PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Fetal Alcohol Spectrum Disorder and attention deficit hyperactivity disorder Stimulant Trial in children: an n-of-1 pilot trial to compare stimulant to placebo (FASST): Protocol
AUTHORS	Crichton, Alison; Harris, Katrina; McGree, James; Nikles, Jane; Anderson, Peter; Williams, Katrina

VERSION 1 – REVIEW

REVIEWER	Kravitz, Richard University of California Davis
REVIEW RETURNED	12-Jan-2023

GENERAL COMMENTS	This is a detailed and well-articulated protocol for an n-of-1 trial series investigating the use of stimulants for children with fetal alcohol syndrome. My main concern is that the rationale for conducting an n-of-1 trial series rather than a parallel group trial is incompletely described. Usually n-of-1 trials are indicated when there is a strong likelihood of substantial HTE (heterogeneity of treatment effects). This should be addressed. In addition, allowing participants to take their usual stimulant (methylphenidate vs dexamphetamine) is pragmatic, but the plan for assessing differences in treatment efficacy are unclear. How do we know that each child is on the optimal choice of drug (for them)? Finally, it seems that this trial may be most appealing to families that question whether the prescribed stimulant is actually helping. This may further limit enrollment and generalizability, and should be mentioned as a limitation. Other issues are minor. Table 1 is confusing for two reasons: 1) it's not clear that sequence 1 and 2 are just 2 possible sequences of 16; and 2) sequence 1 may contain a typo as there are only 3 active treatment periods not 4. In the analysis section, there may be utility in running parallel Bayesian analyses on the n-of-1 series, allowing estimates of the probability of a clinically significant benefit on treatment vs placebo.

REVIEWER	Graham, Tanya King's College London, Florence Nightingale Faculty of Nursing, Midwifery and Palliative Care
REVIEW RETURNED	08-Feb-2023

GENERAL COMMENTS	Fetal Alcohol Spectrum Disorder and attention deficit hyperactivity disorder Stimulant Trial in children: an n-of-1 pilot trial to compare stimulant to placebo (FASST): Protocol Reviewers comments; This is an important study addressing medication for children with

	Fetal Alcohol Spectrum Disorder and ADHD. I have some minor points: Page 20 - I could not see any plans for managing missing data or non-adherence (other than removal from the trial) specifically in the analysis section. Page 20 - are there plans for an interim analysis to see if the trial can indeed needs to run for 8 weeks? May have clinically significant results earlier than 8 weeks/stop the trial earlier than planned? Page 21 and Page 23 regarding patient involvement. There was only one patient representative 'consulted' - was this a parent or a child? What changes were made because of the consultation? Consulting only one person is rather tokenistic if there is a way to involve more parents/families/children in the trial going forward that would be advisable. Page 21 line 45 does not support the notion of patients being involved in treatment decisions. Page 23 lines 25-32 is more in line with the notion of shared decision making which complements meaningful patient involvement in the trial. Using the data to inform treatment decisions is advocated as one of the key strengths of the N-of-1 design (see Duan et al 2013 below) In this way, results from N-of-1 design can be incorporated into shared decision-making during consultations thereby being a powerful tool to by which to integrate patient knowledge and preference into treatment decisions. You have stated that 'Participants and their primary paediatrician will receive a report detailing individual participant's response to the stimulants compared to placebo, to facilitate a decision on further treatment.' I would make this more explicit to include the word shared decision making. If you can collect qualitative data on this aspect of trial design/implementation it will make the results paper far more meaningful in my opinion. But I understand a process evaluation is not included at this stage. The manuscript provided sufficient details as outlined by the Spirit guideline – but would have been useful to also report if the SPIRIT extension and elaboration for n-of-1 trials: SPENT 2019 checklist was also used to identify aspects specific to n-of-1 trial methodology. Best regards Dr Tanya Graham Duan, N., R.L. Kravitz, and C.H. Schmid, Single-patient (n-of-1) trials: a pragmatic clinical decision methodology for patient-centered comparative effectiveness research. Journal of Clinical Epidemiology, 2013. 66(8, Supplement): p. S21-S28.

