ARTICLE DETAILS

TITLE (PROVISIONAL)
The Australian Child Maltreatment Study (ACMS): Protocol for a national survey of the prevalence of child abuse and neglect, associated mental disorders and physical health problems, and burden of disease

AUTHORS
Mathews, Ben; PACELLA, ROSANA; Michael, Dunne; Scott, James; Finkelhor, David; Meinck, Franziska; Higgins, Daryl; Erskine, Holly; Thomas, Hannah J.; Haslam, Divna; Tran, Nam; Le, Ha; Honey, Nikki; Kellard, Karen; Lawrence, David

GENERAL COMMENTS
This is a strong protocol for a national survey of the prevalence of child maltreatment and related outcomes in Australia. The protocol is clear and well-written, while the proposed survey and methods are rigorous, considered, and in line with transparent and open research practices. This is important work for advancing child maltreatment research.

I have a handful of more major methodological considerations, as well as some minor points, which I outline below.

Main points: Methodological considerations

1. I find the use of language around “over/through the lifespan” problematic (e.g., see lines 19, 24, 52). The proposed survey is cross-sectional and thus cannot measure how psychopathology and other problems change across the life-course. I understand that participants will be of varying ages and that this may be informative in determining different associations across age groups. However, comparing such associations speaks to cohort effects rather than developmental effects of maltreatment over an individual’s lifespan, as the current language implies. In other words, the proposed survey design may help to examine differences between age groups but not within individuals over time. Please amend this language to reflect the cross-sectional nature of the survey and its related limitations.

2. The authors intend to sample from the population aged 16 years and over but the proposed recall period for adversities and maltreatment during childhood spans up to 18 years old. I appreciate the different rationalities for choosing these different age ranges however I think it raises a number of potential problems – methodologically and practically.
Methodologically, this renders the survey potentially concurrent for participants aged 16 to 18 years, rather than retrospective, as for the other age groups. This means that some participants will be recalling over a shorter time period for the exposure to have occurred (i.e., 16 years vs 18 years) which means that measure itself is invariant across some sub-populations within the study’s sample. This might incur biases, for example, underrepresentation of sexual abuse for the younger participants. It would be good to see how the authors intend on addressing this in their planned analyses, especially as they intend to oversample from the 16-24 age group.

Practically, this also raises concerns around safeguarding as the survey may identify concurrent cases of maltreatment. The authors do provide an important section on participant safety on page 10. But it would be helpful to supplement this section with more specifics around the safeguarding of participants aged under 18 years old who are identified to be currently experiencing maltreatment.

3. I have some concerns regarding the ability of the Juvenile Victimization Questionnaire to accurately and validly measure the specifics of maltreatment experiences. Specifically, relating to the frequency or duration and timing (age of onset and cessation) of maltreatment experiences. As the authors acknowledge on page 11, retrospective self-report is subject to recall bias and inaccuracy and are notoriously unreliably when it comes to measuring accurate frequency and timings. Accuracy is often related to the temporal distance from the experience itself, again raising some potential measurement invariance biases between the different age groups. I would like to see the protocol consider and try to mitigate these issues.

4. Relatedly, the authors referenced Baldwin et al 2019: a seminal systematic review and meta-analysis adding to the growing evidence that retrospective self-report and official records of maltreatment (e.g., police or child protection services) capture largely different – though overlapping – groups. More recent evidence indicates that these different groups are also important in relation to subsequent risk (Danese & Widom, 2020; doi:10.1038/s41562-020-0880-3). The proposed survey would therefore be greatly strengthened by including a measure relating to official records of child maltreatment (e.g., through linkage with administrative data) or even (though sub-optimal) by including a self-report item on contact with child protection services. Not only would this help to further validate the self-report measures of maltreatment, but it would mean that the survey had the ability to contribute to this important and growing area of research.

5. A nationally representative survey of this size is an incredible amount of work and its value in understanding, preventing, and responding to child maltreatment is undeniable. As evidenced in other countries, US (e.g., Finkelhor’s research) and UK (e.g., Radford’s NSPCC work) the ongoing value of such a cross-sectional survey can be extended even further through repeated data collection at other point(s) in the future. I appreciate that such a replication may not be possible due to funding and resource restrictions. Nevertheless, it would be important to see if and how the authors had considered this possibility and tailored the methods to facilitate a potential replication in the future.
Minor points: Methodological
• Survey instrument:
  o What is the approximate duration (or range of duration) of the questionnaire?
  o Just to clarify, will all questionnaires be conducted by trained interviewers and administered via telephone only, specifically CATI?
• Validation:
  o Cognitive testing; page 8 lines 46-52: “Alternative phrasing and response frames were trialled to improve comprehension and speed of response. Findings supported further refinements to ensure results will be valid, reliable, and complete.” This section reads quite vaguely, may the authors provide an example of a refinement and how the refinements improved the results?
• Pilot data processing and analysis:
  o Page 9, lines 10-22: Again, the wording feels vague in some instances. For example, “missing data was minimal”, “very good reliability”, “alpha was strong”, “Estimates of prevalence... were within expected ranges”, “Few participants found the survey upsetting”. Please provide figures and specifics for these statements.
• Planned analysis:
  o Page 9, lines 45-46: “Occurrence by age, gender and other strata”. What “other strata” are these? Specific socio-demographic characteristics? I could not easily find what socio-demographics would be collected in the survey; this is important information to include in the protocol.
  o Page 9, lines 47-49: “We will use multiple imputation to deal with missing data”. Please be more specific on this methodology (e.g., how many imputed datasets? multiple imputation via chained equations? etc).
  o Page 9, lines 56-59: “Outcomes significantly associated with maltreatment will be analysed using multivariate analyses controlling for demographic characteristics...” I would disagree with this approach of only fully analysing “significant associations”. I would encourage authors to conduct analyses based on previous literature and theory rather than by basing them on statistical significance.
  o Page 10, lines 3-10: the authors refer to logistic regressions and Odds Ratios, presumably indicating that maltreatment will be categorised and analysed as a binary variable. This seems a loss given the emphasis on measuring a variety of dimensions of maltreatment. Either way however, it would be important to detail at this stage how these binary categories will be defined.
• Limitations, page 11, lines 15-21: I was pleased to see a discussion about selection/participation bias by relying on telephones for recruitment. Give this is an anticipated bias of the sample (potentially making it non-representative sample), will authors take additional steps at the recruitment stage to ensure that these sub-populations (eg homeless, Indigenous Australians etc) are sampled. For example, engaging with specific communities or via oversampling. Obtaining data from these populations is more preferable than having to statistical adjust for this bias at a later stage.

