PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Comparison of the effect of immediate versus delayed transfer following a stimulated IVF cycle on the ongoing pregnancy rate of frozen-thawed embryo transfer cycles: a study protocol for a randomised controlled trial.
AUTHORS	LI, HE; LI, LU; LU, XIANG; SUN, XIAOXI; Ng, Ernest

VERSION 1 – REVIEW

REVIEWER	Christos Venetis University of New South Wales, Australia
REVIEW RETURNED	18-Nov-2017

GENERAL COMMENTS	This is the protocol of an original and very interesting RCT. There are a few issues that need to be clarified by the researchers before eventual publication. Issues: The authors argue in their introduction that delaying FET may add to the stress and anxiety accompanying IVF treatment. However, they do not seem to try and explore this hypothesis in their RCT by collecting appropriate data. I would urge the researchers, even if their study has already commenced recruitment, to consider collecting proper data regarding anxiety and distress in the two groups. Selection of primary outcome measure: Could you please clarify why you chose to have as a primary outcome measure ongoing pregnancy rate? It is now widely acknowledged by most Scientific Societies in Reproductive Medicine that live birth rate is the preferred outcome measure. It is also confusing the fact that you chose to collect data on live birth rates but only as a secondary outcome measure. Inclusion criteria: Please define “standard stimulation” which is a very vague term. What type of gonadotrophins, what type of GnRH analogue and most importantly what is the permitted starting dose? Similarly, there are many different definitions on what constitutes “mild stimulation”. Please define the criteria for the diagnosis of severe ovarian hyperstimulation syndrome. Randomization: Please describe in detail the exact timing of randomization. Before they start stimulation? On the day of ET or

	freezing? Also, give details about randomization like whether any type of stratification or block randomization will be used. For example, stratifying according to whether patients had initially a freeze-all cycle or a fresh ET as well as based on the developmental stage of the embryo transferred would probably safeguard this study by imbalances that might spuriously skew the results. Also, describe who will be performing the randomization and the exact routine. Will this researcher be involved in recruitment? What measures are you taking in your study to ensure allocation concealment? Blinding: the embryologist performing the quality assessment of embryos and hence determining the order of transfer of embryos could be blinded to the allocated treatment. Outcome measures: 1) Please define what constitutes a positive b-hCG level 2) When will clinical pregnancy be assessed? At gestational week 6? 3) Please define miscarriage rates, including the denominator 4) Please define live birth rate. What is your cut-off in terms of gestational age and why? Power analysis: Please clarify why you test for equivalence? Also, please clarify why you test for a difference of 10%. Is that based on previously published literature? Statistical analysis: Please clarify which variables you will compare with the use of multivariable logistic regression analysis and why. DMSB: Are there predefined criteria for premature termination of the study for safety?

REVIEWER	Christophe Blockeel Centre for reproductive Medicine Universitair Ziekenhuis Brussel Belgium
REVIEW RETURNED	28-Nov-2017

GENERAL COMMENTS	As the authors state, there is indeed a need for a straight forward RCT comparing the question: To delay or not to delay? Some issues need some clarification: In the inclusion, it states "Undergoing IVF with a standard stimulation". 1/ Is ICSI included, or only IVF, please adjust. 2/ As you describe, both Agonists and antagonists are allowed in the study: this implies a certain bias, so I would strongly advise to stick to the GnRH antagonist protocol only. 3/ Related to the same topic, both GnRH agonist and hCG can be used to trigger final oocyte maturation: this also yields a bias, given the complete different endocrine profile and impact on the endometrium following both agents. I would prefer a GnRH agonist trigger strategy only to draw firm conclusions! 4/ The same is true for D3 and D5. Would there be an option to to blastocyst transfer only?

	More important: 5/ In the statistical analysis, you state that a difference of 10% between both groups can be considered as equivalent. I'm afraid this difference is not equivalent at all. Therefore, I would suggest to revise this statistical analysis.

