PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	TADAFER II: Tadalafil treatment for fetal growth restriction - a study protocol for a multicenter randomized controlled phase II trial.
AUTHORS	Umekawa, Takashi; Maki, Shintaro; Kubo, Michiko; Tanaka, Hiroaki; Nii, Masafumi; Tanaka, Kayo; Osato, Kazuhiro; Kamimoto, Yuki; Tamaru, Satoshi; Ogura, Toru; Nishimura, Yuki; Kodera, Mayumi; Minamide, Chisato; Nishikawa, Masakatsu; Endoh, Masayuki; Kimura, Tadashi; Kotani, Tomomi; Nakamura, Masamitsu; Sekizawa, Akihiko; Ikeda, Tomoaki

VERSION 1 – REVIEW

REVIEWER	Francesc Figueras Hospital Cllinic. University of Barcelona, Spain
REVIEW RETURNED	12-Jan-2018

GENERAL COMMENTS	In the abstract, the main outcome (increased fetal growth) is difficult to understand. I would say: fetal growth velocity... The inclusion criteria (EFW<1.5 SD) is too wide. That will include by definition about 15% of the population, with most cases corresponding to constitutional smallness, where treatment is unnecessary and the ethical principle of “more good than harm” is doubtful to me. A Doppler based criteria (in accordance with the Delphi criteria for FGR) would have yield a more clinically homogeneous population. I found intriguing and a source of selection bias the exclusión criterion: "The investigator decides that entry is inappropriate". The stopping criteria c and d, are somewhat arbitrary and a source of bias. I would recommend an intention-to-treat analysis. Per protocol analysis is known to result in overoptimistic results. In their managing protocol I missed the DV: there´s compelling evidence that DV plus CTG improves the perinatal outcome (TRUFFLE study, Evidence level I). Thus, that will be limitation to the external validity of their results... It is also a limitation that no placebo was planned. Estimated fetal weight is subjected to operator-related variability. I could not tell whether operators will be blinded to the study group. For sure, women will not be blinded.

REVIEWER	J.W. Ganzevoort, A.Pels Academisch Medisch Centrum, Amsterdam, The Netherlands
REVIEW RETURNED	16-Jan-2018

GENERAL COMMENTS	- General: Interesting study with a clear rationale regarding pharmacodynamics.

	- Many authors are from a single centre. It would be good to have neonatologists involved as well. - Regarding safety, it would be good to present an overview about the safety of tadalafil in pregnancy, since it does not seem to be frequently investigated before. - Why was chosen for an open-label design versus placebo-controlled design? - Regarding patient selection: Since doppler abnormalities are not included as inclusion criteria, other causes of fetal growth restriction might be included as well. AEDF and REDF are difficult to include, since it could be an indication for delivery. - In the protocol is stated that antepartum fetal testing will be performed to assess fetal wellbeing. What are the differences in fetal testing before and after viability? - In case of intrauterine death, how will the primary outcome be defined? - Page 10: what is the timepoint where the secondary endpoints fetal head circumference, deepest amniotic fluid pocket, doppler imaging of umbilical arterial blood flow are assessed? - Page 10: Why is preeclampsia no secondary outcome? - Page 11: How will the pediatric development be assessed? Which neurodevelopmental test will be performed? - Page 11: Will the Safety Committee have blinded or unblinded access to the data? - Page 13: What is the detectable alternative you would like to prove with the study? Which difference in fetal growth velocity is the power calculation based on? - Page 13: A per protocol analysis is an analysis in which the patients who received the treatment they were allocated for, are included. I'm not sure what you mean by the 'per protocol set', but it seems to me that you mean the 'intention to treat' analysis here, if your aim is to include all randomized patients in this analysis. - Page 13: It would be good to mention the IQR for the median as well. - Page 19: The survival rates of VLBW and premature children seems quite high to me, e.g. 49% survival at AD 22 weeks and 401-500 gram. It would be good to add some information about how many patients these data are based on. - It might be useful for the authors to reach out to the STRIDER collaboration group to aim to align trial design as much as possible in this phase to be able to come to meaningful meta-analysis. - Why do the authors not aim for datasharing once published?

VERSION 1 – AUTHOR RESPONSE

Reviewer #1

Dr. Francesc Figueras

We would like to thank Dr. Francesc Figueras very much for taking time to comment on our paper.

