tablet counting is not possible (recorded as a percentage). For clomifene/letrozole, this will be reported for the OI phase; for metformin/placebo, this will be reported separately for the OI phase and for the pregnancy phase (participants will need to be adherent in both phases to be considered adherent overall in any analysis).

We will define and summarise adherence in the following categories: number that received ≥75% randomised allocation; number that received <75% randomised allocation; unknown compliance information.

5.5. Handing protocol deviations

A protocol deviation is defined as a failure to adhere to the protocol such as errors in applying the inclusion/exclusion criteria, the incorrect intervention being given, incorrect data being collected or measured, follow-up visits outside the visit window or missed follow-up visits. We will apply a strict definition of the ITT (intention to treat) principle and will include all participants as per the ITT population described in section 5.3 in the analysis, in some form, regardless of deviation from the protocol.[2] This does not include those participants who have specifically withdrawn consent for the use of their data in the first instance; however, these outcomes will be explored as per other missing responses.

5.6. Unblinding

In this study, the unblinding of the Trial Statistician to the intervention code will take place once the database is locked for final analysis, or the data has been downloaded for a planned analysis, unless the DMC request that they review the interim data with knowledge of the intervention groups or the DMC request to be unblinded at an interim analysis.

6. Trial population

6.1. Recruitment

A flow diagram (as recommended by CONSORT[3]) will be produced to describe the participant flow through each stage of the trial. This will include information on the number (with reasons) of losses to follow-up (drop-outs and withdrawals) over the course of the trial. A template for reporting will be given in a separate template report.

6.2. Baseline characteristics

The trial population will be tabulated as per the template report document. Categorical data will be summarised by number of participants, counts and percentages. Continuous data will be summarised by the number of participants, mean and standard deviation if deemed to be normally distributed or number of participants, median and interquartile range if data are skewed, and ranges if appropriate. Tests of statistical significance will not be undertaken, nor confidence intervals presented.[4]

Baseline characteristics will be reported by Letrozole versus clomifene and metformin versus placebo as well as the four treatment arms in the following way:

A) letrozole plus metformin
B) clomifene plus metformin
C) letrozole plus placebo
D) clomifene plus placebo

7. Intervention(s)

7.1. Description of the intervention(s)

A template for reporting information on the intervention(s) will be given in a separate template report.

7.2. Adherence to allocated intervention

A cross-tabulation of allocated intervention by the adherence categories stated in section 5.4 will be produced (proportions and percentages). A template for reporting adherence will be given in a separate template report.

8. Protocol deviations

Frequencies and percentages by group will be tabulated for the protocol deviations as per separate template report.

9. Analysis methods

Intervention groups will be compared using regression models to adjust for all covariates as specified in section 9.1, where possible.

Output for the analysis will be reported via the comparisons as mentioned in section 4.12 Trial comparisons.

9.1. Covariate adjustment

Both group allocations (as a result of the double randomisation required by the factorial design) will be included in all analysis models (for comparisons A and B only as for C, D and E this would not be possible as we are only looking at one of the randomisations. See section 4.12 Trial comparisons for details on the comparisons); no interaction of treatment effect is assumed in the first instance but this will be explored through supportive analyses (section 9.10).

In the first instance, intervention effects between groups for all outcomes will be adjusted for the minimisation parameters listed in section 4.6 in the same form as entered into the minimisation algorithm, i.e. categorical variables such as age will remain as such in the model, rather than converting them to the continuous form (as this creates a problem downstream in terms of presentation of the subgroup analyses). Randomising centre will be included as a random intercept in the model, and all other factors as fixed effects (with the exception of time to event outcomes where centre will be regarded as a shared frailty).

A mixed effects binomial regression model will be used to calculate the risk difference (identity link) and relative risk (log link) with 95% confidence intervals for the primary outcome (live birth ≥ 34 weeks), adjusting for the randomisation variables listed in section 4.6. If the binomial model fails to converge then, in the first instance, a Poisson regression model with randomising centre as a random effect and robust standard errors will be used to estimate the relative risk. In the case that the Poisson regression model still does not converge, randomising centre will be removed and the log-binomial model will be refit for just fixed effects. If this also fails, covariates will be dropped as described in section 9.1. A Poisson model will not estimate the risk difference, so if the initial model fails to converge, the same process of variable removal will be followed, within a log-binomial model framework. The p-value relating to the intervention group parameter as generated by the model estimating the relative risk will be presented.

