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) | Vocal Brain Development in Infants of Mothers with Serious Mental Illness (CAPRI-Voc) - study protocol |
|---------------------|--------------------------------------------------------------------------------------------------|
| AUTHORS             | Stibbs-Eaton, Lucy; Hodgson, Catherine; Kolade, Adekeye; Crowell, Jennifer; Gemignani, Jessica; Hope, Holly; Pierce, Matthias; Elmadih, Alya; Zhao, Chen; Downey, Darragh; Elliott, Rebecca; Abel, Kathryn |

GENERAL COMMENTS

One of the criteria of a protocol paper in BMJ Open is that the study is ongoing. However, the sample size section in this paper indicates that recruitment and testing runs from Sept 2017 – April 2021. Has testing already finished? Perhaps it is only recruitment that occurs in this window and the longitudinal follow ups are still ongoing? Or the pandemic interrupted the study? I think further detail is required here to ensure this publication is valid for this journal, if all testing is already complete then I am not sure that it is eligible under BMJ Open’s guidelines.

Could further detail of the recruitment process be provided in the manuscript please? In the Intro the study is described as a national study based in Manchester, implying that recruitment would be nationwide, however the recruitment section just mentions leaflets in local areas in Manchester. Do the team expect that the CAPRI-Voc social media pages would recruit nationally? And can you be more specific about what these are?

The protocol mentions that for any parent who wishes, the testing will be done in the home. Could further information be provided about the protocol for testing in the home please. How is the equipment transported? What ethical considerations and protocol measures are there for testing in the home for both the participants and researchers? How are these managed? If they were at home for one of the visits, was this kept consistent throughout their involvement in the study or might the measures be undertaken in different settings within an individual? Was this taken into account in the analyses? (For PCI, Bayleys and fNIRS)

Given that some baseline data is allowed for a protocol paper perhaps the authors could provide a summary of what proportion of the sample had visits at home to date.
It is unclear why the protocol was designed with a parent-infant interaction video only conducted at 9mths, and a Bayleys only conducted at the two follow-ups. How do both of these measures fit into the hypotheses, and could the paper elaborate on why they were conducted in this order?

Further to this, the protocol makes no mention of how this PCI and Bayley data will be used in relation to the fNIRS data or other data on parent and environment, except for in the dissemination section which seems an odd choice. Please expand on this in the protocol and analysis sections and explain how this data relates to specific targeted outcomes and relates back to the Aims.

The fNIRS data analysis section could be expanded further to overview the current plans for statistical analysis following data pre-processing. How will the hypotheses be tested, which approach will be used for analysis? On page 19 you mention adding a score of SMI and LofA into a model, but you have not previously mentioned your intention to model data to look at significant results, can you elaborate on plans for modelling further here?

Minor comments:
1) Avoid the term abnormal development (as used in abstract), as this is not commonly used in psychology now, instead perhaps use atypical development.
2) Line 21, pg6 – should read “number”
3) Line 49, pg6 - In this phrase “fNIRS is a relatively new functional imaging technique, particularly suited for use in infants, due to safety (emission of visible light) and the smaller surface area of infant heads” I would suggest that the most important factor regarding fNIRS being ideally suited for infants is its ease of use, portability, tolerance of movement and non-invasive nature. Not the fact that the light is visible or the infant head has a smaller surface area.
4) Line 51, pg6 - It is not “near and infrared” but “near infrared”
5) Line 54, pg6 – Chromophores are molecules not tissues
6) Line 17, pg 11 – can you clarify how long before a session the questionnaires are sent. Also these are paper forms sent in post/email or online?
7) Line 21, pg 17 - Could you clarify further what ‘a’ means? For the vocal stimuli in first 9 min session.
8) Line 48, pg 18 – error in following sentence, should be comma and not a full stop: General linear models appropriate to the distribution of micromolar changes in HbO and HHb will be used to assess differences between groups. age, childhood trauma, substance abuse and social adversity (e.g. income, employment).
9) Line 3, pg 19 – an extra “and” in sentence: 4 5 will combine potential sources of adversity
long-term language development challenges for infants with mothers who have severe mental illness. This innovative approach, should it be sensitive enough to detect early biomarkers, will provide an important tool to identify the need for preventive early intervention for language development in this population.