1. In the abstract, the main outcome (increased fetal growth) is difficult to understand. I would say: fetal growth velocity...

According to the reviewers' suggestion, We changed the following sentences in the Abstract section from:

"Introduction: There is no proven therapy to reverse or ameliorate fetal growth restriction (FGR). Sildenafil, a selective phosphodiesterase 5 (PDE5) inhibitor, has been reported to potentially have a

therapeutic role in FGR, but this has not been established. Tadalafil is also a selective PDE5 inhibitor and has a longer half-life and a more rapid onset of action than sildenafil. We have demonstrated efficacy for tadalafil on fetal growth in FGR and the short-term outcomes and feasibility of tadalafil in FGR. Based on the hypothesis that tadalafil will safely increase the likelihood of increased fetal growth in FGR, we have designed this phase II study to prospectively evaluate the efficacy and safety of tadalafil in FGR.
Methods and analysis: This study is a multicenter randomized controlled phase II trial. A total of 140 fetuses with FGR will be enrolled from major medical centers in Japan. Fetuses will be randomized to receive either the conventional management for FGR, according to the guidelines in Japan, or a once daily treatment with 20 mg of tadalafil along with the conventional management, until delivery. Fetal growth velocity from the first day of the treatment to birth has been defined as the primary endpoint. To minimize bias in terms of fetal baseline conditions and timing of delivery, a fetal indication for delivery is established in this study based on the results from a Japanese multicenter survey. The investigator will evaluate fetal baseline conditions at enrollment and will decide the timing of delivery based on this fetal indication. Infants will be followed up for development until 1.5 years of age.
Ethics and dissemination: This study was approved by the Institutional Review Board of Mie University Hospital and each participating institution. Our findings will be widely disseminated through peer-reviewed publications.
Trial registration: UMIN Clinical Trials Registry UMIN000023778."
To:
"Introduction: There is no proven therapy to reverse or ameliorate fetal growth restriction (FGR). Sildenafil, a selective phosphodiesterase 5 (PDE5) inhibitor, has been reported to potentially have a therapeutic role in FGR, but this has not been established. Tadalafil is also a selective PDE5 inhibitor. We have demonstrated efficacy for tadalafil on fetal growth in FGR and the short-term outcomes and feasibility of tadalafil in FGR. Based on the hypothesis that tadalafil will safely increase the likelihood of increased fetal growth in FGR, we have designed this phase II study to prospectively evaluate the efficacy and safety of tadalafil in FGR.
Methods and analysis: This study is a multicenter randomized controlled phase II trial. A total of 140 fetuses with FGR will be enrolled from medical centers in Japan. Fetuses will be randomized to receive either the conventional management for FGR, according to the guidelines, or a once daily treatment with 20 mg of tadalafil along with the conventional management, until delivery. The primary endpoint is fetal growth velocity from the first day of the protocol-defined treatment to birth (g/day), and is calculated using the following formula:
Fetal growth velocity (g/day)
＝(Birthweight – Estimated fetal weight at the first day of the treatment [g])/(Days of the treatment [days])
To minimize bias in terms of fetal baseline conditions and timing of delivery, a fetal indication for delivery is established in this study. The investigator will evaluate fetal baseline conditions at enrollment and will decide the timing of delivery based on this fetal indication. Infants will be followed up for development until 1.5 years of age.
Ethics and dissemination: This study was approved by the Institutional Review Board of Mie University Hospital and each participating institution. Our findings will be widely disseminated through peer-reviewed publications.
Trial registration: UMIN000023778."

2. The inclusion criteria (EFW<1.5 SD) is too wide. That will include by definition about 15% of the population, with most cases corresponding to constitutional smallness, where treatment is unnecessary and the ethical principle of "more good than harm" is doubtful to me. A Doppler based criteria (in accordance with the Delphi criteria for FGR) would have yield a more clinically homogeneous population.
Because absent or reversed end-diastolic flow in the umbilical artery is associated with an increased risk of perinatal mortality*, we use these findings as a measure for evaluation of fetal well-being in this

study (Figure 1 and Table 1). We will be able to examine the placenta to assess the causes of FGR of the enrolled cases in the other clinical study (UMIN Clinical Trials Registry: UMIN000028772). We believe that we will be able to perform future studies to evaluate the effects of tadalafil treatment for FGR by a more focused population approach, if we analyze the results of this phase II study together with the data of the causes of FGR.
*American College of Obstetricians and Gynecologists. ACOG Practice bulletin no. 134: fetal growth restriction. Obstet Gynecol 2013;121:1122–33.