9.2. Distributional assumptions and outlying responses

Distributional assumptions (e.g. normality of regression residuals for continuous outcomes) will be assessed visually prior to analysis; although in the first instance the proposed primary method of estimation in this analysis plan will be followed. If responses are considered to be particularly skewed and/or distributional assumptions violated, other analytical options will be considered such at the transformation of responses prior to analysis (e.g. log transformation) or the use of medians and interquartile ranges alongside unadjusted differences in medians using bootstrapping methods (repetition=1000, seed=271123).

If extreme values are apparent and considered to be affecting the integrity of the analysis, a sensitivity analysis consisting of removing the outlying response(s) and repeating the analysis will be performed. Output from these analyses, if performed, will be described and presented alongside the original analysis (or included, e.g. in appendices) with the excluded values clearly labelled. See section 9.10 for further details regarding sensitivity analyses.

9.3. Handling missing data

In the first instance, analyses will be completed on received data only with every effort made to follow-up participants to minimise any potential for bias. To examine the possible impact of missing data on the results, and to make sure we are complying with the intention-to-treat principle, sensitivity analysis will be performed on the primary outcome measure.[6] See section 9.10 for further details regarding sensitivity analyses.

9.4. Data manipulations

The trial statistician will derive all responses from the raw data recorded in the database:

Other:

- **Maternal age at randomisation (years)**

(Date of randomisation-date of birth)/365.25.

Outcome Measures:

Primary Outcome

- **Live births at and beyond 34 completed weeks of gestation, as a proportion of all women randomised.**

CRF: LOCI: 07. Pregnancy Outcome Form

Question:

- **'**Was the baby born alive?' = Yes;

AND

- 'Gestation at pregnancy end' = ≥34 weeks

Exclude participants where: pregnancy outcome is miscarriage, ectopic, still birth or termination and any participants who had a live birth <34 weeks.
Participants will be censored if they withdraw or are lost to follow-up before pregnancy end. They will be censored at the point of withdrawal/lost to follow-up.
Secondary outcomes

Treatment outcomes:

- **Number of ovulation induction cycles to live birth**

CRF: LOCI: 04. End of Ovulation Induction/Pregnancy Confirmation Form
Question:

- **'**Number of completed OI treatment cycles?'
- 'Did the participant become pregnant within 240 days (8 months) of randomisation?' = Yes

CRF: LOCI: 07. Pregnancy Outcome Form
Question:

- Was the baby born alive? = Yes

- **No. of women who ovulated**

CRF: LOCI: 03. OI Drug Compliance and Ovulation Tracking Form
Question:

- 'Was an LH surge confirmed?'= Yes

OR

- 'Progesterone concentration' = >20 nmol/l

OR

USS follicular tracking

- 'Was ovulation demonstrated?' = Yes

If any of these criteria are fulfilled at any one of the treatment cycles, then this would indicate that ovulation has occurred.

- **Ovulation rate (total number of ovulations / total number of ovulation induction cycles)**

TOTAL NUMBER OF OVULATIONS:
CRF: LOCI: 03. OI Drug Compliance and Ovulation Tracking Form
Question:
- 'Was an LH surge confirmed?'= Yes

OR
- 'Progesterone concentration' = >20 nmol/l

OR
USS follicular tracking
- 'Was ovulation demonstrated?' = Yes

Obtain total number of ovulations by totalling each 03. OI Drug Compliance and Ovulation Tracking Form per cycle where ovulation has occurred.

TOTAL NUMBER OF OVULATION INDUCTION CYCLES
CRF: LOCI: 04. End of Ovulation Induction/Pregnancy Confirmation form
Question:
- 'Number of completed OI treatment cycles?'