Minor points: Other
• Introduction:
  o It would be helpful for the authors to include a brief definition of how they are defining maltreatment at the beginning of their
introduction. As in the introduction they mention “5 types of maltreatment” and this isn’t necessarily clear given varying definitions of maltreatment.

- Page 4, lines 55-60: “There is no Australian evidence of the nature, prevalence and timing of mental disorders and physical health outcomes associated with child maltreatment, or of other associated health and behavioural outcomes”. I would disagree with this statement and therefore recommend softening this language. For example, there are a handful of important Australian cohorts and case-control studies that have provided relevant evidence on this, including The Mater-University Study of Pregnancy (MUSP), as well as case-control studies, including follow-up of children from Child Protection Units in Sydney (PI: Heather Swanston) and Victoria.

- Limitations

- Page 10/11: “causal relationships between child maltreatment and mental disorders, self-harm and substance use are well established”. Again, I would soften this wording. There is certainly well-established correlational evidence but, due to ethical and practical considerations limiting RCTs (cf. Bucharest Romania adoption Study), I would dispute that there was well-established causal evidence.

- I think it would be beneficial to add a few lines in the limitations section that address the strengths of “behaviour-specific” survey questions in relation to changing definitions of maltreatment and parenting norms over time (especially given the different wide range of age groups being sampled in this study).

REVIEWER
Kimber, Melissa
McMaster University, Department of Psychiatry and Behavioural Neurosciences

REVIEW RETURNED
13-Feb-2021

GENERAL COMMENTS
Peer Review – KIMBER

Manuscript ID: bmjopen-2020-047074

Thank you for the opportunity to review, “The Australian Child Maltreatment Study (ACMS): Protocol for a national survey of the prevalence of child abuse and neglect, associated mental disorders and physical health problems and burden of disease” for publication in BMJ Open. I would like to take the opportunity to share my sincere thanks to the authors and their associated team members for championing such important work in child maltreatment. In addition, I extend my appreciation to the members of the Technical Expert Panel, those who participated in the cognitive testing and test-retest reliability of the survey measures, and most importantly, the four survivors of child maltreatment who reviewed the ACMS items. Given the scope of this research endeavour and the plans to launch the survey in March 2021, I would expect that there is little opportunity to make any changes to the methodology outlined in the protocol. For this reason, I hope that the authors receive my comments with an understanding that much of what I outline below would be helpful to consider – if possible – before the launch of the survey and (if not) in future publications about the ACMS methodology and results.

Comments to consider across the entire protocol:
I suggest using the language of intimate partner violence consistently throughout the abstract, article summary, and body of the manuscript. Using such language would map onto the assessment of “other adverse outcomes” articulated in the methodology and attunes to the importance of recognizing the diversity of genders who are exposed to violence in intimate relationships. Domestic violence is heavily centred on ‘traditional’ notions of heterosexual intimacy and patriarchy. The use of the word ‘sensitive’ to describe research related to child maltreatment in my view, is not necessary. I would suggest that the authors consider the removal of this evocative language throughout the manuscript. There are many concepts, constructs, and exposures that are considered in the ACMS; the interpretation of what is and is not ‘sensitive’ to be asked about is a personal one (e.g., asking about marriage, divorce, sexual activity, alcohol, or drug use, etc) could all be considered sensitive to a given person, family, group, or culture. Its use to describe research on child maltreatment reinforces problematic notions that child maltreatment cannot be asked about in a systematic and respectful way – the proposed study clearly demonstrates (though its own approach and the citation of other work) that in fact, it can be. Suggest that the authors check the ‘tense’ of the entire article. In some cases (e.g., abstract) present vs. past tense are used.

Abstract:

Suggest that the authors add their plan for missing data into the abstract. In addition, can the authors offer the specific strategy that they will use for imputation (e.g, joint modelling vs. fully conditional specification/chained equations)?

Introduction:

First paragraph
suggest moving citations 3-6 to the end of the sentence, as those references cover both mental and physical health consequences of CM.

Can the authors be more specific in what they mean about self-directed violence? (e.g., self-harm and suicide).