REVIEWER	Matheus Roque ORIGEN - Center for Reproductive Medicine, Brazil
REVIEW RETURNED	12-Dec-2017

GENERAL COMMENTS	This is a RCT of interest in Reproductive Medicine, as the freeze-all policy is of increasing interest in the field and there are no available RCT evaluating the best strategy to implemented after performing freeze-all. The study design may have some improvements before publication: * MATERIAL AND METHODS - I suggest you include in ¨Inclusion Criteria¨ that all patients have performed a freeze-all cycle; You mentioned ¨*First FET cycle following ovarian stimulation in IVF¨ . This can also be related to those patients that have already performed a Fresh embryo transfer and then will perform their First FET cycle. Thus, ¨First FET cycle¨ doesn´t mean first embryo transfer in a specific cycle; - in INTERVENTIONS: the number of IVF cycles with antagonist and agonist protocols should be the same in both groups. It is not clearly stated who are the patients that you perform agonist and antagonist protocol. And this should lead to differences in the groups; - Oocyte retrieval performed 34-36 hours triggering - if the groups present different number of patients with different timing of triggering, this may be subject of bias. If in one group there are more patient with 34 hours between trigger and OPU and the other group more patients with 36 hours, this may be subject of bias. Different time intervals between the trigger and OPU may lead to differences in laboratory. - It is not clear which patients will perform cleavage and blastocyst embryo transfer - Concerning about blastocysts, you will follow and freeze embryos until which day (day 5, 6 ,7) ? - The patients will be randomized by cleavage and also by blastocyst groups? If the randomization doesn´t take in account the embryo development stage, there is a risk of having more cleavage transfers in one group and blastocyst in theater group; - Will Serum hCG be checked 14 days after FET for cleavage and also blastocyst transfer? - what is the definition of miscarriage rates? In your protocol in Clinical Trials it is presented that Miscarriage is the pregnancy loss before the 28th week of pregnancy. Is this correct? * DISCUSSION - in this section you stated that to increase the generalizability or your results, you´ll include both patients having the first FET cycle after a failed stimulated IVF cycle or undertaking freeze all strategy. This was not clearly stated in Material and Methods. Moreover, it seems that these are not similar patients: freeze-all and patients that failed in a previous fresh embryo transfer that will perform a first FET

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Christos Venetis
Institution and Country: University of New South Wales, Australia
Competing Interests: None declared

This is the protocol of an original and very interesting RCT.
There are a few issues that need to be clarified by the researchers before eventual publication.

Issues:
The authors argue in their introduction that delaying FET may add to the stress and anxiety accompanying IVF treatment. However, they do not seem to try and explore this hypothesis in their RCT by collecting appropriate data. I would urge the researchers, even if their study has already commenced recruitment, to consider collecting proper data regarding anxiety and distress in the two groups.
Response: This is a very good point as we do not have data regarding the stress and anxiety in this group of patients undergoing IVF treatment. We can measure the stress and anxiety levels by the standard questionnaire before the randomization and at the time of starting FET. L186-189

Selection of primary outcome measure: Could you please clarify why you chose to have as a primary outcome measure ongoing pregnancy rate?
It is now widely acknowledged by most Scientific Societies in Reproductive Medicine that live birth rate is the preferred outcome measure. It is also confusing the fact that you chose to collect data on live birth rates but only as a secondary outcome measure.
Response: Thanks for the suggestion. We feel that ongoing pregnancy is an appropriate endpoint for this study. Clarke et al. investigated earlier that conclusions on the effectiveness of a treatment based on either clinical pregnancy or live birth as endpoints are comparable. (Clarke et al Fertil Steril 2010). Braakhekke M et al suggested ongoing pregnancy best serves the many purposes of a primary outcome and best reflects the effectiveness of a treatment (Braakhekke M et al Fertil Steril 2014). We do trace and report the live birth as a secondary outcome. In some patients, we may not be able to trace the live birth outcome of all patients because of loss of follow up.

- Clarke JF, van Rumste MM, Farquhar CM, et al. Measuring outcomes in fertility trials: can we rely on clinical pregnancy rates? Fertil Steril 2010;94(5):1647-51.
- Braakhekke M, Kamphuis EI, Dancet EA, et al. Ongoing pregnancy qualifies best as the primary outcome measure of choice in trials in reproductive medicine: an opinion paper. Fertil Steril 2014;101(5):1203-4.