OVULATION RATE = Total number of ovulations / Total number of ovulation induction cycles

- **Time from randomisation to pregnancy**

CRF: LOCI: 04. End of Ovulation Induction/Pregnancy Confirmation Form
Question:
  - 'Did the participant become pregnant within 240 days (8 months) of randomisation?' = Yes;

AND
  - 'If yes, date of positive pregnancy test:' = Use date

Time to pregnancy (weeks) = ((Date of positive pregnancy test)-(Randomisation Date))/7

- **Number of ovulation induction cycles required for pregnancy**

CRF: LOCI: 04. End of Ovulation Induction/Pregnancy Confirmation Form
Question:
  - 'Number of completed OI treatment cycles?'
  - 'Did the participant become pregnant within 240 days (8 months) of randomisation?' = Yes

Pregnancy end outcomes:

- **Ongoing pregnancy at 12 weeks (range 11 to 14 weeks) of gestation**

CRF: LOCI: 05. Pregnancy Booking Scan Form
Question:
- 'Did the woman undergo a booking scan at 11-14 weeks?' = Yes

AND
- 'Gestation according to ultrasound scan:' = range 11 to 14 weeks
- 'Fetal heart beat present' = For each fetus

- **Pregnancy loss (defined as pregnancy loss before 24 weeks of gestation)**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- **'**Was the baby born alive?' = No;

AND
- 'Gestation at pregnancy end' = <24 weeks;

AND
- 'If no, what was the pregnancy outcome:' = Miscarriage, Ectopic pregnancy, Molar pregnancy, Pregnancy of unknown location, Termination

- **Twin live births**

CRF: LOCI: 05. Pregnancy Booking Scan Form
Question:
- 'Did the woman undergo a booking scan at 11-14 weeks?' = Yes

AND
- 'Fetal heart beat present' = For each fetus

AND
- 'If more than 1 fetus, type of multiple pregnancy:' = List type

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
CHECK FOR EACH FORM FOR EACH BABY
- 'Was the baby born alive?' = Yes

- **Gestational age at live birth**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = Use gestation

Where live birth ≥24 weeks:

- **Time from conception to delivery (gestational age)**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = ≥24 weeks

AND
- 'Gestation at pregnancy end:' = Use Gestation

- **Gestational age <28/<32/<37 weeks**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = ≥24 weeks

Categorise Gestational age at pregnancy end by: <28/<32/<37 weeks.

- **Singleton live births at and beyond 34 completed weeks of gestation**

CRF: LOCI: 05. Pregnancy Booking Scan Form
Question:
- 'Number of fetus(es)?' = 1

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = ≥34 weeks

- **Live births at and beyond 37 completed weeks of gestation**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = ≥37 weeks

- **Mode of birth (unassisted vaginal, instrumental vaginal, elective caesarean section, emergency caesarean section, vaginal breech birth, other)**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = ≥24 weeks

AND
- 'Mode of birth' = unassisted vaginal, instrumental vaginal, elective caesarean section, emergency caesarean section, vaginal breech birth, other

- **Birth weight**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = ≥24 weeks

AND
- **'**Birth weight'

- **APGAR score <7 out of 10 (at 1, 5 and 10 minutes)**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Gestation at pregnancy end:' = ≥24 weeks

AND
- 'Apgar at 1 minute:' = <7
- 'Apgar at 5 minutes:' = <7
- 'Apgar at 10 minutes:' = <7

Antenatal outcomes:

- **Antepartum haemorrhage**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Antepartum haemorrhage?' = Yes

CRF: LOCI: 08. Maternal Outcome Form
Question:
- 'Did the woman suffer antepartum haemorrhage?' = Yes

- **Pregnancy-induced hypertension**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Hypertension / pre-eclampsia?' = Yes

CRF: LOCI: 08. Maternal Outcome Form
Question:
- 'Did the woman suffer pregnancy induced hypertension?' = Yes

- **Pre-eclampsia**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Hypertension / pre-eclampsia?' = Yes

CRF: LOCI: 08. Maternal Outcome Form
Question:
- 'Did the woman suffer pre-eclampsia?' = Yes

- **Obstetric cholestasis**

CRF: LOCI: 08. Maternal Outcome Form
Question:
- 'Did the woman suffer obstetric cholestasis?' = Yes

- **Preterm (<37 weeks) pre-labour rupture of membranes**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
  - 'Was the baby born alive?' = Yes

AND
  - 'Gestation at pregnancy end:' = <37 weeks

AND
  - **'**Pre-PROM?' = Yes

CRF: LOCI: 08. Maternal Outcome Form
Question:
  - 'Did the woman suffer preterm (<37 weeks) pre-labour rupture of membranes?' = Yes