Fifth paragraph
The comment that there is “no Australian evidence of the nature, prevalence and timing of mental disorders and physical health outcomes associated with child maltreatment…” is not accurate (which I am positive that the authors know this, given that a number of them are authors on important studies; e.g.,
https://doi.org/10.1016/j.chiabu.2016.12.002;
10.1016/j.childyouth.2016.11.014;
https://doi.org/10.1016/j.chiabu.2015.05.006). Thus, this sentence needs to be revised so that readers who are less familiar with the available evidence in Australia are clear that the evidence in the country is (1) disparate; (2) lacks the voices of young people (especially children/adolescents – which is like most countries); (3) and does not tend to use representative approaches that can characterize population-level incidence, prevalence, correlates, and burdens.

Sixth paragraph
Suggest that the authors offer a definition of ‘corporeal punishment’ that falls in line with policy and practice guidelines within the Australia context.
Also suggest that authors make explicit what the definitions of child maltreatment are in the Australian context and what level/type of
Methods

First paragraph
Please add the source of the sampling frame in the “Sample Selection and Setting” section of the methods. In addition, a further description of the sampling frame should be noted (how does an individual become a part of the sampling frame, who does it miss?)

Description of Survey Instrument Section

Maltreatment Questions Section
Add “(JQV)” following Juvenile Victimization Questionnaire

Third paragraph of this section – I am concerned that the emphasis on physical and sexual abuse for items focused on “whether the participant told anyone” and “to whom the disclosure occurred” perpetuates the notion that other forms of child maltreatment are: (a) less impactful over the short and long term; and (b) less important to respond too/less important to have helpful responses too. As the experts on this team would be aware of – emotional abuse, neglect, and child exposure to IPV can have equally damaging impacts on the mental and physical wellbeing of survivors. I appreciate that there is the need (from a child protection perspective) to prioritize imminent and severe risks to children (which are invoked in physical and sexual abuse disclosure), but there is a real opportunity to leverage data (and recommendations) around disclosure experiences and response experiences by asking these questions to respondents who indicate emotional abuse, neglect, and exposure to IPV. If the authors cannot do this, a justification needs to be added here and this should be appropriately acknowledged as a limitation in papers that report on findings for these variables.

A citation is needed for the MINI, as is clarification of the version being used.

Given the breadth of disorders captured by the MINI, there should be justification here for why the authors have selected these specific mental health conditions when there are strong associations, for example, between CM exposure and other conditions - like eating disorders, which are often chronic, debilitating, and deadly if intervention is not offered early.

Instrument Development and Validation Section

First paragraph:
Suggest replacing “Study” with ACMS here and throughout the remainder of the paper where authors are referring to the ACMS, more generally. Will reduce confusion and make it clear that the authors are only referring to one study, which is the ACMS.

Second paragraph:
I was surprised to see that children’s exposure to IPV is being limited to direct witnessing of IPV behaviours and I think this relates to how this exposure may be conceptualized related to child protection in the Australian context. This re-iterates the need to clearly define these exposures/experiences in the introduction, the rationale for using those definitions. A rationale for not considering awareness of IPV as a form of exposure should be provided or acknowledged in the limitations (work by Wathen, MacMillan, Hamby, and others discusses the impact of awareness of IPV).

Pilot data processing and analysis
I appreciate that some of the analysis related to the pilot is ongoing, however, given that the likelihood of the protocol being
routinely reviewed and cited by other jurisdictions in the future, there are a few details that I think would be helpful here. For example, can the proportion of data (range for the different modules in the survey) and whether the data was missing at random be added to the protocol?

Main Study

Second paragraph
Please provide the timeframe in which the advance text message to notify potential participants of the survey is provided before the phone call.
Please also clarify if individuals could, after receiving the text message, can: (1) reply to the text and request to complete the survey right then; (2) click on a link in the text and complete the survey electronically (and if not, why not?)

Planned Analysis section:
Please provide an indication if weights will be generated prior to analysis to ensure estimates are representative of the population. Please provide clarity of what type of imputation will be utilized and when this will occur (will it occur for the entire data set or will imputation procedures be completed per analysis focus?)

My sense is that having the expert’s complete imputation procedures PRIOR to releasing the data to other researchers would be important to ensuring consistency in estimates and analytical approaches.

Estimating Disease Burden Attributable to Child Maltreatment
A citation is needed after the description of approach for generating YLLs, YLDs, and DALYS.
Can the authors offer more information about how they will estimate the economic burden of CM?

DISCUSSION

I feel strongly that a brief paragraph, at the start of the discussion, needs to be added to articulate the methodological and substantive contributions of this protocol. This team’s work presents an incredible opportunity to present a ‘standard’ by which this important data can be collected in a robust and respectful way.
Suggest that the Distress Protocol and the Risk of Harm protocol be attached as supplementary files. My apologies if I missed this. Again, this reiterates the need to support the translational impact of this work to other countries.

Can you strengthen the mobile phone administration section of the work by providing an indication of the proportion of individuals in Australia who have access to a mobile phone? In addition (and to address potential concerns around missing Indigenous Australians), one could use geomapping methods to examine the extent to which mobile phone ownership attached to post codes (for example) are located areas with a high concentration of Indigenous peoples.
Suggest removing the use of “mainstream” populations in the paragraph right before the Ethics and Dissemination section.