3. I found intriguing and a source of selection bias the exclusión criterion: "The investigator decides that entry is inappropriate".
Regarding exclusion criteria No.9 “The investigator decides that entry is inappropriate”, this study excludes mothers with mental or psychiatric problems, since poor judgment capabilities that are often associated with such conditions may not be compatible with Inclusion Criterion No. 6 “Signed written informed consent”, as we have described in the Rationale for Eligibility Criteria section in the Study protocol on page 13.
We added the following sentence in the Patient selection section,
“**Regarding exclusion criteria No.9 “The investigator decides that entry is inappropriate”, this study excludes mothers with mental or psychiatric problems, since poor judgment capabilities that are often associated with such conditions may not be compatible with Inclusion Criterion No. 6 “Signed written informed consent.”

4. The stopping criteria c and d, are somewhat arbitrary and a source of bias. I would recommend an intention-to-treat analysis. Per protocol analysis is known to result in overoptimistic results.
We really appreciate your suggestion. We will add an intention-to-treat analysis to the Study Protocol in the next revision.

5. In their managing protocol I missed the DV: there´s compelling evidence that DV plus CTG improves the perinatal outcome (TRUFFLE study, Evidence level I). Thus, that will be limitation to the external validity of their results...
Although TRUFFLE study reported that DV plus CTG improves the perinatal outcome, the guidelines for obstetrical practice in Japan have not strictly recommended DV yet. However, the investigator will enter the patients’ data, including fetal testing practices, into the Case Report Form on the website of this clinical trial.

6. It is also a limitation that no placebo was planned. Estimated fetal weight is subjected to operator-related variability. I could not tell whether operators will be blinded to the study group. For sure, women will not be blinded.
This study is the first nation-wide intervention study in the field of obstetrics in Japan. We selected an open-label study design with a strict fetal management algorithm on the basis of the results from the multicenter Japanese survey instead of a placebo-controlled design because of operational challenges including low acceptability by pregnant women in Japan.
We added the following sentences in the Discussion section,
“This study is the first nation-wide intervention study in the field of obstetrics in Japan. We selected an open-label study design with a strict fetal management algorithm on the basis of the results from the multicenter Japanese survey instead of a placebo-controlled design because of operational challenges including low acceptability by pregnant women in Japan.”

Reviewer #2
Dr. J.W. Ganzevoort and Dr. A.Pels
We would like to thank Dr. J.W. Ganzevoort and Dr. A.Pels very much for taking time to comment on our paper.
1. Many authors are from a single centre. It would be good to have neonatologists involved as well.

Because we have received advice on the protocol of pediatric developmental assessment from Dr. Hirofumi Sawada (Department of Pediatrics, Anesthesiology, and Critical Care Medicine, Mie University Graduate School of Medicine), we added his contribution to this protocol in the Acknowledgements section.

2. Regarding safety, it would be good to present an overview about the safety of tadalafil in pregnancy, since it does not seem to be frequently investigated before.
As we have described in the manuscript, the Food and Drug Administration in the United States has rated tadalafil as pregnancy category B (presumed safety based on animal studies). Ladouceur et al. reported pregnancy outcomes in patients with pulmonary arterial hypertension associated with congenital heart disease treated with tadalafil. They did not point out adverse effects associated with tadalafil. Doimon et al. also reported no side effects of tadalafil on mothers or offsprings in cases with pulmonary arterial hypertension treated with tadalafil.
We added the following sentences in the Introduction section,
“Ladouceur et al. reported pregnancy outcomes in patients with pulmonary arterial hypertension associated with congenital heart disease treated with tadalafil. They did not point out adverse effects associated with tadalafil.[17] Doimon et al. also reported no side effects of tadalafil on mothers or offsprings in cases with pulmonary arterial hypertension treated with tadalafil.[18]”
We added two references to the Reference section,
17 Ladouceur M, Benoit L, Radojevic J, et al. Pregnancy outcomes in patients with pulmonary arterial hypertension associated with congenital heart disease. Heart Br Card Soc 2017;103:287–92. doi:10.1136/heartjnl-2016-310003
18 Daimon A, Kamiya CA, Iwanaga N, et al. Management of pulmonary vasodilator therapy in three pregnancies with pulmonary arterial hypertension. J Obstet Gynaecol Res 2017;43:935–938.