- **Gestational diabetes**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
  - 'Was the baby born alive?' = Yes

AND
  - **'**Gestational diabetes mellitus / type I diabetes / type II diabetes?' = Yes

CRF: LOCI: 08. Maternal Outcome Form
Question:
  - 'Did the woman suffer gestational diabetes, diagnosed by GTT?' = Yes

<u>Intrapartum outcomes:</u>

- **Chorioamnionitis**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
  - 'Was the baby born alive?' = Yes

AND
  - 'Were there any intrapartum complications?' = Yes

AND
  - **'**Chorioamnionitis?' = Yes

- **Fetal growth restriction**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
  - Suspected IUGR = Yes

- **Macrosomia**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
  - 'Was the baby born alive?' = Yes

AND
  - **'**Fetal macrosomia?' = Yes

Post-partum outcomes:

- **Haemorrhage**

CRF: LOCI: 08. Maternal Outcome Form
Question:
  - 'Did the woman suffer any other complications?'= Yes

AND
  - 'Haemorrhage?' = Yes

Maternal outcomes:

- **Admission to high dependency unit (HDU)**

CRF: LOCI: 08. Maternal Outcome Form
Question:
  - 'Was the participant admitted to HDU (level 2 care) prior to post-natal discharge?'= Yes

- **Admission to intensive therapy unit (ITU)**

CRF: LOCI: 08. Maternal Outcome Form
Question:
  - 'Was the participant admitted to ITU (level 3 care) prior to post-natal discharge?'= Yes

Neonatal outcomes:

- **Discharge to hospital**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Was the baby discharged from hospital? = Yes

AND
- 'Discharged to' = 'Other hospital'

- **Early infection**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Early infection'= Yes

- **Retinopathy of prematurity**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Retinopathy of prematurity requiring treatment?' = Yes

- **Necrotising enterocolitis**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Necrotising enterocolitis?' = Yes

- **Intraventricular haemorrhage**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND

- 'Grade 3 or Grade 4 intraventricular haemorrhage?' = Yes

- **Respiratory distress syndrome**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Respiratory distress syndrome?'= Yes

- **Ventilation or oxygen support**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'Ventilation or oxygen support?'= Yes

- **Survival at 28 days of neonatal life.**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Was the baby born alive?' = Yes

AND
- 'What was the outcome at 28 days post-birth?' = Alive

<u>Safety outcomes:</u>

- **Neonatal congenital or chromosomal abnormalities**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
- 'Congenital or genetic anomalies?' = Yes

AND
- 'If yes' = List details

- **Multiple pregnancies**

CRF: LOCI: 05. Pregnancy Booking Scan Form
Question:
- 'Number of fetus(es)?' = >1

- **Ectopic pregnancies**

CRF: LOCI: 07. Pregnancy Outcome Form
Question:
  - 'Was the baby born alive?' = No

AND
  - 'If no, what was the pregnancy outcome'= Ectopic

- **Ovarian hyperstimulation syndrome (OHSS)**

CRF: LOCI: 08. Maternal Outcome Form
Question:
  - 'Did the woman suffer ovarian hyperstimulation syndrome (OHSS)?' = Yes

AND
  - 'If yes, please specify:' = Mild/Moderate/Severe

- **Serious adverse events**

CRF: LOCI: 10. Serious Adverse Event Form

Subgroup Analyses:

- Androgen excess confirmed clinically or through laboratory testing (yes and no) confirmed as yes if:
  - 'Does the woman suffer from acne?' = Yes (LOCI: 02. Baseline Medical Details Form)

OR
  - 'Does the woman suffer from hirsutism (unwanted male pattern hair growth on the face or body)?' = Yes (LOCI: 02. Baseline Medical Details Form)

OR
  - 'Does the woman suffer from alopecia = Yes (excessive hair loss)?' (LOCI: 02. Baseline Medical Details Form)

OR
  - 'Has the participant previously had Sex Hormone Binding Globulin (SHBG) concentrations measured? = Yes AND 'Sex Hormone Binding Globulin (SHBG)' <19.0 nmol/L (LOCI: 02. Baseline Medical Details Form)

OR
  - High levels of testosterone ' Has the participant previously had free testosterone concentrations measured?' = Yes AND ' Testosterone' >2.0 ng/Ml