Ethics and Dissemination

I just want to note my appreciation to the authors for their plan to make the instrument available through a creative commons license.
**Version 1 – Author Response**

Authors’ revisions in response to reviewers’ comments

<table>
<thead>
<tr>
<th>Reviewer 1 Comments</th>
<th>Authors’ response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments to the Author:</td>
<td>We appreciate Reviewer 1’s recognition of the quality of our work, and we are grateful for their deep engagement and thoughtful suggestions.</td>
</tr>
<tr>
<td>This is a strong protocol for a national survey of the prevalence of child maltreatment and related outcomes in Australia. The protocol is clear and well-written, while the proposed survey and methods are rigorous, considered, and in line with transparent and open research practices. This is important work for advancing child maltreatment research. I have a handful of more major methodological considerations, as well as some minor points, which I outline below. Main points: Methodological considerations</td>
<td></td>
</tr>
<tr>
<td>We appreciate Reviewer 1’s recognition of the quality of our work, and we are grateful for their deep engagement and thoughtful suggestions.</td>
<td></td>
</tr>
<tr>
<td><strong>1.</strong> I find the use of language around “over/through the lifespan” problematic (e.g., see lines 19, 24, 52). The proposed survey is cross-sectional and thus cannot measure how psychopathology and other problems change across the life-course. I understand that participants will be of varying ages and that this may be informative in determining different associations across age groups. However, comparing such associations speaks to cohort effects rather than developmental effects of maltreatment over an individual’s lifespan, as the current language implies. In other words, the proposed survey design may help to examine differences between age groups but not within individuals over time. Please amend this language to reflect the cross-sectional nature of the survey and its related limitations.</td>
<td></td>
</tr>
<tr>
<td>Thanks for this comment. As the reviewer notes, our analyses will be able to compare associations across age groups, enabling findings about relative frequency at different stages of life of mental health disorders and physical health problems associated with maltreatment. We agree that while it was not our intention, the term lifespan may suggest a longitudinal developmental approach. We have changed the terminology to “throughout life” (e.g., in the Article Summary, we now state “The study also measures associations between child maltreatment and mental disorders, physical health and health risk behaviours that occur throughout life”, rather than “through the lifespan”). We have also added an acknowledgment of this in the limitations, stating: “Also, given that the ACMS is cross-sectional we are unable to measure individuals’ mental and physical health over the life-course. However, we can compare associations between child maltreatment and different outcomes across different age groups. Despite its limitations this is an appropriate way of being able to estimate the mental and physical health impacts of child maltreatment.”</td>
<td></td>
</tr>
</tbody>
</table>
2. The authors intend to sample from the population aged 16 years and over but the proposed recall period for adversities and maltreatment during childhood spans up to 18 years old. I appreciate the different rationalities for choosing these different age ranges however I think it raises a number of potential problems – methodologically and practically. Methodologically, this renders the survey potentially concurrent for participants aged 16 to 18 years, rather than retrospective, as for the other age groups. This means that some participants will be recalling over a shorter time period for the exposure to have occurred (i.e., 16 years vs 18 years) which means that measure itself is invariant across some sub-populations within the study’s sample. This might incur biases, for example, underrepresentation of sexual abuse for the younger participants. It would be good to see how the authors intend on addressing this in their planned analyses, especially as they intend to oversample from the 16-24 age group.

Our view is that the survey is still retrospective for participants aged 16-17. It is true that compared with other participants aged 18 and over, these participants will be reporting on their experiences over a period slightly less than the full span of childhood, namely age 0-18. We agree that for 16-17 year olds, the results may underrepresent the experience of some types of maltreatment. We note that other prevalence studies with children aged under 18 do not make statistical adjustments to accommodate this, and present estimates of combined samples with an implicit acknowledgment of this limitation. We will generate estimates for the entire sample, but will also examine 16 and 17 year olds separately from 18-24 year olds, and from the entire sample. Data we obtain on the mean age at which the abuse last occurred will also allow us to statistically model patterns of abuse using the data provided by the other group. We have added some material to reflect this in the “Planned analysis” section (Maltreatment prevalence measurement).

Practically, this also raises concerns around safeguarding as the survey may identify concurrent cases of maltreatment. The authors do provide an important section on participant safety on page 10. But it would be helpful to supplement this section with more specifics around the safeguarding of participants aged under 18 years old who are identified to be currently experiencing maltreatment.

As explained below,* we have added a paragraph to the Participant safety section explaining that the required comprehensive treatment of this complex topic will be provided in full elsewhere in forthcoming work.

3. I have some concerns regarding the ability of the Juvenile Victimization Questionnaire to accurately and validly measure the specifics of maltreatment experiences. Specifically, relating to the frequency or duration and timing (age of onset and cessation) of maltreatment experiences. As the authors acknowledge on page 11, retrospective self-report is subject to recall bias and inaccuracy and are notoriously unreliable when it comes to measuring accurate frequency and timings. Accuracy is often related to the temporal distance from the experience itself, again raising some potential measurement invariance biases between the different age groups. I would like to see the protocol consider and try to mitigate these issues.

In project outputs we will discuss at greater length the complex issues related to recall as applied generally to studies of child maltreatment prevalence; we are reluctant to attempt to deal with them here in the protocol manuscript.

However, we are comfortable making two responses to the Reviewer’s concerns, and we have provided some additional text in the second paragraph of the “Limitations” section.

