Updated statistical analysis plan for FIIX

**Trial registration:** ACTRN12619001003167.

This draft authored by Jack Wilkinson, with input from Melissa Vercoe, 15/11/22.

This version supersedes the original statistical analysis plan dated 5/5/20.

Introduction

## Rationale for update to the statistical analysis plan

The FIIX study is a multicentre, open-label, two-arm parallel non-inferiority randomised controlled trial with 1:1 allocation in New Zealand. Participants are allocated to receive either up to four cycles of IUI followed by two complete cycles of IVF, or two complete cycles of IVF. The study was conceived before the COVID-19 pandemic. Measures to control the spread of COVID-19 have affected the delivery of the study interventions in such a way that the originally planned analysis may not provide a useful answer to the study question. The planned analyses have been updated in response to the pandemic, to ensure that the study provides useful information about the relative effectiveness of the study interventions. The updated analyses have been informed by published guidance for analysis of randomised trials affected by a pandemic [1].

## Impact of the pandemic on the conduct of the trial

Due to pandemic control measures, the study treatments were paused in two periods while the study was active. This means that some participants who had been recruited to the study and randomised to one of the study arms faced a considerable delay to the start of their treatment. Some participants who had not yet received any study treatment did not commence treatment during this period, while participants who had commenced treatment but were not yet pregnant had to wait to undergo further attempts (further frozen embryo transfers in the IVF arm, or further insemination cycles in the IUI arm). Moreover, treatments in the two arms were paused for differing lengths of time, such that there were periods where clinics continued to perform IVF, but did not perform IUI. This means that there are periods where participants received IVF if they were randomised to the IVF arm, but received no treatment if they were randomised to the IUI arm. We therefore have periods in the study where the difference in outcomes between the groups reflects a comparison of ‘IVF versus no IUI’. While this might yield useful information about the relative effectiveness of these two interventions during a pandemic, the goal of the study is to provide information about the relative effectiveness of the two interventions outside of a pandemic.

We anticipate that an additional consequence of the pandemic, albeit one that is harder to quantify, is that there will have been delays to treatments in both arms throughout the study period. This could arise due to illness in clinic staff or in patients, for example. This could plausibly affect the two arms in the study differently.

In the following sections, we describe the updated statistical analysis plan, highlighting departures from the previous version.

Study methods

Estimand

We are interested in the effect of IVF and IUI as *treatment policies* [2]. This would usually mean that we are interested in the outcome measures in all participants, regardless of whether they are adherent to the intended treatment protocols. Moreover, participants would be analysed in their allocated groups, regardless of what they actually received. These two principles amount to a description of what has traditionally, albeit inconsistently, referred to as the ‘intention to treat effect’. However, in light of the COVID-19 pandemic, this description requires further elaboration [1]. Specifically, given the impact of the pandemic on the trial, we must indicate whether our interest is in estimating the relative effect of the study interventions within or outside of a pandemic context. Anticipating that the delivery of care will not continue to be interrupted in the same manner (or at least, not the same extent) by the COVID-19 pandemic in future, our primary interest is in estimating the relative effect of IUI compared to IVF outside of a pandemic context (or, more accurately, in a scenario where the treatments are not seriously disrupted by a pandemic). We have updated our primary analysis to reflect this estimand (estimand 1). An additional analysis of the outcome ‘live birth’ is planned corresponding to a different estimand: the relative effect of IUI compared to IVF in the context of an ongoing pandemic (estimand 2).

Trial design

The overall trial design remains unchanged. A multicentre, open label, two-arm parallel non-inferiority randomised controlled trial with 1:1 allocation. Participants are allocated to receive either up to four cycles of IUI followed by two complete cycles of IVF, or two complete cycles of IVF.

## Randomisation

Randomisation remains unchanged. Couples are randomly assigned to either the IUI followed by IVF arm or the IVF arm with a 1:1 allocation using a variable block design. The block sizes are not disclosed, to ensure concealment. The randomisation is stratified by centre and by age (<36 OR >=36).