3. Why was chosen for an open-label design versus placebo-controlled design?
This study is the first nation-wide intervention study in the field of obstetrics in Japan. We selected an open-label study design with a strict fetal management algorithm on the basis of the results from the multicenter Japanese survey instead of a placebo-controlled design because of operational challenges including low acceptability by pregnant women in Japan.
We added the following sentences in the Discussion section,
“This study is the first nation-wide intervention study in the field of obstetrics in Japan. We selected an open-label study design with a strict fetal management algorithm on the basis of the results from the multicenter Japanese survey instead of a placebo-controlled design because of operational challenges including low acceptability by pregnant women in Japan.”

4. Regarding patient selection: Since doppler abnormalities are not included as inclusion criteria, other causes of fetal growth restriction might be included as well. AEDF and REDF are difficult to include, since it could be an indication for delivery.
Because absent or reversed end-diastolic flow in the umbilical artery is associated with an increased risk of perinatal mortality*, we use these findings as a measure for evaluation of fetal well-being in this study (Figure 1 and Table 1). We will be able to examine the placenta to assess the causes of FGR of the enrolled cases in the other clinical study (UMIN Clinical Trials Registry: UMIN000028772).
*American College of Obstetricians and Gynecologists. ACOG Practice bulletin no. 134: fetal growth restriction. Obstet Gynecol 2013;121:1122–33.

5. In the protocol is stated that antepartum fetal testing will be performed to assess fetal wellbeing. What are the differences in fetal testing before and after viability?
As we have described in the Study Protocol, fetal well-being is evaluated by ultrasonography, including Doppler imaging of umbilical arterial blood flow, non-stress test, contraction stress test, and biophysical profile scoring depending on gestational age. The investigator will enter the patients’ data, including fetal testing practices, into the Case Report Form on the website of this clinical trial.

6. In case of intrauterine death, how will the primary outcome be defined?
The primary outcome in case of intrauterine death is calculated using the following formula:
Fetal growth velocity (g/day)
＝(Birthweight – Estimated fetal weight at the first day of the treatment [g])/(Days of the treatment [days])

7. Page 10: what is the timepoint where the secondary endpoints fetal head circumference, deepest amniotic fluid pocket, doppler imaging of umbilical arterial blood flow are assessed?
Fetal head circumference, deepest amniotic fluid pocket, and doppler imaging of umbilical arterial blood flow are evaluated according to the flow chart as shown in the Secondary endpoints section.
We added the following sentence and flow chart in the Secondary endpoints section,
“iv) Fetal head circumference, vi) deepest amniotic fluid pocket, and v) doppler imaging of umbilical arterial blood flow are evaluated according to the flow chart as shown below."

8. Page 10: Why is preeclampsia no secondary outcome?
We will assess “Incidence rate of pre-eclampsia” as secondary outcome, as we have mentioned in the Secondary endpoints section.

9. Page 11: How will the pediatric development be assessed? Which neurodevelopmental test will be performed?
We will assess the neurological development of the enrolled infants using the Kyoto Scale of Psychological Development* in the other clinical study (UMIN Clinical Trials Registry: UMIN000028773).
*Kato T, Mandai T, Iwatani S, et al. Extremely preterm infants small for gestational age are at risk for motor impairment at 3 years corrected age. Brain Dev 2016;38:188–95.

10. Page 11: Will the Safety Committee have blinded or unblinded access to the data?
Because the Safety Evaluation Committee is independent from research organization, it will have unblinded access to the data.