REVIEWER	García-Algar, O. Hospital Clínic, Neonatology
REVIEW RETURNED	06-May-2023

GENERAL COMMENTS	The subject of this protocol is very interesting and a unsolved question. The protocol is complicated amb perhaps an more clear algorith will aid to understand it completely.

REVIEWER	O'Neill, Joseph UCLA, Child Psychiatry
REVIEW RETURNED	08-May-2023

GENERAL COMMENTS	1. This MS is not a report of a completed study. It is an experimental protocol for a planned study. The investigators propose a clinical trial

	of neurostimulant medication for the ADHD-symptoms of pediatric fetal alcohol spectrum disorders (FASD). They will use an “N-of-1” design. This does not mean that they will be doing a case study, they will actually study 20 FASD patients of either sex aged 4-18. But each patient will be treated as a trial unto him- or herself. Results will be analyzed both for individual patients and across the collective. Each 8-week individual trial will consist of 4 2-week periods. During the first week of each period, the patient will receive either active drug (A) or placebo (P) Monday thru Friday with both patient and clinicians blinded. Then there will be a medication holiday on the weekend. (Note that the half-life of stimulants is quite short, which favors rapid washout.) For Monday thru Friday of the second week, the patient will receive placebo or drug, i.e., will cross over. Which treatment goes first will be randomized within each patient for each period. Thus, the treatment regimen will have one of 16 possible sequences APAPAPAP, PAPAPAPA, APPAAPAP,… The investigators will only enroll patients who have already been taking stimulant medication for at least 1 month prescribed by an outside doc. The two permitted agents (both oral) are methylphenidate and dextroamphetamine. Patients currently taking other classes of ADHD-meds (e.g., atomoxetine, guanfacine) will be excluded. The dose will be the same dose that the patient is already taking. This is a frequent feature of N-of-1 research and clinical practice: the physician DCs or switches an existing med in order to find out whether it is “really working”. Hence, this style of prescriptive practice greatly exploits the ability of controlled longitudinal interventions to establish cause-and-effect. Apart from feasibility and tolerability, there will be several assessments, including the Teachers and Parents Conners 3 to evaluate severity of core ADHD symptoms, the Top Problems Assessment (TPA; a caregiver rating of ADHD challenges) and selected tests from the CANTAB. The investigators have conducted a power analysis and prepared a statistical approach suitable for an N-of-1 design. 2. Generally, I do not advocate pre-trial publication of study protocols. Such papers clutter an already inflated scientific literature. With respect to N-of-1 designs, I am agnostic. Having said that, I nonetheless welcome this manuscript. There is a pressing need for more research aimed at identifying effective treatments for FASD. With respect to stimulants, for years it has been a highly significant underinvestigated open question whether or not they work for ADHD in FASD. Thus, the present trial is to be greeted with open arms. 3. The N-of-1 design should not be taken as a reason to reject the manuscript. Particularly in Australia, N-of-1 designs have been applied with some success in the past. N-of-1 is a defensible alternative to the evidence-based model for many scenarios, including the present one. While evidence-based models are concerned with maximizing mean response across a group and with minimizing variation in therapeutic practice, the N-of-1 approach (much in the spirit of personalized medicine and precision medicine) focusses on maximizing clinical response in the individual patient. 4. The manuscript states that FASD will be diagnosed according to the criteria of the Australian Guide to the diagnosis of FASD. Presumably, they mean Bower et al. (2017). If so, this should be cited in Table 2 and/or in the text. More importantly, the main criteria should be listed explicitly in the text. 5. Table 2 states that children “at risk” for FASD will be included. It