- Insulin resistance confirmed clinically or laboratory testing (yes and no) confirmed as yes if:
  - 'Does the woman suffer from type 2 diabetes?' = Yes (LOCI: 02. Baseline Medical

Details Form)

OR

- 'Other relevant medical history?' = Yes AND 'Diabetes' = Yes (LOCI: 02. Baseline Medical Details Form)

OR

- 'Onset of labour' = Induced AND 'Gestational diabetes mellitus / type I diabetes / type II diabetes?' = Yes AND 'If yes to Gestational diabetes, please provide GTT:' >7.8 mmol/L (LOCI: 07. Pregnancy Outcome Form)

OR

- 'Did the woman suffer gestational diabetes, diagnosed by GTT?' = Yes (LOCI: 08. Maternal Outcome Form)

OR

- 'Has the participant previously had HbA1c concentrations measured?' = Yes AND 'HbA1c' <48.0 mmol/mol (LOCI: 02. Baseline Medical Details Form)

9.5. Analysis methods – primary outcome(s)

A template for reporting the primary outcome will be given in a separate template report.

The primary outcome is the proportion of women randomised who experience a live birth at and beyond 34 weeks of gestation. The denominator of this proportion will be all women randomised and the numerator will be those women who have conceived during the six ovulation induction cycles (maximum 240 days) and have gone on to have a live birth at or beyond 34 weeks.
A mixed effects binomial regression model will be used to calculate the risk difference (identity link) and relative risk (log link) with 95% confidence intervals for the primary outcome (live birth ≥ 34 weeks), adjusting for the randomisation variables listed in section 4.6. If this primary method doesn't converge, the plan detailed in section 9.1 will be followed.

Statistical significance of the treatment group parameter will be determined (p-value generated) through examination of the associated chi-squared statistic (this will be obtained from the log-binomial model which produces risk ratios).

Summary data of the primary outcome in the form of frequencies and percentages will be produced that looks at the primary outcome for all the combinations of treatments (combinations A-D as per section 6.2 Baseline characteristics).

In the event of babies from a multiple pregnancy having different pregnancy outcomes, if at least one of the babies had a live birth ≥34 weeks then this will count as a live birth ≥34 weeks overall. If all those babies within that multiple pregnancy have the same pregnancy outcome (e.g. all are live births) then this will only be counted as one outcome (event).

See section 9.1 for covariate adjustment and model convergence.

9.6. Analysis methods – secondary outcomes

A template for reporting the secondary outcomes will be given in a separate template report. See section 9.1 for covariate adjustment and model convergence.

Dichotomous outcomes (e.g. pregnancy loss and APGAR score <7) will be analysed using the same regression approach as the primary outcome, using the denominators detailed in section 5.3. Pregnancy events will be limited to those who conceived during the six ovulation induction cycles (maximum 240 days). For pregnancy end outcomes, in the event of babies from a multiple pregnancy having different outcomes, for example one live birth and one miscarriage, both the events will be counted in the separate categories (i.e. they will contribute to both a live birth event and a miscarriage event.

Mixed-effects Poisson regression will be used for count data (number of ovulation induction cycles to live birth) with an Incidence Rate Ratio (IRR) and confidence interval generated (if there is evidence of overdispersion then a negative-binomial model will be fitted instead). This outcome will be reported as count data (i.e. 1, 2, 3, 4, 5, and 6).

For continuous secondary outcome measures (e.g. birthweight and gestational age) means and standard deviations will be reported alongside adjusted mean differences (with confidence intervals) estimated using a mixed effects linear regression model.

A Cox Proportional Hazard (PH) model (provided the assumptions of proportionality are met) will be utilised for time to event data (time from randomisation to pregnancy). Data will be summarised with medians and interquartile ranges and estimates of difference presented as adjusted hazard ratios (with confidence intervals). A Kaplan Meier plot will be produced to assess the data visually. In this model, randomising centre will be regarded as a shared frailty.

Mode of birth will be analysed as unassisted vaginal versus any other category; all categories of data will be presented.

For birthweight, babies from multiple births will be included as separate babies. To allow for the clustering at the mother level a robust variance estimator clustered by mother will be included in the model. If this does not work (e.g. failure to converge) then analysis on the mother's level will be used instead and the outcome interpreted as such.