11. Page 13: What is the detectable alternative you would like to prove with the study? Which difference in fetal growth velocity is the power calculation based on?
We retrospectively analyzed 11 Japanese singleton pregnant women with FGR who received tadalafil along with conventional management for FGR at Mie University Hospital from July 2015 to February 2016 (tadalafil group). These women were matched for maternal age, parity, gestational age (GA), and estimated fetal weight at enrollment with 14 singleton pregnant women who had received only the conventional management for FGR in 2014 (conventional management group). The conventional management for FGR was performed according to the guidelines for obstetric practice in Japan. This retrospective study showed that both fetal growth velocity from enrollment to birth and birth weight were significantly higher in the tadalafil group than in the conventional management group (Tadalafil group 17.7 (10.6 to 23.0) g/day vs. conventional management group 12.8 (0 to 17.2) g/day. Data were presented as median (interquartile range). Statistical significance ($P<0.05$) was calculated by the Mann-Whitney U test)*. As we have mentioned in the Rationale for the Target Sample Size section, the power calculation was performed based on this data.
*Kubo M, Umekawa T, Maekawa Y, et al. Retrospective study of tadalafil for fetal growth restriction: Impact on maternal and perinatal outcomes. J Obstet Gynaecol Res 2017;43:291–297.

12. Page 13: A per protocol analysis is an analysis in which the patients who received the treatment they were allocated for, are included. I’m not sure what you mean by the ‘per protocol set’, but it seems to me that you mean the ‘intention to treat’ analysis here, if your aim is to include all randomized patients in this analysis.

We really appreciate your suggestion. We will add an intention-to-treat analysis to the Study Protocol in the next revision.

13. Page 13: It would be good to mention the IQR for the median as well.
We appreciate your suggestion. We added “interquartile range” in the Statistical analysis section.

14. Page 19: The survival rates of VLBW and premature children seems quite high to me, e.g. 49% survival at AD 22 weeks and 401-500 gram. It would be good to add some information about how many patients these data are based on.
The multicenter survey of very low birth weight (VLBW) infants in Japan consisted of 5855 VLBW infants. 50 infants who weighed between 401 and 500 gram at birth were born at 22 weeks of GA. 26 cases of these infants were died in neonatal intensive care units. The infant survival rate data acquired from the survey were preprocessed with the moving average method.

15. It might be useful for the authors to reach out to the STRIDER collaboration group to aim to align trial design as much as possible in this phase to be able to come to meaningful meta-analysis.
Why do the authors not aim for data sharing once published?
We appreciate your suggestion. When the chief researcher receives requests after publication, we will prepare for data sharing to the extent permitted by the Japanese ethics guidelines.

We would like to exclude the URL of the Clinical Trial Data Management System from this manuscript for the security reasons. Thank you very much for your consideration.

We cite the supplementary file just before the Discussion section of the main text. Thank you very much.

VERSION 2 – REVIEW

REVIEWER	Francesc Figueras Hospital Clinic. University of Barcelona.
REVIEW RETURNED	12-Feb-2018

GENERAL COMMENTS	The authors have reasonably answered my queries and concerns.

REVIEWER	Wessel Ganzevoort AMC Amsterdam, Netherlands
REVIEW RETURNED	20-Feb-2018

GENERAL COMMENTS	This is an important protocol, and it is important to get this published. Nonetheless I have significant objections to the present form that I have raised in first review and that have not been adequately dealt with or discussed. • It is good that the authors now acknowledge the significant limitations of their open-label design and share their motivations. They should be more elaborate and reference qualitative studies supporting the statement and explaining which cultural facets do not allow placebo-controlled design. • The authors should be more explicit that in fact they are including SGA rather than FGR, because they do not require Doppler abnormalities as inclusion criteria, therefore not excluding other causes of fetal growth restriction such as congenital anomalies. • The Japanese consensus management with delivering all AEDF and REDF are difficult, it is not uniform international practice these ar an indication for delivery. The authors should discuss this

	limitation. • There are no explicit guidelines regarding antepartum fetal testing in the protocol to assess fetal wellbeing before and after viability. • The authors have not answered how the primary outcome will be defined in case of fetal death. • Preeclampsia or hypertensive disorders of pregnancy should be an outcome. • There is yet no description of how pediatric development will be assessed. Which neurodevelopmental test will be performed? By whom? • Page 11: Will the Safety Committee have blinded or unblinded access to the data? • Which difference in fetal growth velocity is the power calculation based on? • Page 13: A per protocol analysis is an analysis in which the patients who received the treatment they were allocated for, are included. I'm not sure what you mean by the 'per protocol set', but it seems to me that you mean the 'intention to treat' analysis here, if your aim is to include all randomized patients in this analysis. • Page 19: The survival rates of VLBW and premature children seems quite high to me, e.g. 49% survival at AD 22 weeks and 401-500 gram. It would be good to add some information about how many patients these data are based on and perhaps to reference these. • Why do the authors not aim for datasharing once published? Modern ethics require or support international collaboration to do the most possible with data, and would consider not sharing the data after the initial publication unethical.