	might be better to exclude these children. Per Bower et al., the “at risk” category potentially includes some children without confirmed FASD. The projected sample size of N=20 is small to begin with and the sample will probably contain considerable heterogeneity as it is. Even a series of N-of-1 trials should take some measures against unnecessarily diluting the sample in advance. That could lead to issues later when trying to publish the results. 6. Side effects will be evaluated. The authors should include a list of the most likely side effects anticipated. 7. Both methylphenidate and dextroamphetamine will be tested. The investigators should contemplate restricting the study to methylphenidate only, again for the sake of a cleaner study. If not, the manuscript needs to discuss how many patients are expected to be on methylphenidate and how many on dextroamphetamine. And how that will be dealt with statistically. 8. Will extended release formulations be included? Again, maybe better to exclude for the sake of uniformity. 9. Teacher ratings on the Conners are to be the primary outcome with parent ratings a secondary outcome. It might be wiser to make the parent ratings the primary outcome. The authors cite a paper claiming that teacher ratings more accurately predict FASD diagnosis over parent ratings—but we’re not primarily concerned with diagnosis here. In the past, our group at least has experienced poor compliance amongst teachers in reporting ratings. In contrast, we typically find parents or guardians of children with FASD to be highly motivated to participate in research, especially treatment research. The wrong choice could possibly undermine the trial. 10. The manuscript has a tendency to cite reviews rather than primary sources, e.g., instead of the well-known O’Malley et al. (2000) retrospective study of stimulants in FASD, they cite the (high-quality) review of Peadon & Elliott. Or, the review of Popova et al. is cited rather than primary epidemiological sources on the prevalence of FASD. 11. Is the statement in the text, “…cleared from the body from between 35 – 2.5 days”, an error? 12. It appears that the study will draw patients pre-diagnosed with FASD from “VicFAS”. That should be made clear in the text. Is VicFAS a patient database from the state of Victoria or is it something else? Appropriate description and/or references should be made in the text.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1
Comment Response
My main concern is that the rationale for conducting an n-of-1 trial series rather than a parallel group trial is incompletely described. Usually n-of-1 trials are indicated when there is a strong likelihood of substantial HTE (heterogeneity of treatment effects). This should be addressed. We agree with this comment and have addressed this in the manuscript. Data on variability of stimulant effect in FASD

population included. See ‘Rationale for trial’, “Furthermore, existing studies in the FASD population have found significant within-subject variability (F = 4.02 df 4, p <0.01) potentially obscuring group treatment effects. Oesterheld, 1998” p. 7
In addition, allowing participants to take their usual stimulant (methylphenidate vs dexamphetamine) is pragmatic, but the plan for assessing differences in treatment efficacy are unclear. How do we know that each child is on the optimal choice of drug (for them)? The dose is the individualized dose titrated by treating clinician. The study data will aim to inform the question of whether their current stimulant medication and dose is optimal for them. If participants have no or some benefit and continue to have high behavioural ratings, for example, they could subsequently explore, via their treating Paediatrician whether further dose optimisation or trial of a different stimulant drug option would be worthwhile. Statement added: ”Individualised efficacy data could be used to explore whether further dose optimisation or trial of a different stimulant drug option would be worthwhile.”, p.9.
Finally, it seems that this trial may be most appealing to families that question whether the prescribed stimulant is actually helping. This may further limit enrolment and generalizability, and should be mentioned as a limitation. We have amended the ‘strengths and limitations’ to ensure this is clear. See amended ‘strengths and limitations’ points, p. 5.

Table 1 is confusing for two reasons: 1) it's not clear that sequence 1 and 2 are just 2 possible sequences of 16; and 2) sequence 1 may contain a typo as there are only 3 active treatment periods not 4. Thank you for this constructive feedback. The table has been revised to clarify the sequence, with words ‘example sequences’, and …[Sequ 16] added for clarity. See Table 2: Schedule of assessments for n-of-1 trial.
In the analysis section, there may be utility in running parallel Bayesian analyses on the n-of-1 series, allowing estimates of the probability of a clinically significant benefit on treatment vs placebo. The authors agree that there is potential benefit in running parallel Bayesian analyses on the n-of-1 series. Accordingly, this has now been included in the Analysis section. This statement has been added: “Bayesian analyses on the n-of-1 series will also be considered, allowing for the probability of a clinically significant difference between medication relative to placebo conditions to be estimated. For these analyses, weakly informative prior information (calibrated by prior predictive checks) will be used such that the conclusions are essentially data driven”. See Analysis section, p. 18.