Inclusion criteria:
Please define “standard stimulation” which is a very vague term. What type of gonadotrophins, what type of GnRH analogue and most importantly what is the permitted starting dose?
Similarly, there are many different definitions on what constitutes “mild stimulation”.
Response: Standard and mild stimulation is defined according to the published terminology for ovarian stimulation for IVF. L136-137

- Nargund G, Fauser BC, Macklon NS, et al. The ISMAAR proposal on terminology for ovarian stimulation for IVF. Hum Reprod 2007;22(11):2801-4.

Please define the criteria for the diagnosis of severe ovarian hyperstimulation syndrome.
Response: OHSS is diagnosed and classified according to the RCOJ guideline. L137-138

- Green-top guideline No.5: Ovarian Hyperstimulation Syndrome, https://www.rcog.org.uk/en/guidelines-research-services/ guidelines/gtg5/. (26 February 2015, date last accessed).

Randomization: Please describe in detail the exact timing of randomization. Before they start stimulation? On the day of ET or freezing?
Response: The exact timing of randomization is on the day of embryo freezing for patients taking the freeze all strategy and on the day of blood HCG test on 14 days after fresh-ET for the failed fresh-ET women. L145-148

Also, give details about randomization like whether any type of stratification or block randomization will be used. For example, stratifying according to whether patients had initially a freeze-all cycle or a fresh ET as well as based on the developmental stage of the embryo transferred would probably safeguard this study by imbalances that might spuriously skew the results.
Response: We do not take any type of stratification or block randomization and we just randomize all recruited women into two groups. The sample size will be too small to perform subgroup analysis.

Also, describe who will be performing the randomization and the exact routine. Will this researcher be involved in recruitment?
Response: The randomization is carried out by a project nurse who is not involved in the recruitment and clinical management of patients using an online randomization program through the website www.randomization.com. Then the nurse will prepare the randomization arm and put it into opaque envelops for use. On the randomization day, the recruited women will be randomized according to the opaque envelops into one of the two groups. We have added this in the text. L148-154

What measures are you taking in your study to ensure allocation concealment?
Response: The randomisation is carried out by a project nurse using the online randomization program through the website www.randomization.com. Then the nurse will prepare the randomization arm and put it into opaque envelops for use. The study nurse is not involved in recruitment and clinical management of patients.

Blinding: the embryologist performing the quality assessment of embryos and hence determining the order of transfer of embryos could be blinded to the allocated treatment.
Response: Yes, the embryologist performing the quality assessment of embryos is blinded to the allocated treatment. We have adjusted the blinding part as “Both the researchers and the participants cannot be blinded because the nature of the study. The embryologist performing the quality assessment is blinded to the allocated treatment.” L161-163

Outcome measures:
1) Please define what constitutes a positive b-hCG level
2) When will clinical pregnancy be assessed? At gestational week 6?
3) Please define miscarriage rates, including the denominator
4) Please define live birth rate. What is your cut-off in terms of gestational age and why?
Response:1) serum β-hCG is measured to determine pregnancy 14 days after embryo transfer. Conception is defined with the result of serum β-hCG ≥10 mIU/mL.
2) Clinical pregnancy will be assessed by transvaginal ultrasound at gestational week 6.
3) Miscarriage rate is defined as a clinically recognized pregnancy loss before the 22 weeks of pregnancy. The denominator is the clinical pregnancy.
4) A live birth is defined as the delivery of any number of newborns ≥22 weeks’ gestation with heartbeat and breath.
We revised the text. L223-232

Power analysis:
Please clarify why you test for equivalence?

Response: As most of the published studies show comparable ongoing pregnancy rates of immediate and delayed FET, the hypothesis is of the present study is to test for equivalence.

Also, please clarify why you test for a difference of 10%. Is that based on previously published literature?
Response: It was assumed that an absolute difference of more than 10% in ongoing pregnancy rate will be of clinical significance. A recent reference supported this point.