## Timing of the primary outcome

The timing of the primary outcome has been changed for participants randomised from 19th August 2021 onwards. For these participants, the primary outcome, cumulative live birth rate (CLBR), will be assessed at 7 months (215 days) post-randomisation. Specifically, any live birth arising from a pregnancy initiated in the first 7 months from randomisation will be included. For participants randomised prior to this date, the timing of the primary outcome assessment remains unchanged (6 months/185 days). This change was made to accommodate the delays to treatment in both arms of the trial which was evident before and exacerbated by COVID-19. This amendment was approved by HDEC (reference code: 2022 AM 8410) February 2022.

Sample size

The originally planned sample size was 580 participants. This was based on the primary outcome, CLBR at 6 months from randomisation, using the following estimates from the literature:

* Estimated cumulative live birth rate (CLBR) of 30% after four cycles of IUI at 185 days (6 months).
* Estimate CLBR of 30% for a single completed IVF cycle at 185 days (6 months).
* Sample size calculated based on the hypothesis of noninferiority for CLBR at 6 months. Estimated CLBR is 30% in each group, and requires 580 patients (290 in each arm) for 80% power to reject the null hypothesis that the groups differ by more than 10 percentage points in favour of IVF at the 5% level, allowing for 10% withdrawals.

However, both of the planned analyses in the updated SAP will reduce power, either by excluding some of the randomised participants from the analysis, or by including data from participants during periods in which IUI was paused but IVF was not (such that we anticipate participants in the IUI arm to have inferior outcomes during these periods). We increased the recruitment target as a result, although anticipate that this may not be achieved due to resource limitations (cost and time). We targeted a sample size of 730.

Confidence and Significance

Type 1 error of the primary analysis will be controlled at 5%, by comparing the lower limit of a 90% two-sided confidence interval to the inferiority margin. A significance threshold of 1% will be used for secondary outcomes, which will be analysed using 99% confidence intervals.

No adjustment for multiplicity will be made for sensitivity analyses or for the alternative analyses of the primary outcome variable, CLBR, added in response to COVID-19 pandemic.

Compliance and protocol violation

This section details major changes to the planned analysis due to COVID-19 interruption. The primary analysis of CLBR will exclude all participants during periods where treatment was impacted by COVID-19 lockdowns. This means 151 participants will be removed from the analysis. March 2020 lockdown affected participants randomised between 22/09/2019 – 25/03/2020. August 2021 lockdown affected participants randomised between 14/02/2021 – 18/08/2021. The exclusions are based on date of randomization. Participants randomized in these periods will be excluded regardless of whether their individual treatment was or was not successful, and whether it was or was not significantly delayed. This decision is intended to prevent the introduction of bias, which would be a concern if we implicitly selected participants to exclude on the basis of their outcome (if we opted to include participants if they fell pregnant, for example). This analysis is intended to target the relative effect of IUI compared to IVF as treatment policies outside of a pandemic context (or more specifically, in a scenario where there are not major disruptions to the treatments due to a pandemic (estimand 1)).

A secondary analysis of CLBR will include all randomised participants, regardless of COVID interruptions. This analysis is intended to target the relative effect of IUI compared to IVF in the context of an ongoing pandemic (estimand 2).

## Interim analysis

No formal interim analysis is planned. Summaries of outcome data to date will be included in reports for the DMC.

## Timing of analysis

Statistical analysis for the final report will commence following the end of follow up of the final participant. However, analysis of 6/7 month (185/215 days) data (the primary outcome timepoint) may commence once the 6/7 month outcome has been obtained in all participants.

Missing Data

Missing covariate data

Since only the stratification variables are to be adjusted for in analyses, and these are anticipated to be complete (since they are required for randomisation), no action is required for missing covariate data.

Missing outcome data

Missing outcome data are anticipated due to the duration of the study. Any analysis in the presence of missing data will be subject to assumptions. We will conduct the primary analysis under one set of assumptions, but will perform sensitivity analyses allowing the assumption to vary.

For the analysis of CLBR at 6 and 7 months, we will perform analyses assuming:

* That the data are ‘missing at random’ given site and age (complete case analysis).
* That the patients with missing outcome data did not have a live birth, unless they were pregnant at the time of loss to follow up.