Reviewer: 2
Comment Response
Page 20 - I could not see any plans for managing missing data or non-adherence (other than removal from the trial) specifically in the analysis section Thank you for this comment, and the omission of this was an oversight and has now been included. The heading Handing of missing data section was added with: “The occurrence of missing data will be reported. These occurrences will be explored to assess whether there are any patterns in the missingness. If concerns arise about such patterns, the potential impact of missing data will be explored via single value imputation, implementing a best-worst-case and worst-best-case sensitivity analysis. Within the Bayesian analyses, standard Bayesian imputation will be implemented such that the uncertainty in the missing data will be accounted for when evaluating the treatment effect.” (see p.14)
Page 20 - are there plans for an interim analysis to see if the trial can indeed needs to run for 8 weeks? May have clinically significant results earlier than 8 weeks/stop the trial earlier than planned? We do not plan to do an interim analysis. Although there are some benefits to adaptive trial design, they require resources beyond the scope of this trial including complicated trial design and statistician input. Furthermore, there are practical constraints to this approach within our proposed trial timeframe. For example, if interim analysis was undertaken after week 4, as a potential time where adaptions could be made, it would probably take around 2 weeks for a decision to be made about adapting (i.e. data cleaning, data analysis, reporting, review by team, decision). This would mean there is potentially only benefit in the last 2 weeks of the trial for a participant. Given the significant additional costs and resource requirements for the limited benefit, we did not feel this was warranted in this case. However, the finding will provide estimates that can be used to more accurately estimate future trial length and sample size (observations) for future trials. N/A
Page 21 and Page 23 regarding patient involvement. There was only one patient representative 'consulted' - was this a parent or a child? What changes were made because of the consultation? Consulting only one person is rather tokenistic if there is a way to involve more parents/families/children in the trial going forward that would be advisable. The consumer was a parent of a child with FASD. We acknowledge the extent of consumer consultation is small, and due funding and timing for this trial further consultation was beyond the scope of the current trial. There is

now a broader research framework within the clinic, led by the study PI, that includes broader consumer engagement via an expression of interest in future projects, including a larger trial of stimulants for children with FASD. We have added the specific changes that our consumer representative assisted with. The following comment has been added: “Key areas of input and subsequent protocol revision included planning a washout day on Sunday rather than Monday for long acting formulations, minimize school disruption, reducing questionnaire length to reduce overall carer burden, and revisions to increase useability of the online platform (RedCAP).” (see pp. 22-23)
Page 21 line 45 does not support the notion of patients being involved in treatment decisions. Page 23 lines 25-32 is more in line with the notion of shared decision making which complements meaningful patient involvement in the trial. Using the data to inform treatment decisions is advocated as one of the key strengths of the N-of-1 design (see Duan et al 2013 below) In this way, results from N-of-1 design can be incorporated into shared decision-making during consultations thereby being a powerful tool to by which to integrate patient knowledge and preference into treatment decisions. You have stated that 'Participants and their primary paediatrician will receive a report detailing individual participant’s response to the stimulants compared to placebo, to facilitate a decision on further treatment.'
I would make this more explicit to include the word shared decision making. If you can collect qualitative data on this aspect of trial design/implementation it will make the results paper far more meaningful in my opinion. But I understand a process evaluation is not included at this stage.
Duan, N., R.L. Kravitz, and C.H. Schmid, Single-patient (n-of-1) trials: a pragmatic clinical decision methodology for patient-centered comparative effectiveness research. Journal of Clinical Epidemiology, 2013. 66(8, Supplement): p. S21-S28. We would like to thank the reviewer for their considered feedback. We agree with this suggestion, and have made changes to the manuscript to reflect shared decision making more accurately. Text revised to “Recording post-trial stimulant prescribing decision made by the participant’s carer and prescribing paediatrician after receiving N-of-1 trial data.”p.8, and the participant information statement was updated to reflect this (see Supplemental Material 2).
The manuscript provided sufficient details as outlined by the Spirit guideline – but would have been useful to also report if the SPIRIT extension and elaboration for n-of-1 trials: SPENT 2019 checklist was also used to identify aspects specific to n-of-1 trial methodology. Thank you for this advice and direction to the SPENT checklist. This has now been included as Supplemental Material 1. See SPENT Checklist, Supplemental Material 1.