For the following neonatal outcomes the denominator will be number of mothers and not the number of babies. In the event of multiple babies from the same pregnancy having different outcomes, e.g. if one twin had an early infection, and the other did not, this event will still be classified as having occurred:

- Discharge to hospital
- Early infection
- Retinopathy of prematurity
- Necrotising enterocolitis
- Intraventricular haemorrhage
- Respiratory distress syndrome
- Ventilation or oxygen support

Ovulation rate will be converted to a percentage score (Total number of ovulations / Total number of ovulation induction cycles) and so will be treated as continuous proportion data. Data will be summarised separately by allocation group using mean and standard deviation. Since this is a proportion, the data will be bounded between the values [0, 1] and so a fractional regression model using a logit link will be fitted with cluster-robust standard errors. Marginal effects will be derived as mean differences in proportions between the groups and the associated 95% confidence intervals.

9.7. Analysis methods – other outcomes

Any data that does not form a pre-specified outcome will be presented using simple summary statistics by intervention group (i.e. numbers and percentages for binary data and means (or medians) and standard deviations (or inter-quartile ranges) for continuous normal (or non-normal) data.

9.8. Safety data

The number and percentage of participants experiencing any serious adverse events (SAEs) and suspected unexpected serious adverse reactions (SUSARs) will be presented by intervention group. Statistical significance will be determined (p-value generated) through examination of the associated chi-squared statistic. The total number of SAEs and SUSARs in each group will also be given along with a descriptive table of the details of the events. For side effect data, summary statistics (e.g. frequency and percentage) will be detailed only.

A template for reporting this safety data will be given in a separate template report.

9.9. Planned subgroup analyses

Subgroup analyses will be restricted to the primary outcome only.
BMI (<30 and ≥30 kg/m^2) is the single key subgroup of interest (i.e. we propose to be able to draw firm conclusion about any differential effect with respect to this variable only).

The following subgroup analyses will be considered exploratory:

From the minimisation algorithm -

- Maternal age (<35 and ≥35 years);
- Body mass index (<30 and ≥30 kg/m^2);
- Any previous pregnancy (yes and no);
- Previous exposure to either clomifene or letrozole;
- Any menstrual periods in the preceding 6 months (yes and no);

As well as the following variables –

- Ethnicity (White/South Asian/Black/Other)
- History of previous miscarriage (0/1-2/≥3)
- Androgen excess confirmed clinically or through laboratory testing (yes and no) see section 9.4 Data manipulations on derivation details
- Insulin resistance confirmed clinically or laboratory testing (yes and no) see section 9.4 Data manipulations on derivation details.

Tests for interaction will be performed by including the treatment group by subgroup interaction parameter in the statistical model. A p-value for this parameter will be produced. Given there is only one key subgroup of interest, no adjustments will be made for multiple testing on this p-value.

9.10. Sensitivity and supportive analyses

Sensitivity analyses will be limited to the primary outcome and will consist of:

- Per-protocol analysis; this will comprise the adherent population as described in sections 5.4;
- No interaction of treatment effect (as a result of the two trial allocations) is assumed. To explore this assumption, a treatment group interaction parameter (randomisation 1 by randomisation 2) will be added to the base model for trial comparisons A and B only as comparisons C-E only include one of the randomisations[4.13]). A p-value for this interaction parameter will be produced and examined as well as the interaction coefficient and 95% CI. If this parameter is significant, then treatment effect estimates within randomisation 2 allocation (for randomisation 1) and randomisation 1 allocation (for randomisation 2) will be explored as per the subgroup analysis methods described in section 9.9. It should be acknowledged that any test for interaction will be underpowered to detect a significant effect.
- An analysis to assess the effect of missing responses for the primary outcome only. Unadjusted models will be utilised. This analysis will explore the possibility that missing

responses are 'missing not at random' (MNAR) using a tipping point approach. In this analysis, for women with missing outcome data, events will be added sequentially in each of the groups in turn to determine the point where the general conclusion changes (for example from positive to inconclusive in terms of the CI). An assessment can then be made about whether the event rate in the missing responses between groups is likely to be plausible when compared with the event rate in the non-missing data. Two scenarios will be considered as follows:

- Scenario A: In the 'treatment A' group, assume all missing responses are non-events (i.e. women randomised who experience a live birth is no). For missing responses in the 'treatment B' group, we will first replace X missing responses with events (i.e. X additional cases of live birth) where X is the number of events such that the event rate in the missing responses is equal to the event rate in the non-missing responses in the 'treatment B' group. This will be regarded as the base case. All other missing responses in the 'treatment B' group will be considered as non-events. An unadjusted model will be fitted. Event rates will be compared between groups. The CI from the treatment estimate (from the RR of the log-binomial model) will be examined and stored. Following the base case, an additional event will be added to the 'treatment B' group and the above procedure repeated (unadjusted model fitted and the RR and CI examined). This process will end when the number of events added to the 'treatment B' group are equal to the original number of missing responses in this group. The tipping point for the 'treatment B' group will occur when enough events have been added such that the upper/lower limit of the CI from the corresponding model differs from that of the primary ITT finding (in regards to whether the CI crosses the null value of one).

- Scenario B: In the 'treatment B' group, assume all missing responses are non-events (i.e. women randomised who experience a live birth is no). For missing responses in the 'treatment A' group, we will first replace X missing responses with events (i.e. X additional cases of live birth) where X is the number of events such that the event rate in the missing responses is equal to the event rate in the non-missing responses in the 'treatment A' group. This will be regarded as the base case. All other missing responses in the 'treatment A' group will be considered as non-events. An unadjusted model will be fitted. Event rates will be compared between groups. The CI from the treatment estimate (from the RR of the log-binomial model) will be examined and stored. Following the base case, an additional event will be added to the 'treatment A' group and the above procedure repeated (unadjusted model fitted and the RR and CI examined). This process will end when the number of events added to the 'treatment A' group are equal to the original number of missing responses in this group. The tipping point for the 'treatment A' will occur when enough events have been added such that the upper/lower limit of the CI from the corresponding model differs from that of

the primary ITT finding (in regards to whether the CI crosses the null value of one).

10. Analysis of sub-randomisations

Not applicable.

11. Health economic analysis

As indicated in the protocol there will also be an economic analysis. The details of this analysis are documented separately.

12. Statistical software

Statistical analysis will be undertaken in the following statistical software packages: SAS software and Stata.