VERSION 2 – AUTHOR RESPONSE

Reviewer:1
Reviewer Name: Francesc Figueras
Institution and Country: Hospital Clinic. University of Barcelona, Spain
Competing Interests: None

Thank you for accepting our revision.

Reviewer: 2
Reviewer Name: Wessel Ganzevoort
Institution and Country: AMC Amsterdam, Netherlands
Competing Interests: None declared (PI for sildenafil trial Dutch STRIDER)

It is good that the authors now acknowledge the significant limitations of their open-label design and share their motivations. They should be more elaborate and reference qualitative studies supporting the statement and explaining which cultural facets do not allow placebo-controlled design.
● Among Japanese phase 2 and 3 trials, 27.2% are randomized double-blind trials and 7.6% are placebo-controlled double-blind trials※. There are major reasons for the low rate of placebo-controlled double-blind trials in Japan. As a nation, Japan has the tendency to decline administering drugs with unclear efficacies. Second, the Japan Medical Association opposed the production a guideline from a misinterpretation in the year 2000. We introduce two recent open-label design trials in Japan.
1. Aoyagi R, et al. Study protocol for a phase III multicentre, randomised, open-label, blinded-end

point trial to evaluate the efficacy and safety of immunoglobulin plus cyclosporin A in patients with severe Kawasaki disease (KAICA Trial).BMJ Open. 2015 1;5:e009562. doi: 10.1136/bmjopen-2015-009562.
2. Tsurudome Y, et al. Immunogenicity and safety of an inactivated quadrivalent influenza vaccine in healthy adults: a phase II, open-label, uncontrolled trial in Japan. Microbiol Immunol. 2015;59:597-604.

※ Tsutani K. Lessons from the EBM movement in Japan: to avoid repeating past mistakes. National Institute for Educational Policy Research Bulletin. 2011;140:45-54.

The authors should be more explicit that in fact they are including SGA rather than FGR, because they do not require Doppler abnormalities as inclusion criteria, therefore not excluding other causes of fetal growth restriction such as congenital anomalies.
• Thank you for pointing out this important concern. It is possible that SGA was included among cases of FGR without Doppler abnormalities in this study. Therefore, we have added this in the Limitation section as follow: “It is possible that SGA was included among the cases of FGR without Doppler abnormalities in this study.” On the other hand, we think that FGR with Doppler abnormalities has a progressive disease status and does not have much treatment success with tadalafil administration.

The Japanese consensus management with delivering all AEDF and REDF are difficult, it is not uniform international practice these are an indication for delivery. The authors should discuss this limitation.
There are no explicit guidelines regarding antepartum fetal testing in the protocol to assess fetal wellbeing before and after viability.
● No consensus has been reached worldwide regarding the indications for delivery of AEDF and REDF, as suggested by the reviewer. AEDF and REDF appear around 20 days before IUFD in FGR ※. However, we think that it is important to standardize the criterion of delivery in this study. In addition, we introduce two international practices these are an indication for delivery1, 2.
※ Turan OM, Turan S, Gungor S, et al. Progression of Doppler abnormalities in intrauterine growth restriction. Ultrasound Obstet Gynecol. 2008 Aug;32(2):160-7.
1. Lausman A, Kingdom J1; MATERNAL FETAL MEDICINE COMMITTEE. Intrauterine growth restriction: screening, diagnosis, and management.J Obstet Gynaecol Can. 2013;35:741-748.
2. Institute of Obstetricians and Gynaecologists,Royal College of Physicians of Ireland and Directorate of Clinical Strategy and Programmes,Health Service Executive. CLINICAL PRACTICE GUIDELINE FETAL GROWTH RESTRICTION - RECOGNITION, DIAGNOSIS & MANAGEMENT. https://www.hse.ie/eng/services/publications/clinical-strategy-and-programmes/fetal-growth-restriction.pdf

The authors have not answered how the primary outcome will be defined in case of fetal death.
● Thank you for your comment. We have excluded the cases of fetal death in the primary outcome. We have added this in the Primary Endpoint section.