* First, considering the JVQ generally, we are confident in its ability to measure maltreatment experiences with sufficient accuracy and validity, even while
we accept the limitation common to all retrospective studies of recall bias and inaccuracy. The JVQ has been employed repeatedly in large national studies in the U.S with children and youth, in the U.K with children and young adults up to age 24. It has been employed with remarkable success, demonstrated by its consistent results over time. It employs behaviourally-specific screeners and follow-ups that they been subjected to rigorous testing and repeated administration and analysis. The JVQ research team has repeatedly analysed the instrument’s robustness in detail, and while making refinements especially to suit contemporary settings, has not identified major limitations. Therefore, in assessing its robustness generally, we are confident in its soundness.

- Second, in relation to the measurement of frequency of maltreatment and its timing, we can note that we have made some modifications to our adapted JVQ to accommodate the different age span of our sample which extends to adults of more advanced years, to mitigate potential greater recall bias. We did this by providing extra response options for two key follow-ups and tested this approach in piloting. We provide an example of this in the new added text in the revised manuscript (paragraph 2 of the Limitations section).

- Again, we will be considering these types of issues in detail in the separate forthcoming work in which we explain the process of adapting the JVQ for the ACMS. Accordingly, we would prefer not to provide further detailed coverage of this here.
4. Relatedly, the authors referenced Baldwin et al 2019: a seminal systematic review and meta-analysis adding to the growing evidence that retrospective self-report and official records of maltreatment (e.g., police or child protection services) capture largely different – though overlapping – groups. More recent evidence indicates that these different groups are also important in relation to subsequent risk (Danese & Widom, 2020; doi:10.1038/s41562-020-0880-3). The proposed survey would therefore be greatly strengthened by including a measure relating to official records of child maltreatment (e.g., through linkage with administrative data) or even (though sub-optimal) by including a self-report item on contact with child protection services. Not only would this help to further validate the self-report measures of maltreatment, but it would mean that the survey had the ability to contribute to this important and growing area of research.

We can respond first by saying that we do indeed include a question asking participants if they ever had contact with the out-of-home care system, and this will provide some useful although limited data.

We have built into the design of our ACMS design the capacity to conduct subsequent separate studies that explore issues connected with the benchmark ACMS. Namely, we ask participants if they would be willing to be recontacted for the purpose of participating in future connected studies. We envisage that, appropriately designed, such studies could include data linkage studies, which could be particularly useful in considering in more detail health service utilisation, educational outcomes generally, and educational disciplinary outcomes. We have added comments about this to the first paragraph of the Discussion.

However, for several reasons, it is not possible to adopt the suggestion to add a connected data linkage study to this benchmark administration of the ACMS. The main reason for this is that participation in the ACMS is anonymous. We would need to seek consent from participants in the ACMS to access sufficient identifying information (such as birth name) to enable concurrent data linkage and this would almost certainly reduce participation. A second reason is that, connected with this, we did not build into our ACMS design (or its ethical approval) required measures to protect participants in a data linkage study. Third, nor did we conduct preparatory work in eight different States and Territories with data custodians required to conduct such a study, including...
assessing the nature, scope and coverage of linked datasets.

We would also note that while we concur with the Reviewer in acknowledging the general importance of the findings from Baldwin et al., we do not regard a data linkage study using child protection agency child maltreatment records as being able to provide reliable data that can substantially mitigate any limitations regarding prevalence in a retrospective self-report study. There are a number of reasons for this, which are worthy of separate treatment; most significantly, data linkage using official child protection agency records of substantiated maltreatment will provide extremely limited data confirming actual maltreatment, because so few cases of maltreatment come to official agencies’ attention at all, and of those that do, so few are officially substantiated for a range of technical reasons. Accordingly, we do not feel such data linkage would serve to further validate the self-report measures of maltreatment; in fact of the two respective measures, properly-designed self-report measures are far superior (this is acknowledged by several statements in Baldwin et al., 591-592; note, e.g., “the low agreement between prospective and retrospective measures cannot be interpreted to directly indicate poor validity of retrospective measures. For example, prospective measures may have lower sensitivity…and the higher prevalence of retrospective measures could, this, indicate greater ability to identify true cases of childhood maltreatment.”). The limits of child protection agency data on maltreatment are numerous. Moreover, given different approaches to these data in eight different Australian States and Territories, and with further multiple variations occurring over time, this would face insurmountable obstacles.

5. A nationally representative survey of this size is an incredible amount of work and its value in understanding, preventing, and responding to child maltreatment is undeniable. As evidenced in other countries, US (e.g.,
Finkelhor's research) and UK (e.g., Radford's NSPCC work) the ongoing value of such a cross-sectional survey can be extended even further through repeated data collection at other point(s) in the future. I appreciate that such a replication may not be possible due to funding and resource restrictions. Nevertheless, it would be important to see if and how the authors had considered this possibility and tailored the methods to facilitate a potential replication in the future.

Minor points: Methodological

<table>
<thead>
<tr>
<th>Survey instrument:</th>
</tr>
</thead>
<tbody>
<tr>
<td>o What is the approximate duration (or range of duration) of the questionnaire?</td>
</tr>
<tr>
<td>o Just to clarify, will all questionnaires be conducted by trained interviewers and administered via telephone only, specifically CATI?</td>
</tr>
</tbody>
</table>

Survey instrument:

We have added information about the approximate duration (para 1, last line of “Recruitment/procedures” now reads: “Depending on participants’ responses, the average duration of interviews is approximately 30 minutes.”)