- Tournaye H, Sukhikh GT, Kahler E, et al. A Phase III randomized controlled trial comparing the efficacy, safety and tolerability of oral dydrogesterone versus micronized vaginal progesterone for luteal support in in vitro fertilization. Human Reproduction 2017;32(5):1019.

Statistical analysis:
Please clarify which variables you will compare with the use of multivariable logistic regression analysis and why.
Response: We use the multivariable logistic regression to adjust for potentially con- founding factors and results, namely female age (as a continuous variable), retrieved oocytes, COH protocol, ovulation trigger, number of good quality embryos produced (as a continuous variable) and number of embryos transferred (one versus two), developmental stage (cleavage versus blastocyst stage) and quality of the embryos transferred (quality of the embryo transferred). We have revised the text accordingly (L263-269).

DMSB: Are there predefined criteria for premature termination of the study for safety?
Response: Since FET in HRT cycles is a standard procedure in IVF centers, and there is no agreement regarding the time interval between the stimulated IVF and the subsequent FET in the literature, there are not predefined criteria for premature termination of the study. There is no interim analysis during the study. L276-280

Replies to Reviewer 2

Reviewer: 2
Reviewer Name: Christophe Blockeel
Institution and Country: Centre for reproductive Medicine, Universitair Ziekenhuis Brussel, Belgium
Competing Interests: None declared

As the authors state, there is indeed a need for a straight forward RCT comparing the question: To delay or not to delay?
Some issues need some clarification:
In the inclusion, it states "Undergoing IVF with a standard stimulation".
1/ Is ICSI included, or only IVF, please adjust.
Response: Both IVF and ICSI are included. We have revised as “Women aged <=43 years at the time of IVF/ICSI treatment”.L124

2/ As you describe, both Agonists and antagonists are allowed in the study: this implies a certain bias, so I would strongly advise to stick to the GnRH antagonist protocol only.
Response: Since both agonists and antagonists having the same chance to be randomized into the immediate FET group and the delayed FET group, we have included both women undertaken the agonist and antagonists protocol to make the results generally applicable.

3/ Related to the same topic, both GnRH agonist and hCG can be used to trigger final oocyte maturation: this also yields a bias, given the complete different endocrine profile and impact on the endometrium following both agents.
I would prefer a GnRH agonist trigger strategy only to draw firm conclusions!

Response: Since both GnRH agonist and hCG trigger having the same chance to be randomized into the immediate FET group and the delayed FET group, we have included both women undertaken the GnRH agonist and hCG trigger to make the results generally applicable.

4/ The same is true for D3 and D5. Would there be an option to blastocyst transfer only?
Response: The exact timing of randomization is on the day of embryo freezing for patients taking the freeze all strategy and on the day of blood HCG test on 14 days after fresh-ET for the failed fresh-ET women. Since both D3 and blastocyst embryo having the same chance to be randomized into the immediate FET group and the delayed FET group, we have included women transferring D3 embryo and who transferring blastocyst embryo to make the results generally applicable.

More important:
5/ In the statistical analysis, you state that a difference of 10% between both groups can be considered as equivalent. I'm afraid this difference is not equivalent at all. Therefore, I would suggest to revise this statistical analysis.
Response: It was assumed that an absolute difference of more than 10% in ongoing pregnancy rate will be of clinical significance. A recent reference supported this point.

- Tournaye H, Sukhikh GT, Kahler E, et al. A Phase III randomized controlled trial comparing the efficacy, safety and tolerability of oral dydrogesterone versus micronized vaginal progesterone for luteal support in in vitro fertilization. Human Reproduction 2017;32(5):1019.

Replies to Reviewer 3

Reviewer: 3
Reviewer Name: Matheus Roque
Institution and Country: ORIGEN - Center for Reproductive Medicine, Brazil
Competing Interests: None

This is a RCT of interest in Reproductive Medicine, as the freeze-all policy is of increasing interest in the field and there are no available RCT evaluating the best strategy to implemented after performing freeze-all.
The study design may have some improvements before publication:

* MATERIAL AND METHODS
- I suggest you include in ¨Inclusion Criteria¨ that all patients have performed a freeze-all cycle; You mentioned ¨*First FET cycle following ovarian stimulation in IVF¨ . This can also be related to those patients that have already performed a Fresh embryo transfer and then will perform their First FET cycle. Thus, ¨First FET cycle¨ doesn´t mean first embryo transfer in a specific cycle;
Response: We have no evidence to support that there is difference between the freeze-all group and those who failed to get pregnant after transfer in a fresh cycle. We recruit both freeze-all and those who fail to get pregnant after transfer in the fresh cycle as this is a randomized study and two groups of patients have equal chance to be randomized into one of the allocation arms. This will make the results more applicable to the practical situation.