Reviewer: 3
Comment Response
The subject of this protocol is very interesting and a unsolved question.
The protocol is complicated and perhaps an more clear algorith will aid to understand it completely.
Thank you for this feedback. We have welcomed specific suggestions to improve clarify through the protocol manuscript in responding to the reviewers comments, and revised the Table 1 (Trial sequence) for clarity.

Reviewer: 4
Comment Response

With respect to stimulants, for years it has been a highly significant underinvestigated open question whether or not they work for ADHD in FASD. Thus, the present trial is to be greeted with open arms.
Thank you for this comment.
The N-of-1 design should not be taken as a reason to reject the manuscript. Particularly in Australia, N-of-1 designs have been applied with some success in the past. N-of-1 is a defensible alternative to the evidence-based model for many scenarios, including the present one. While evidence-based models are concerned with maximizing mean response across a group and with minimizing variation in therapeutic practice, the N-of-1 approach (much in the spirit of personalized medicine and precision medicine) focusses on maximizing clinical response in the individual patient. We would like to thank reviewer 4 for their positive feedback.
The manuscript states that FASD will be diagnosed according to the criteria of the Australian Guide to the diagnosis of FASD. Presumably, they mean Bower et al. (2017). If so, this should be cited in Table 2 and/or in the text. More importantly, the main criteria should be listed explicitly in the text
Thank you for this observation. Each participant will complete FASD diagnostic assessment according to Australian Guide to the diagnosis of FASD. The Guide states the correct citing is Bower et al 2016,