We have clarified that all questionnaires will be conducted by trained interviewers and administered via CATI (first line of “Recruitment/procedures” now reads: “All interviews in the ACMS will be conducted by trained interviewers using CATI, as piloted.”)

Validation:

| o Cognitive testing; page 8 lines 46-52: “Alternative phrasing and response frames were trialled to improve comprehension and speed of response. Findings supported further refinements to ensure results will be valid, reliable, and complete.” This section reads quite vaguely, may the authors provide an example of a refinement and how the refinements improved the results? |

Validation:

We have provided examples of refinements made as a result of cognitive testing, adding text to the cognitive testing section to state: “Examples of this include revisions of wording to enhance the clarity of screener questions on generalised sexual harassment and internet sexual victimization.”

Pilot data processing and analysis:

| o Page 9, lines 10-22: Again, the wording feels vague in some instances. For example, “missing data was minimal”, “very good reliability”, “alpha was strong”, “Estimates of prevalence… were within expected ranges”, “Few participants found the survey upsetting”. Please provide figures and specifics for these statements. |

Pilot data processing and analysis:

We have added a statement at the end of the main paragraph in this section to state that: “In forthcoming work we report full details of the process of developing and testing the modified instrument, which will include comprehensive psychometric data.”

We have adopted this approach because before submission of the protocol the research team discussed whether or not to include figures and specifics of the psychometric data. We concluded it was not appropriate in a protocol manuscript to include this data, because: (1) some journals expressly prohibit inclusion of this data in protocol manuscripts and
encourage separate publication of it as a research article; (2) we did not wish to publish pilot data about the experience of maltreatment types since it is not representative and there is a risk of misuse of such data; (3) we conducted an extensive process of instrument configuration and testing and we are writing a separate comprehensive article explaining why this was required, how it was done, why we conducted the range of analyses, the nature and significance of the pilot data, and a full discussion of its implications.

| Planned analysis: |  |
|-------------------|  |
| o Page 9, lines 45-46: “Occurrence by age, gender and other strata”. What “other strata” are these? Specific socio-demographic characteristics? I could not easily find what socio-demographics would be collected in the survey; this is important information to include in the protocol. | We have added a sentence to indicate that depending on cell sizes we anticipate conducting analyses also by ethnicity, socio-economic status, sexuality, and out of home care involvement. |
| o Page 9, lines 47-49: “We will use multiple imputation to deal with missing data”. Please be more specific on this methodology (e.g., how many imputed datasets? multiple imputation via chained equations? etc) | We have provided additional information on the strategy for dealing with missing data. This will be partially data driven. In our experience, most participants in interviewer-led surveys will either not participate in the study at all or will answer most or all questions. Weighting will be used to adjust for complete survey non-response. Where individual items are left missing, the imputation strategy will be based on the extent of the missing data and the nature of the variable. Where the amount of missing data is small (less than 1% of survey responses), the benefit of multiple imputation is trivially small compared to the impact of weighting and overall non-response, so a single random hot-deck imputation will be used. Where the amount of missing data is greater than 1%, a theoretical and empirical assessment will be undertaken of whether it is reasonable to assume missing at random. For instance, it is possible that individuals who do not answer questions about sexual assault may be qualitatively different from those that answer these questions, eg because they feel stigma or embarrassment about the issue. Where missing at random is reasonable to assume, multiple |
imputation will be undertaken using chained equations. For variables where there is reason to suspect systematic bias in refusals, the “don’t know” or “refused” category will be treated as a separate category in the analysis. For transparency of reporting prevalence estimates of types of maltreatment we will report both conservative estimates based on assuming refusals did not suffer maltreatment, as well as estimates produced using the imputation procedures.

The questions selected for inclusion in the survey have all been chosen based on previous literature and theory. We agree that it is appropriate to carry out planned analyses based on existing literature and theory, and we have amended the manuscript accordingly. We note that one of the key reasons for the study is the limited current knowledge and large gaps in the evidence base regarding child maltreatment. Where appropriate, to extend existing theory and assist in developing new theory, exploratory data analysis approaches will be adopted, e.g. to address objective 3: generate new evidence of institutional sexual abuse and physical abuse.

The analysis section has been expanded to more fully cover the types of analyses that are planned. Logistic regression analyses of binary exposure outcomes will also be supplemented by appropriate analyses regarding frequency and severity of maltreatment and multi-type patterns.

For reasons of cost, viability and alignment with primary aims of the ACMS, we are not taking additional steps in this benchmark population survey to recruit or oversample from specific sub-populations.

In the Australian context, as shown by multiple other national studies, the common practice is to employ statistical adjustment for any resulting bias that becomes evident after application of the sample-wide recruitment strategy, rather than at a later stage.
This is also consistent with the advice provided by our partner survey research agency. We have added a reference to support this at line 21.

As indicated by our acknowledgment of this in the Limitations section of this protocol manuscript, in any project outputs we will acknowledge as a limitation any biases in the sample.

We originally included these definitions in the Instrument Development and Validation section and feel they are best placed there.

We have made revisions to qualify this statement in the way suggested, by acknowledging literature about other evidence in this broad field, including the MUSP, the National Survey of Adolescent Mental Health, and the Longitudinal Study of Women's Health (and we have added the relevant citations). We have also slightly restructured this section to accommodate this added nuance.