- in INTERVENTIONS:
the number of IVF cycles with antagonist and agonist protocols should be the same in both groups. It is not clearly stated who are the patients that you perform agonist and antagonist protocol. And this should lead to differences in the groups;
Response: Patients will have antagonist and agonist protocols according to the doctors` and patients` preference and make the discussion together. There were no specific rules who will take antagonist and who will take agonist protocols. Since both agonists and antagonists having the same chance to be randomized into the immediate FET group and the delayed FET group, we have included both women undertaken the agonist and antagonists protocol to make the results generally applicable.

- Oocyte retrieval performed 34-36 hours triggering - if the groups present different number of patients with different timing of triggering, this may be subject of bias. If in one group there are more patient with 34 hours between trigger and OPU and the other group more patients with 36 hours, this may be subject of bias. Different time intervals between the trigger and OPU may lead to differences in laboratory.
Response: Since patients with different timing of triggering will have the same chance to be randomized into the immediate FET group and the delayed FET group, we think the bias of time intervals between the trigger and OPU may be avoided.

- It is not clear which patients will perform cleavage and blastocyst embryo transfer
Response: Patients ≥6 embryos of on day 3 were eligible for a blastocyst culture and transfer. We added this in the draft. L184-185

- Concerning about blastocysts, you will follow and freeze embryos until which day (day 5, 6 ,7) ?
Response: Concerning about blastocysts, we will follow and freeze embryos until day 6. Frozen blastocysts include Day 5 and Day 6 embryos.

- The patients will be randomized by cleavage and also by blastocyst groups? If the randomization doesn´t take in account the embryo development stage, there is a risk of having more cleavage transfers in one group and blastocyst in theater group;
Response: The exact timing of randomization is on the day of freezing. Since both cleavage and blastocyst embryo having the same chance to be randomized into the immediate FET group and the delayed FET group, we have included women transferring cleavage embryo and who transferring blastocyst embryo to make the results generally applicable.

- Will Serum hCG be checked 14 days after FET for cleavage and also blastocyst transfer?
Response: Serum hCG level will be checked 14 days after FET for both cleavage and blastocyst transfer.

- what is the definition of miscarriage rates? In your protocol in Clinical Trials it is presented that Miscarriage is the pregnancy loss before the 28th week of pregnancy. Is this correct?
Response: Thank you for your comments. Miscarriage rate is defined as a clinically recognized pregnancy loss before the 22 weeks of pregnancy. We have revised the text accordingly L228-230. We will revise the protocol in Clinical Trials later.

* DISCUSSION
- in this section you stated that to increase the generalizability or your results, you´ll include both patients having the first FET cycle after a failed stimulated IVF cycle or undertaking freeze all strategy. This was not clearly stated in Material and Methods. Moreover, it seems that these are not similar patients: freeze-all and patients that failed in a previous fresh embryo transfer that will perform a first FET
Response:
We have no evidence to support that there is difference between the freeze-all group and those who failed to get pregnant after transfer in a fresh cycle.
We recruit both freeze-all and those who fail to get pregnant after transfer in the fresh cycle as this is a randomized study and two groups of patients have equal chance to be randomized into one of the allocation.