and this has been added to Table 2.
Please refer to the 'secondary outcomes' where this was already stated in text (p.13) as follows: "As per the Australian Guide to the Diagnosis of FASD (Bower et al 2016) baseline demographics (child age, sex), comorbidities (developmental diagnoses including ADHD, ASD, learning difficulty) and neurodevelopmental functioning will be obtained through the VicFAS research database and/or medical record, with consent. Neurodevelopmental impairment across the 10 domains assessed for the purposes of FASD diagnostic assessments will be categorised according to level of impairment (none, mild, moderate, severe as per the FASD Guide).(Bower, 2016).
If this requires further clarification we would be happy to revise this. Note added to Table 2 as follows: FASD diagnostic assessment = as per Australian guide to the diagnosis of FASD (Bower et al, 2016).
Table 2 states that children "at risk" for FASD will be included. It might be better to exclude these children. Per Bower et al., the "at risk" category potentially includes some children without confirmed FASD. The projected sample size of N=20 is small to begin with and the sample will probably contain considerable heterogeneity as it is. Even a series of N-of-1 trials should take some measures against unnecessarily diluting the sample in advance. That could lead to issues later when trying to publish the results. While we agree that strict inclusion and exclusion criteria help reduce sample heterogeneity, we chose to include children at risk of FASD for three key reasons. First, inclusion of FASD, 'at risk' of FASD and children with confirmed PAE is common practice in Australian research studies (1), and is common practice in FASD research that participants are included who have confirmed heavy PAE (see research from the CIFASD group e.g. Sarah Mattson, Julie Kable or Claire Coles), such as (2, 3). Second, this group of children represent a continuum of impairments resulting from PAE, and diagnostic terminology and criteria vary between countries, and of these differing criteria, Australian guidelines specify the highest degree of neurodevelopmental impairment. Importantly, we have clarified that all children required a confirmed prenatal alcohol exposure for inclusion in the study (as part of admission to the VicFAS clinic). This has been more clearly articulated in the manuscript (Table 1, inclusion criteria). See Table 1 for addition of inclusion criteria: "Have confirmed prenatal alcohol exposure (PAE)."
Side effects will be evaluated. The authors should include a list of the most likely side effects anticipated. Thank you for this comment. Side effects are commonly reported across MHP and DEX. Efron et al (1997) reported on the side effects in a paediatric sample with ADHD across these stimulant drugs on the measure of interest, and this is used to estimate the most likely side effects now included in the text. The side effects are outlined to the potential participants (as stipulated by human ethics requirements) in the participant information and consent form – 'What are the possible risks, side-effects, discomforts and/or inconveniences?' Text included: "The most common side effects reported for Methylphenidate and Dexamphetamine on this measure in Australian samples are decreased appetite, sleep difficulties, not happy and overly meticulous behaviour (mild or moderate severity).(4)", p. 12
See Supplemental Material 2 - participant information and consent form – 'What are the possible risks, side-effects, discomforts and/or inconveniences?
Both methylphenidate and dextroamphetamine will be tested. The investigators should contemplate restricting the study to methylphenidate only, again for the sake of a cleaner study. If not, the manuscript needs to discuss how many patients are expected to be on methylphenidate and how many on dextroamphetamine. And how that will be dealt with statistically. Response: The trial is a taken from a clinical sample of children seen through the VicFAS clinic, who are on the VicFAS database. Initial, pilot data (first 12 months) provided the following rates of simulant medication: Of children who were prescribed stimulants for treatment of ADHD, the most common stimulants were Methylphenidate (SA, 68%), Methylphenidate (LA, 22%), Dexmethylphenidate (10%) (with non-stimulants (guanfacine) at around 17%). All stimulant medications were included in order to: (1) increase potential sample size, (2) be inclusive of those interested in the trial due to potential to benefit from trial inclusion in post-trial clinical decision making; (3) provide pilot data to inform future larger trials, including data from different stimulant medication types. This was considered feasible as the trial design accommodates the range of pharmacokinetic characteristics of all stimulants in the trial. Importantly, there is no convincing evidence from comparison among stimulants (mainly methylphenidate and amphetamines) that one class outperformed the other in terms of ADHD symptom control, such that effects should not be pooled.(5) Studies comparing different formulations of the same drug revealed no significant differences in terms of symptom control (5), suggesting it is feasible to pool data for the current study purposes, while providing pilot data to inform future larger trials, including data from different stimulant medication types. Furthermore, given we anticipate that the size of groups for each medication would small, the study would not be sufficiently powered to compare effects between stimulant type. The protocol has been amended to articulate the rationale

for inclusion of methylphenidate or dexamphetamine (long or short acting formulations) as follows "The existing trial is not sufficiently powered to examine stimulant to stimulant differences between methylphenidate and dexamphetamine due to the small overall sample size and anticipated number in each medication group. We will report the treatment effects for both methylphenidate and dexamphetamine separately as exploratory analysis in order to providing pilot data to inform future larger trials, including data from different stimulant medication types See comment: "The existing trial is not sufficiently powered to examine stimulant to stimulant differences between methylphenidate and dexamphetamine due to the small overall sample size and anticipated number in each medication group. Importantly, studies comparing different formulations of the same drug revealed no significant differences in terms of symptom control (5), suggesting it is feasible to pool data for the current study purposes, while providing pilot data to inform future larger trials, including data from different stimulant medication types", p.18

Will extended release formulations be included? Again, maybe better to exclude for the sake of uniformity. Children were included if they were prescribed extended release formulations. There is a precedent in the literature to include all stimulant types and formulations to assist in generalisability to the clinical population. See Nikles et al 2014 for precedent of this method in N-of-1 stimulant trials, in which Children already on long-acting MPH were offered long acting MPH trials at their clinician's discretion. See analysis section: "The existing trial is not sufficiently powered to examine stimulant to stimulant differences between methylphenidate and dexamphetamine due to the small overall sample size and anticipated number in each medication group", p. 18.