13. References

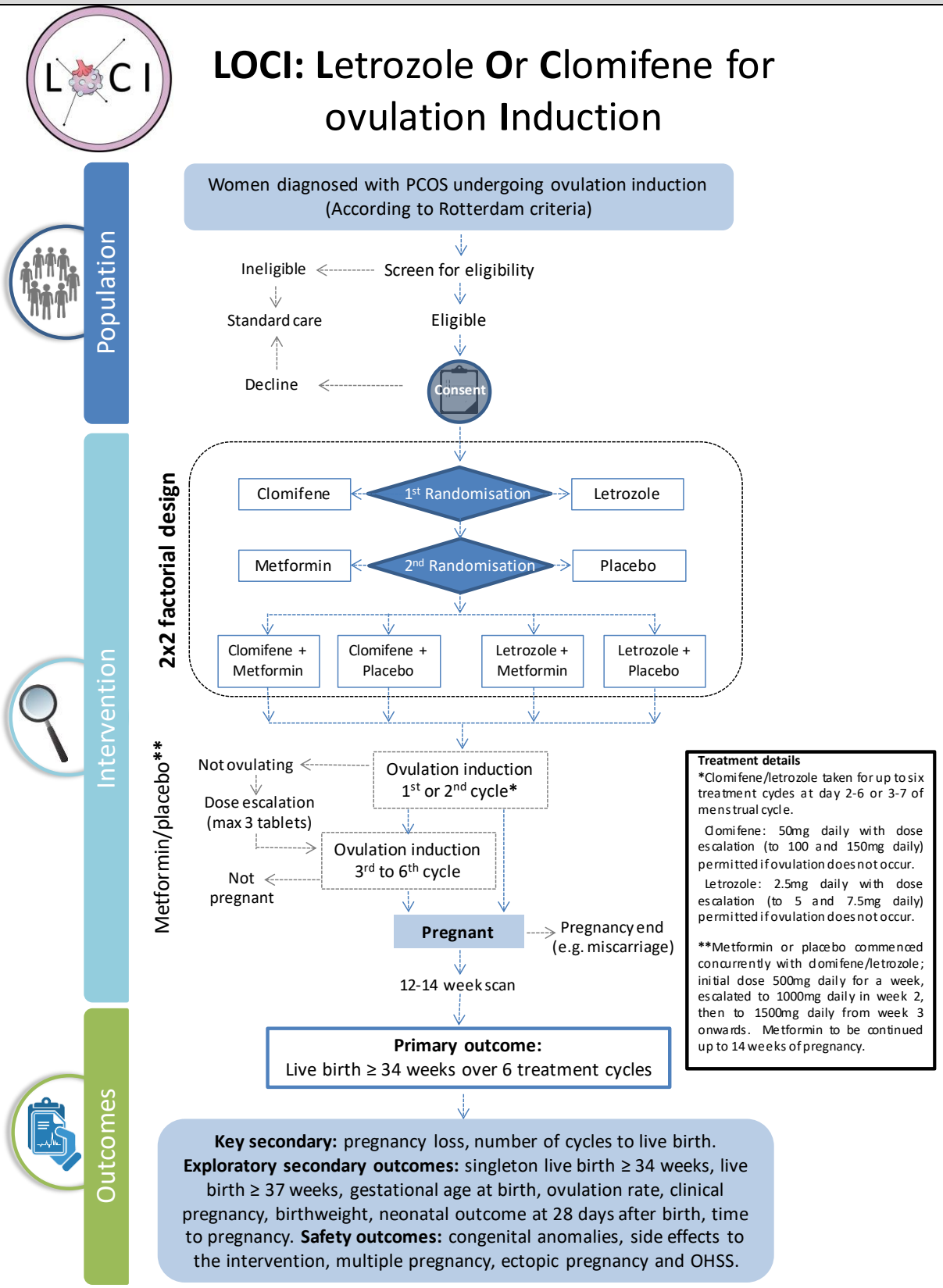
1. Huque MFD, A. ; D'Agostino, R. Multiplicity Issues in Clinical Trials With Multiple Objectives. Statistics in Biopharmaceutical Research 2013;5.Gupta SK. Intention-to-treat concept: A review. Perspect Clin Res. 2011;2(3):109-112.
2. Gupta SK. Intention-to-treat concept: A review. Perspect Clin Res. 2011;2(3):109-112.
3. 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.
4. Altman DG, Dore CJ. Randomisation and baseline comparisons in clinical trials. Lancet. 1990;335:149–53.
5. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol. 2004;159(7):702-6.
6. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ. 2011;342:d40.

Appendix A: Deviations from SAP

This report below follows the statistical analysis plan dated *<insert effective date of latest SAP>* apart from following:

Section of report not following SAP	**Reason**
<insert section >	<insert, e.g. exploratory analyses request by TMG>

Appendix B: Trial schema



LOCI Trial Schema v2.0 17-Jun-2020

Appendix C: Schedule of assessments

			From randomisation							
Visit	**Screening before clinic**	**Baseline clinic**	**Month 1**	**Month 2**	**Month 3**	**Month 4**	**Month 5**	**Month 6**	**Month 4-9***	**Month 2-18****
Eligibility check	x	x								
Valid informed consent		x								
Relevant medical history taken	x	x								
Concomitant medication		x	x	x	x	x	x	x	x	x
Randomisation		x								
Routine blood tests	x	x								
D21 progesterone concentration blood test (where applicable)			x	x	x	x	x	x		
Dispensing of IMP		x								
Ultrasound monitoring of ovulation (optional, depending on local practice)			x							
Drug returns (participant visit)								x†		
Phone call to participant								x	x#	
EQ-5D-5L questionnaires		x							x	x
11-14 week ultrasound scan (if participant becomes pregnant)									x	
Final outcomes (conception failure; pregnancy, antenatal, intrapartum, post-partum, Day 28 neonatal and maternal outcomes)										x

*Exact month determined by when the participant becomes pregnant (e.g. if they become pregnant in month 1, subsequent assessments will be made in month 4). **Exact month determined by when participant becomes pregnant, and if they miscarry or have a successful live birth. †Pregnancy to be confirmed by month 6, visit 2 scheduled accordingly. #Second phone call to be made at week 15 of pregnancy.

Appendix D: Template report

A template report for the final analyses will be provided in a separate document.