Replies to editors
Editorial Requests:

1. Can you improve the description of the study design in the title and abstract? It should be clear that it is a randomised controlled trial.
Response: We have adjusted the title according to your suggestion into: " Comparison of the effect of immediate versus delayed transfer following a stimulated IVF cycle on the ongoing pregnancy rate of frozen-thawed embryo transfer cycles: a study protocol for a randomised controlled trial". We have adjusted the abstract part: “Methods/analysis: This is a randomized controlled trial.” L25

2. Please expand the 'strengths and limitations' section on page 4. This section should contain four or five short bullet points, no longer than one sentence each, that relate specifically to the methods or design of the study reported (see: http://bmjopen.bmj.com/site/about/guidelines.xhtml#articletypes).
Response: We have added another two bullet points to the “strengths and limitations” section.
2. This is the first trial that seeks to add significantly to the clinical evidence base and to allow conclusions to be made on the time interval in the FET following a stimulated IVF cycle.
L49-51
3. The study includes women aged 20–43 years undergoing the first FET after GnRH agonist and GnRH antagonist ovarian stimulation in IVF/ICSI; thus, results can be extrapolated to the majority of the infertile population. L52-55

3.Please thoroughly check the manuscript for grammatical/ typographical errors e.g. page 4: “ClinicalTrial.gov” should be “ClinicalTrials.gov”.
Response: Thank you for your comment and please excuse us for the carelessness. We have checked the grammatical/typographical errors and revised it.

4.Can you include a schematic diagram as recommended in Item 13 of the SPIRIT checklist?
Response: I have supplemented the schematic diagram.

5.Please check that your study protocol is adequately reported. The SPIRIT checklist needs to be filled in better. Some items have not been completed. Please remember to include the relevant page number(s) from the manuscript next to each reporting item or state 'n/a' next to items that are not applicable to your study. For help and guidance completing the checklist see: http://www.bmj.com/content/346/bmj.e7586
Response: I have revised the text and completed the SPIRIT checklist.

We hope that the revised paper can now be acceptable for publication and we look forward to your final decision on the manuscript.

VERSION 2 – REVIEW

REVIEWER	Matheus Roque ORIGEN - Center for Reproductive Medicine, Brazil
REVIEW RETURNED	12-Jan-2018

GENERAL COMMENTS	In the present form, this manuscript can be published in this journal

REVIEWER	Christophe Blockeel Centre for Reproductive Medicine Universitair Ziekenhuis Brussel Laarbeeklaan 101 1090 Brussels Belgium
REVIEW RETURNED	13-Jan-2018

GENERAL COMMENTS	The authors have answered to the questions.

REVIEWER	Christos Venetis University of New South Wales, Australia
REVIEW RETURNED	01-Feb-2018

GENERAL COMMENTS	Then authors have provided a response to the comments made during the first review round and have performed relevant adjustments in most cases. There are some remaining or new points from my side that the authors might want to consider before eventual publication 1) Lines 184-185: ". Patients ≥6 good quality embryos of on day 3 were eligible for a blastocyst culture and transfer." Please rephrase this sentence since it is grammatically incorrect. Furthermore, since this is a protocol for a future or ongoing study, past tense should not be used. 2) Please correct "RCOJ" to "RCOG" 3) I understand that the researchers' choice of selecting 10% as the equivalence threshold has been probably dictated by the sample size calculation and the overall feasibility of the study, however, this still represents a significant drawback of this (and other) research. Clinicians consider the threshold of clinical significance to be much lower (usually ~5%) and research has shown that patients are even more strict and would switch treatment for an efficacy difference of ~2%. If the current study shows a non-significant difference of -9%, the authors might claim equivalence, but no one else would actually come to the same conclusion. I think that this limitation should be discussed extensively. 4) I remain sceptical as to why the authors need to use a multivariable analysis to adjust for potential confounding since this is a randomized trial of a decent size. It will most likely lead to the same conclusion but I suppose it is reassuring to check, for the odd scenario where randomization has not achieved perfect balance between important confounders. Secondly, I am also not convinced that all the reported variables represent true confounders, especially if they are all included in the same multivariable model. Including many collinear and sometimes miss-specified covariates in a multivariable model can significantly backfire. Finally, the authors have omitted from their analysis probably one of the most important covariates which would be whether the cycle from which the embryo originates was a freeze-all or not.