Preeclampsia or hypertensive disorders of pregnancy should be an outcome.
● We have previously added preeclampsia or hypertension disorder in the outcomes.

There is yet no description of how pediatric development will be assessed. Which neurodevelopmental test will be performed? By whom?
● In the neurodevelopment test in this study, the Kyoto Scale of Psychological Development 2001※ was used. Evaluation of neurodevelopment was performed by a pediatric neurologist.
※ Aoki S, Hashimoto K, Ikeda N, et al. Comparison of the Kyoto Scale of Psychological Development 2001 with the parent-rated Kinder Infant Development Scale (KIDS). Brain Dev. 2016;38:481-90.

Page 11: Will the Safety Committee have blinded or unblinded access to the data?
● We have added the relevant sentence in the Safety Evaluation Committee section as follows, as suggested by the reviewer:
“The safety committee had blind access to the data.”

Which difference in fetal growth velocity is the power calculation based on?
● Thank you for your comment. The power calculation of fetal growth velocity was based on the phase 1 trial.

Page 13: A per protocol analysis is an analysis in which the patients who received the treatment they were allocated for, are included. I’m not sure what you mean by the ‘per protocol set’, but it seems to me that you mean the ‘intention to treat’ analysis here, if your aim is to include all randomized patients in this analysis.
● Thank you for your comment. We have changed “per protocol set” to “full analysis set.”

Page 19: The survival rates of VLBW and premature children seems quite high to me, e.g. 49% survival at AD 22 weeks and 401-500 gram. It would be good to add some information about how many patients these data are based on and perhaps to reference these.
● The survival rates of VLBW are data from the Japan Clinical Study Network (total number: 5855 cases, 2003–2007). Unfortunately, these data are unpublished.

Why do the authors not aim for datasharing once published? Modern ethics require or support international collaboration to do the most possible with data, and would consider not sharing the data after the initial publication unethical.
● Thank you for pointing out this important concern. We hope to collaborate with the STRIDER group. Please share the data after the initial publication.

VERSION 3 – REVIEW

REVIEWER	Wessel Ganzevoort AMC Netherlands
REVIEW RETURNED	16-Apr-2018

GENERAL COMMENTS	This is my third revision. I thank the authors for their constructiveness. Most of my previous points were addressed appropriately. I still feel there are a couple of important issues that would improve the paper if improved. - They have now decided to leave out fetal deaths after inclusion the study, or so it seems. They should make it clear this is part of the composite outcome because it is an important diagnosis that can happen after randomization and fetal deaths should not be excluded. -They have described the statistical test at 1.5year follow-up. Please describe in discussion how this corresponds with other more mainstream evaluation tools for international literature such as Bayleys or Griffiths.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Wessel Ganzevoort
Institution and Country: AMC Netherlands
Competing Interests: PI of Dutch STRIDER

This is my third revision. I thank the authors for their constructiveness.
Most of my previous points were addressed appropriately. I still feel there are a couple of important issues that would improve the paper if improved.

- They have now decided to leave out fetal deaths after inclusion the study, or so it seems. They should make it clear this is part of the composite outcome because it is an important diagnosis that can happen after randomization and fetal deaths should not be excluded.
We have changed from ‘excluded’ to ‘included’.

-They have described the statistical test at 1.5year follow-up. Please describe in discussion how this corresponds with other more mainstream evaluation tools for international literature such as Bayleys or Griffiths.
Thank you for pointing it out. Bayleys or Griffiths was used as a mainstream evaluation tool in Japan. Here, we selected ‘the Kyoto Scale of Psychological Development 2001’, which is currently used as a mainstream evaluation tool in Japan. Also, ‘The Kyoto Scale of Psychological Development 2001’ is correlated with Bayleys or Griffiths.
※ Aoki S, Hashimoto K, Ikeda N, Takekoh M, Fujiwara T, Morisaki N, Mezawa H, Tachibana Y, Ohya Y. Comparison of the Kyoto Scale of Psychological Development 2001 with the parent-rated Kinder Infant Development Scale (KIDS). Brain Dev. 2016;38:481-90.