The strengths of this manuscript include a strong rationale for an early detection measure for this area of need, clear explanation of fNIRS with infants and how the study team has piloted it with healthy infants and mothers, clear description and outline (Table 1) of study measures and modification of measures for the control group, use of measures that account for neurological processing (fNIRS) and behavioral parameters of outcomes, and detailed description of fNIRS procedures, data analysis and how it can assess infant vocal processing, ethics training of study team, and the use of mothers to inform future research with regards to recruitment, burden of time and emotional strain, as well as the inclusion of mothers in the dissemination of study processes and outcomes throughout the study in addition to the traditional paths of dissemination taken by researchers.

As a reviewer there were several things that could be revised prior to publication. They are as follows:

- There are several places that the authors use acronyms or vernacular that may not be clear or known for international readers. Use of the acronym SMI first appears without identifying its full meaning, assumed by this reviewer as severe mental illness (see p. 7, line 25 in Study Aims and Hypotheses #2). In Table 1 Instrument Details column for the measure Infant Medical Notes and Growth Trajectories (p. 13, line 36) the authors state “Using information from the ‘little red book’ to record…” but it is not clear what the ‘little red book’ is. As BMJ is an international journal it could be helpful to use a more universal term or descriptor to represent the ‘little red book.’ Finally, SOP is an acronym used in the ethical considerations section. This reviewer assumes it refers to Standard Operating Procedures, but clarification of the acronyms full meaning should be used in its first appearance (p. 18, line 27).

- The study dates for recruitment and data collection are identified as September 2017 until April 2021 (p. 8, line 13 & p. 19, line 6). The latter half of the recruitment and data collection falls within the global pandemic time period that changed human subjects research access and progress. This reviewer believes it is important to explicitly address any ways in which this study protocol was altered, delayed or otherwise changed in response the pandemic, particularly since this was a home visit protocol.

- The authors hypothesize an interaction between adversity and parental mental illness. The authors are encouraged to clearly define the parameters considered as adversity (e.g. poverty), in addition to parental mental health or if adversity is to be discovered as part of the data collection to explain that process (p. 7, line 42, 43).

- There are a large number of measures for mothers to complete during the study time period. Efforts to offset the level of burden were described as offering compensation at the end of the study period and using medical records for mothers with SMI. However,
it was unclear how the medical records could be used to complete aspects of the measures outlined in Table 1. It is recommended the authors be more explicit with how the medical records will decrease the level of burden on the mothers and if any other actions were taken to support the completion of the measures.

Under ethical considerations (p. 18, line 39), the authors state that they will “escalate if necessary” in the event no contact is made by the study team with the researcher who is at a participants’ home collecting data; however, it is unclear what actions are in place as representative of escalate. Although it is likely described in the SOP, a brief description of what escalating actions mean in this instance would be reassuring and helpful for future researchers conducting home-based data collection with vulnerable populations.

Thank you for the opportunity to review this study protocol.

**VERSION 1 – AUTHOR RESPONSE**

Reviewer: 1
Dr. Sarah Lloyd-Fox, University of Cambridge Comments to the Author:

3. One of the criteria of a protocol paper in BMJ Open is that the study is ongoing. However, the sample size section in this paper indicates that recruitment and testing runs from Sept 2017 – April 2021. Has testing already finished? Perhaps it is only recruitment that occurs in this window and the longitudinal follow ups are still ongoing? Or the pandemic interrupted the study? I think further detail is required here to ensure this publication is valid for this journal, if all testing is already complete then I am not sure that it is eligible under BMJ Open’s guidelines.
- This comment is resolved in the text, located via comment number

4. Could further detail of the recruitment process be provided in the manuscript please? In the Intro the study is described as a national study based in Manchester, implying that recruitment would be nationwide, however the recruitment section just mentions leaflets in local areas in Manchester. Do the team expect that the CAPRI-Voc social media pages would recruit nationally? And can you be more specific about what these are?
- The SMI cohort are recruited nationally through NHS mental health trust sites. The control cohort are mainly recruited from the Greater Manchester areas as that is the reach of the advertising campaign. Both of these procedures are detailed in the ‘Participant and Recruitment Procedures’ section.