Nevertheless, we remain on firm ground in maintaining the key point originally made – namely, that evidence of child maltreatment prevalence to date is neither nationally representative, nor comes from studies directed towards the primary question of establishing reliable prevalence data. For this reason alone, the field is characterised by a major gap in evidence. This is further supported by the associated gaps in evidence presented by other measurement limitations (inattention to maltreatment types, and to a broad range of health outcomes, and burden of disease).

We have added a sentence to this paragraph which further supports our acknowledgment of the evidence of causal associations, and we have referred to the study by Norman et al (2012) and its use of the Bradford Hill Criteria.
I think it would be beneficial to add a few lines in the limitations section that address the strengths of “behaviour-specific” survey questions in relation to changing definitions of maltreatment and parenting norms over time (especially given the different wide range of age groups being sampled in this study).

We agree with the Reviewer that the use of behaviourally-specific screener items, and their conception and testing and final design, is an important topic. We feel that this topic warrants comprehensive discussion and the best place for this is in our forthcoming work detailing the process of instrument configuration, adaptation and testing for the ACMS. While agreeing with the Reviewer, we are reluctant to explore this topic here in the protocol article as we feel in such a short space we cannot properly do it justice in relation to even one, let alone five forms of maltreatment (each of which have multiple sub-dimensions corresponding to the 20 screener items), and so we would prefer to address this in full elsewhere.

**Reviewer 2 Comments**

**Authors’ response**

Comments to the Author:
I have attached my comments in a word document. Congratulations on this excellent work.

We thank Reviewer 2 for this acknowledgment and we appreciate this Reviewer’s deep engagement and thoughtful suggestions.

Thank you for the opportunity to review, “The Australian Child Maltreatment Study (ACMS): Protocol for a national survey of the prevalence of child abuse and neglect, associated mental disorders and physical health problems and burden of disease” for publication in BMJ Open. I would like to take the opportunity to share my sincere thanks to the authors and their associated team members for championing such important work in child maltreatment. In addition, I extend my appreciation to the members of the Technical Expert Panel, those who participated in the cognitive testing and test-retest reliability of the survey measures, and most importantly, the four survivors of child maltreatment who reviewed the ACMS items. Given the scope of this research endeavour and the plans to launch the survey in March 2021, I would expect that there is little opportunity to make any changes to the methodology outlined in the protocol. For this reason, I hope that the authors receive my comments with an understanding that much of what I outline below would be helpful to consider – if possible – before the launch of the survey and (if not) in future publications about the ACMS methodology and results.

Comments to consider across the entire protocol:
I suggest using the language of intimate partner violence consistently throughout the abstract, article summary, and body of the manuscript. Using such language would map onto the assessment of “other adverse outcomes” articulated in the methodology and attunes to the importance of recognizing the diversity of genders who are exposed to violence in intimate relationships. Domestic violence is heavily centred on ‘traditional’ notions of heterosexual intimacy and patriarchy.

We understand the point being made. However, we need to use both concepts because we measure both: (1) participants’ childhood exposure to domestic violence (where this concept is different to and broader than IPV) and we do this in a way consistent with Finkelhor’s JVQ; and (2) participants’ experience as adults of intimate partner violence (where this concept of IPV is...
The use of the word ‘sensitive’ to describe research related to child maltreatment in my view, is not necessary. I would suggest that the authors consider the removal of this evocative language throughout the manuscript. There are many concepts, constructs, and exposures that are considered in the ACMS; the interpretation of what is and is not ‘sensitive’ to be asked about is a personal one (e.g., asking about marriage, divorce, sexual activity, alcohol, or drug use, etc) could all be considered sensitive to a given person, family, group, or culture. Its use to describe research on child maltreatment reinforces problematic notions that child maltreatment cannot be asked about in a systematic and respectful way – the proposed study clearly demonstrates (though its own approach and the citation of other work) that in fact, it can be.

Thank you; we have removed these references.

Suggest that the authors check the ‘tense’ of the entire article. In some cases (e.g., abstract) present vs. past tense are used.

Thank you – we have done this and made several revisions accordingly.

Abstract:

* Suggest that the authors add their plan for missing data into the abstract. In addition, can the authors offer the specific strategy that they will use for imputation (e.g., joint modelling vs. fully conditional specification/chained equations)?

To respect the word limits of the journal, the missing data analysis plan has not been added to the abstract. However, as indicated above, we have provided additional information on the planned approach to handling missing data in the manuscript.

Introduction:

First paragraph

• suggest moving citations 3-6 to the end of the sentence, as those references cover both mental and physical health consequences of CM.

• Can the authors be more specific in what they mean about self-directed violence? (e.g., self-harm and suicide).

We have revised the sentence to clarify that we refer to self-harm and suicidality.

Fifth paragraph

o The comment that there is “no Australian evidence of the nature, prevalence and timing of mental disorders and physical health outcomes associated with child maltreatment…” is not accurate (which I am positive that the authors know this, given that a number of them are authors on important studies; e.g., https://doi.org/10.1016/j.chiabu.2016.12.002;10.1016/j.childyouth.2016.11.014; https://doi.org/10.1016/j.chiabu.2015.05.006). Thus, this sentence needs to be revised so that readers who are less familiar with the available evidence in Australia are clear that the evidence in the country is (1) disparate; (2) lacks the voices of young people (especially children/adolescents – which is like most countries); (3) and does not tend to use representative approaches that can characterize population-level incidence, prevalence, correlates, and burdens.