Teacher ratings on the Conners are to be the primary outcome with parent ratings a secondary outcome. It might be wiser to make the parent ratings the primary outcome. The authors cite a paper claiming that teacher ratings more accurately predict FASD diagnosis over parent ratings—but we're not primarily concerned with diagnosis here. In the past, our group at least has experienced poor compliance amongst teachers in reporting ratings. In contrast, we typically find parents or guardians of children with FASD to be highly motivated to participate in research, especially treatment research. The wrong choice could possibly undermine the trial. Thank you for this insightful comments. The FASST Trial cohort will be selected from participants already in the VicFAS database. Through this database, it has been observed that in the FASD cohort, there is generally fewer missing data in the teacher questionnaires relative to carer/parent. In the VicFAS cohort at the time of recruitment, 81% of participants were in out of home care; this is in contrast to the 3.74 % in the Conners3 normative sample.(6) Therefore, for this cohort, teacher ratings were collected as the primary outcome, and carer/parent as secondary.

The manuscript has a tendency to cite reviews rather than primary sources, e.g., instead of the well-known O'Malley et al. (2000) retrospective study of stimulants in FASD, they cite the (high-quality) review of Peadon & Elliott. Or, the review of Popova et al. is cited rather than primary epidemiological sources on the prevalence of FASD. Thank you for this constructive feedback. The original sources (O'Malley et al, 2000; Lange et al , 2017) have been cited in the revised manuscript. See References

Is the statement in the text, "…cleared from the body from between 35 – 2.5 days", an error? Thank you for this observation. There was an error which is now corrected to read 35 hours – 2.5 days. See p.9

It appears that the study will draw patients pre-diagnosed with FASD from "VicFAS". That should be made clear in the text. Is VicFAS a patient database from the state of Victoria or is it something else? Appropriate description and/or references should be made in the text. The reviewer is correct in their understanding. Additional detail has been provided about the VicFAS clinic and database, including inclusion criteria (confirmed PAE) in Table 1 for greater clarity. PAE confirmation, as a condition of inclusion to the VicFAS clinic and research database has been added to Table 1. Text added: "The VicFAS Database was established as a prospective study by Dr Crichton as principal investigator in 2019 and captures core data on children seen for FASD diagnostic assessment from August 2019 onwards, against the Australian Guide to the Diagnosis of FASD", p.19

Key references in response:

1. Young SL, Gallo LA, Brookes DSK, Hayes N, Maloney M, Liddle K, et al. Altered bone and body composition in children and adolescents with confirmed prenatal alcohol exposure. Bone. 2022;164:116510.

2. Lee KT, Mattson SN, Riley EP. Classifying children with heavy prenatal alcohol exposure using measures of attention. J Int Neuropsychol Soc. 2004;10(2):271-7.

3. Mattson SN, Roesch SC, Fagerlund A, Autti-Rämö I, Jones KL, May PA, et al. Toward a neurobehavioral profile of fetal alcohol spectrum disorders. Alcohol Clin Exp Res. 2010;34(9):1640-50.

4. Efron D, Jarman F, Barker M. Side Effects of Methylphenidate and Dexamphetamine in Children With Attention Deficit Hyperactivity Disorder: A Double-blind, Crossover Trial. Pediatrics. 1997;100(4):662.
5. Brown RT, Amler RW, Freeman WS, Perrin JM, Stein MT, Feldman HM, et al. Treatment of Attention-Deficit/Hyperactivity Disorder: Overview of the Evidence. Pediatrics. 2005;115(6):e749-e57.
6. Gallant S, Conners CK, Rzepa SR, Pitkanen J, Marocco M, Sitarenios G. Psychometric properties of the Conners 3. Poster presented at the annual meeting of the American Psychological Association2007.

VERSION 2 – REVIEW

REVIEWER	O'Neill, Joseph UCLA, Child Psychiatry
REVIEW RETURNED	14-Aug-2023

GENERAL COMMENTS	The authors have replied adequately to my critiques.