5. The protocol mentions that for any parent who wishes, the testing will be done in the home. Could further information be provided about the protocol for testing in the home please. How is the equipment transported? What ethical considerations and protocol measures are there for testing in the home for both the participants and researchers? How are these managed? . If they were at home for one of the visits, was this kept consistent throughout their involvement in the study or might the measures be undertaken in different settings within an individual? Was this taken into account in the analyses? (For PCI, Bayleys and fNIRS)
- This comment is resolved in the text, located via comment number. In addition, the ethical considerations are addressed in the ‘ethical considerations’ section'
6. Given that some baseline data is allowed for a protocol paper perhaps the authors could provide a summary of what proportion of the sample had visits at home to date.
- This comment is resolved in the text, located via comment number

7. It is unclear why the protocol was designed with a parent-infant interaction video only conducted at 9mths, and a Bayleys only conducted at the two follow-ups. How do both of these measures fit into the hypotheses, and could the paper elaborate on why they were conducted in this order?
- This comment is resolved in the text, located via comment number

8. Further to this, the protocol makes no mention of how this PCI and Bayley data will be used in relation to the fNIRS data or other data on parent and environment, except for in the dissemination section which seems an odd choice. Please expand on this in the protocol and analysis sections and explain how this data relates to specific targeted outcomes and relates back to the Aims.
- This comment is resolved in the text, located via comment number

9. The fNIRS data analysis section could be expanded further to overview the current plans for statistical analysis following data pre-processing. How will the hypotheses be tested, which approach will be used for analysis? On page 19 you mention adding a score of SMI and LofA into a model, but you have not previously mentioned your intention to model data to look at significant results, can you elaborate on plans for modelling further here?
- This comment is resolved in the text, located via comment number

Minor comments:
a) Avoid the term abnormal development (as used in abstract), as this is not commonly used in psychology now, instead perhaps use atypical development.
b) Line 21, pg6 – should read “number”
c) Line 49, pg6 - In this phrase “fNIRS is a relatively new functional imaging technique, particularly suited for use in infants, due to safety (emission of visible light) and the smaller surface area of infant heads” I would suggest that the most important factor regarding fNIRS being ideally suited for infants is its ease of use, portability, tolerance of movement and non-invasive nature. Not the fact that the light is visible or the infant head has a smaller surface area.
d) Line 51, pg6 - It is not “near and infrared” but “near infrared”
e) Line 54, pg6 – Chromophores are molecules not tissues
f) Line 17, pg 11 – can you clarify how long before a session the questionnaires are sent. Also these are paper forms sent in post/email or online?
g) Line 21, pg 17 - Could you clarify further what ‘a’ means? For the vocal stimuli in first 9 min session.
h) Line 48, pg 18 – error in following sentence, should be comma and not a full stop: General linear models appropriate to the distribution of micromolar changes in HbO and HHb will be used to assess differences between groups. age, childhood trauma, substance abuse and social adversity (e.g. income, employment).
i) Line 3, pg 19 – an extra “and” in sentence: 4
5 will combine potential sources of adversity
- All minor comments resolved in the text and located via comment letter

Reviewer: 2
Dr. Deanna Hanson-Abromeit , The University of Kansas Comments to the Author:

This article outlines a study protocol that examines the feasibility of noninvasive imaging using fNIRS with infants of mothers with severe mental illness to identify vocal stimuli processing compared to infants of mothers without severe mental illness. This study hopes to determine if fNIRS can be used
as an early detector of long-term language development challenges for infants with mothers who have severe mental illness. This innovative approach, should it be sensitive enough to detect early biomarkers, will provide an important tool to identify the need for preventive early intervention for language development in this population.