## Assumptions

The specified primary analysis is not contingent on checking of assumptions.

# **Study population**

## Screening data

No formal analysis of factors associated with participation are planned.

## Eligibility

The proportion of eligible participants from those screened, and reasons for ineligibility, will be tabulated.

## Recruitment and Attrition

The numbers screened, eligible, consenting, participating and contributing outcome data at each point (6/7 and 18/20 month assessments) will be presented in the form of a CONSORT flow diagram.

## Baseline Characteristics

The following table will be presented to summarise baseline characteristics of all randomised participants.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Total | Group A | Group B |
| Randomized | N | N | N |
| Age | Mean | Mean | Mean |
| Age by strata <36yo | % | *%* | *%* |
| Age by strata >/= 36yo | *%* | *%* | *%* |
| BMI (kg/m2) | *mean* | *mean* | *mean* |
| Duration of infertility (mo) | Median | Median | Median |
| Ethnicity | % | % | % |
| Māori |  |  |  |
| Pacific |  |  |  |
| Asian |  |  |  |
| MELAA |  |  |  |
| Other European  |  |  |  |
| NZ European |  |  |  |
| Previous births | % | % | % |
| Previous IVF | % | % | % |
| Previous IUI | % | % | % |
| Sperm Concentration | Mean/median | Mean/median | Mean/median |
| Sperm Motility | Mean/median | Mean/median | Mean/median |
| Total motile sperm count | Mean/median | Mean/median | Mean/median |
| Prediction score (Hanualt) | Mean | Mean | Mean |
| AMH level | Mean | Mean | Mean |

# Analysis

## Data manipulation

All data will be captured in a central database[3], such that merging of multiple datasets by the trial statistician will not be required. Time to event variables will be collected using date of randomisation and date of event variables. Data will be checked for internal coherence prior to analysis.

Statistical methods

Primary outcome (CLBR at 6 or 7 months/185 or 215 days)

Primary analysis of CLBR will be conducted using logistic regression. The outcome of any ongoing pregnancies at 6 months/185 days (7 months/215 days), following the change of the primary endpoint date) post randomisation will be included. This will be adjusted for the stratification variables: age and centre. For the purpose of comparison with the non-inferiority margin, a risk difference will be obtained from the fitted model, by predicting risk under both allocations for each patient. A bootstrap procedure will be used to calculate the 95% confidence interval, which will be used to test the hypothesis of non-inferiority.

Secondary outcomes

For continuous and binary outcomes, linear and logistic regression will be employed, adjusting for the stratification variables. For time to pregnancy leading to live birth, a Cox regression model will be used, and cumulative incidence will be plotted. Analyses of the secondary outcomes will be performed using the analysis sets corresponding to estimands 1 and 2 (see *Compliance and protocol violation*).

Because women have multiple chances at conception in the study, miscarriage and live birth are not mutually exclusive events. For example, a woman could have a miscarriage in one attempt, and a live birth in a subsequent attempt. In this case, she would be included both as having a miscarriage and as having a live birth in the respective analyses of these outcomes. Because a woman may have several pregnancies and miscarriages in this study, we will report these outcomes both as ‘any per woman’ (a binary outcome) and as ‘number per woman’ (a count outcome).

Subgroup analyses

Exploratory subgroup analyses for CLBR will be performed, but the trial is not powered to this end. These will involve tests of interaction between treatment and age and between treatment and number of previous treatment attempts. These analyses will be considered hypothesis generating.

Sensitivity analyses

Sensitivity analyses will be conducted around the missing data assumptions used in the analysis of the primary outcome (see section, Missing outcome data).

Safety data

Serious adverse events will be analysed descriptively, by treatment group.

Software

Analysis will be conducted in R.

References

[1] Cro S, Morris TP, Kahan BC, Cornelius VR, Carpenter JR. A four-step strategy for handling missing outcome data in randomised trials affected by a pandemic. BMC Med Res Methodol. 2020;20:208.

[2] European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials 2020.

[3] Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019;95:103208.