VERSION 2 – AUTHOR RESPONSE

Reviewers' Comments to Author:

Reviewer: 1
Reviewer Name: Christos Venetis
Institution and Country: University of New South Wales, Australia
Competing Interests: None

Then authors have provided a response to the comments made during the first review round and have performed relevant adjustments in most cases.

There are some remaining or new points from my side that the authors might want to consider before eventual publication

1) Lines 184-185: ". Patients ≥6 good quality embryos of on day 3 were eligible for a blastocyst culture and transfer." Please rephrase this sentence since it is grammatically incorrect. Furthermore, since this is a protocol for a future or ongoing study, past tense should not be used.
Response: This was corrected L174-176.

2) Please correct "RCOJ" to "RCOG"
Response: This was corrected.

3) I understand that the researchers' choice of selecting 10% as the equivalence threshold has been probably dictated by the sample size calculation and the overall feasibility of the study, however, this still represents a significant drawback of this (and other) research. Clinicians consider the threshold of clinical significance to be much lower (usually ~5%) and research has shown that patients are even more strict and would switch treatment for an efficacy difference of ~2%. If the current study shows a non-significant difference of -9%, the authors might claim equivalence, but no one else would actually come to the same conclusion. I think that this limitation should be discussed extensively.
Response: Many studies have also assumed that an absolute difference of more than 10% in ongoing pregnancy rate will be of clinical significance for sample size calculation. This limitation is added in the text L57-60.

4) I remain sceptical as to why the authors need to use a multivariable analysis to adjust for potential confounding since this is a randomized trial of a decent size. It will most likely lead to the same conclusion but I suppose it is reassuring to check, for the odd scenario where randomization has not achieved perfect balance between important confounders. Secondly, I am also not convinced that all the reported variables represent true confounders, especially if they are all included in the same multivariable model. Including many collinear and sometimes miss-specified covariates in a multivariable model can significantly backfire. Finally, the authors have omitted from their analysis probably one of the most important covariates which would be whether the cycle from which the embryo originates was a freeze-all or not.
Response: We have added and revised the relevant section about multivariable analysis in the text L254-264. We have added the important covariates which is whether the patient is under failed fresh ET or freeze-all policy in the text L257.

Reviewer: 2
Reviewer Name: Christophe Blockeel
Institution and Country: Centre for Reproductive Medicine, Universitair Ziekenhuis Brussel
Laarbeeklaan 101
1090 Brussels, Belgium
Competing Interests: None

The authors have answered to the questions.

Reviewer: 3
Reviewer Name: Matheus Roque
Institution and Country: ORIGEN - Center for Reproductive Medicine, Brazil
Competing Interests: None declared

In the present form, this manuscript can be published in this journal

Replies to editors
Editorial Requests:

- Page 2 (Abstract): Please revise “Trial registration numbers: NCT03201783” to “Trial registration number: NCT03201783”.
Response: This is corrected.

Please carefully proofread the rest of the paper one more time. The presentation of your work could be improved in places. For example, please do not use one sentence paragraphs (such as page 11: "An audit trail will be designed as another security measure.")
Response: Thank you for your comment, we have rechecked the paper and made further improvement in the presentation of the work.

- You have included a statement about ethics approval on page 12. Can you please remove this section and instead include this information in the 'Ethics and dissemination' section on page 10?
Response: We have removed the ethics approval on page 12 and included the information on Page 11.

- If you are going to include an (optional) 'Discussion' section, then can you please improve the quality of this section? For example, you discuss some of the study's strengths but you do not discuss the study's limitations. Unless you intend to elaborate further on the strengths and limitations reported on page 3 then we suggest removing this section. However, you should include a 'Study status' section that includes the following information: "The study was designed in May 2017, and the first participant was randomised on 9 August 2017. At the time of the manuscript preparation, we have recruited 200 women and the recruitment is ongoing."
Response: We have removed the discussion section and include a study status section.

We hope that the revised paper can now be acceptable for publication and we look forward to your final decision on the manuscript.

VERSION 3 – REVIEW

REVIEWER	Christos Venetis UNSW, Australia
REVIEW RETURNED	09-Mar-2018

GENERAL COMMENTS	The authors have satisfactorily addressed the points of the previous review round and this protocol is now ready for publication.