The strengths of this manuscript include a strong rationale for an early detection measure for this area of need, clear explanation of fNIRS with infants and how the study team has piloted it with healthy infants and mothers, clear description and outline (Table 1) of study measures and modification of measures for the control group, use of measures that account for neurological processing (fNIRS) and behavioral parameters of outcomes, and detailed description of fNIRS procedures, data analysis and how it can assess infant vocal processing, ethics training of study team, and the use of mothers to inform future research with regards to recruitment, burden of time and emotional strain, as well as the inclusion of mothers in the dissemination of study processes and outcomes throughout the study in addition to the traditional paths of dissemination taken by researchers.

As a reviewer there were several things that could be revised prior to publication. They are as follows:

10. There are several places that the authors use acronyms or vernacular that may not be clear or known for international readers. Use of the acronym SMI first appears without identifying its full meaning, assumed by this reviewer as severe mental illness (see p. 7, line 25 in Study Aims and Hypotheses #2). In Table 1 Instrument Details column for the measure Infant Medical Notes and Growth Trajectories (p. 13, line 36) the authors state “Using information from the ‘little red book’ to record…” but it is not clear what the ‘little red book’ is. As BMJ is an international journal it could be helpful to use a more universal term or descriptor to represent the ‘little red book.’ Finally, SOP is an acronym used in the ethical considerations section. This reviewer assumes it refers to Standard Operating Procedures, but clarification of the acronyms full meaning should be used in its first appearance (p. 18, line 27).
- This comment is resolved in the text, located via comment number

11. The study dates for recruitment and data collection are identified as September 2017 until April 2021 (p. 8, line 13 & p. 19, line 6). The latter half of the recruitment and data collection falls within the global pandemic time period that changed human subjects research access and progress. This reviewer believes it is important to explicitly address any ways in which this study protocol was altered, delayed or otherwise changed in response the pandemic, particularly since this was a home visit protocol.
- This comment is resolved in the text, located via comment number

12. The authors hypothesize an interaction between adversity and parental mental illness. The authors are encouraged to clearly define the parameters considered as adversity (e.g. poverty), in addition to parental mental health or if adversity is to be discovered as part of the data collection to explain that process (p. 7, line 42, 43).
- This comment is resolved in the text, located via comment number

13. There are a large number of measures for mothers to complete during the study time period. Efforts to offset the level of burden were described as offering compensation at the end of the study period and using medical records for mothers with SMI. However, it was unclear how the medical records could be used to complete aspects of the measures outlined in Table 1. It is recommended the authors be more explicit with how the medical records will decrease the level of burden on the mothers and if any other actions were taken to support the completion of the measures.
- This comment is resolved in the text, located via comment number
14. Under ethical considerations (p. 18, line 39), the authors state that they will “escalate if necessary” in the event no contact is made by the study team with the researcher who is at a participants’ home collecting data; however, it is unclear what actions are in place as representative of escalate. Although it is likely described in the SOP, a brief description of what escalating actions mean in this instance would be reassuring and helpful for future researchers conducting home-based data collection with vulnerable populations.

- This comment is resolved in the text, located via comment number

VERSION 2 – REVIEW

| REVIEWER                 | Hanson-Abromeit, Deanna                                           |
|--------------------------|---------------------------------------------------------------------|
| RESEARCH INSTITUTION     | The University of Kansas, Music Education & Music Therapy          |
| REVIEW RETURNED          | 05-Jan-2022                                                        |

GENERAL COMMENTS

Thank you for your careful attention to the recommendations for revision made by the reviewers. There are 2 minor edits I noticed in the revision.

• Sample recruitment dates have been revised to account for the pandemic, stated as September 2017-October 2022 in the Sample Size section for recruitment and testing (p. 5) and September 2017 until July 2022 in the Study Status section on recruitment and data collection (p. 19). Please clarify the incongruency of dates and correct the end date (either October or July 2022).

• Bottom of p. 18 – last sentence of statistical analysis section, the word significant should be significance (“…will be set using the p-value threshold…”).