Given the preceding paragraphs we are confident that we clearly indicated in our manuscript that this absence of evidence related to evidence at the population level; but in any case we thank the reviewer for this suggestion and have made revisions to this paragraph to respond to this requested revision.

Sixth paragraph

o Suggest that the authors offer a definition of ‘corporeal punishment’ that falls in line with policy and practice guidelines within the Australia context.

We have added the appropriate definition here which incorporates a robust social science conceptual model, and we have cited the authoritative sources and reordered subsequent citations. For the research purposes of the ACMS, this approach is required, and is more important than employing a definition different to DV) and we do this by using the CAS - Short Form.
Also suggest that authors make explicit what the definitions of child maltreatment are in the Australian context and what level/type of exposure would fall under the purview of child protection - this is important for discerning policy and practice implications that could arise from the findings of the survey, including the adoption of similar measurement and survey approaches in other countries.

On this point we would prefer – especially in a protocol manuscript – not to engage in discussion. This is an extremely complex topic to discuss, since such definitions referred to are not simple, not consistent across States and Territories, and can only be understood by referring in much greater depth to policy and agency practice about the interpretation and application of the concepts in the statutory child protection setting.

Moreover, the purpose of the ACMS is not to measure child maltreatment using statutory child protection definitions (or policy-based definitions when the statutes are silent or ambiguous), but to use robust social science understandings of these concepts and hence to adopt a similar approach to other prevalence studies including the JVQ.

To the extent possible, in project outputs we will consider the implications of our findings for a range of stakeholders and systems, including child protection systems, and for their frameworks of legislation, policy and practice, including by taking into account the many variations that occur in this context across the Australia’s eight States and Territories. This will be further facilitated by engagement with our Advisory Board.

Methods

First paragraph

Please add the source of the sampling frame in the “Sample Selection and Setting” section of the methods. In addition, a further description of the sampling frame should be noted (how does an individual become a part of the sampling frame, who does it miss?)

We have added material in the second paragraph of the Recruitment/procedures section to explain the source and coverage of the sampling frame.

Description of Survey Instrument Section

Maltreatment Questions Section

- Add “(JVQ)” following Juvenile Victimization Questionnaire
- Third paragraph of this section – I am concerned that the emphasis on physical and sexual abuse for items focused on “whether the participant told anyone” and “to whom the disclosure occurred” perpetuates the notion that other forms of child maltreatment are: (a) less impactful over the short and long term; and (b) less important to respond too/less important to have helpful responses too. As the experts on this team would be aware of – emotional abuse, neglect, and child exposure to IPV can have equally damaging impacts on the

We have added “JVQ” as requested.

On the disclosure issue:

We have added a justification for this approach as the reviewer suggested, and will acknowledge this as a limitation in papers reporting on findings for these variables.
I appreciate that there is the need (from a child protection perspective) to prioritize imminent and severe risks to children (which are invoked in physical and sexual abuse disclosure), but there is a real opportunity to leverage data (and recommendations) around disclosure experiences and response experiences by asking these questions to respondents who indicate emotional abuse, neglect, and exposure to IPV. If the authors cannot do this, a justification needs to be added here and this should be appropriately acknowledged as a limitation in papers that report on findings for these variables.

We elected not to include disclosure questions for all five types of maltreatment for three main reasons.

The first reason is that the most salient national and international scientific and policy questions around non-disclosure, delayed disclosure, and the nature of responses to disclosure, relates to sexual abuse (and to a lesser extent physical abuse).

The second is that disclosure of these types of maltreatment (and responses to any such disclosure) is particularly important to child protection systems and policy, including reporting duties. In Australia, only sexual abuse is required by all eight jurisdictions to be reported by designated reporters, and physical abuse is required to be reported by seven out of eight jurisdictions. Accordingly, there was less reason to include the other three forms of maltreatment. Similarly, apart from life-threatening cases of other maltreatment, no other legal or policy-based duty is present requiring reports of other maltreatment.

The third reason is that it was simply not viable to include disclosure questions for all maltreatment types due to the enormous burden of time and cost that this would have entailed.

---

<table>
<thead>
<tr>
<th>A citation is needed for the MINI, as is clarification of the version being used.</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Given the breadth of disorders captured by the MINI, there should be justification here for why the authors have selected these specific mental health conditions when there are strong associations, for example, between CM exposure and other conditions - like eating disorders, which are often chronic, debilitating, and deadly if intervention is not offered early.</td>
</tr>
</tbody>
</table>

We have added a citation for the MINI and reordered subsequent citations.

We have added a justification for the selection of some disorders and not others.

---

<table>
<thead>
<tr>
<th>Instrument Development and Validation Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>First paragraph: Suggest replacing “Study” with ACMS here and throughout the remainder of the paper where authors are referring to the ACMS, more generally. Will reduce confusion and make it clear that the authors are only referring to one study, which is the ACMS.</td>
</tr>
</tbody>
</table>

We have made this suggested change throughout.

---

<table>
<thead>
<tr>
<th>Second paragraph:</th>
</tr>
</thead>
<tbody>
<tr>
<td>I was surprised to see that children’s exposure to IPV is being limited to direct witnessing of IPV behaviours and I think this relates to how this exposure may be conceptualized related to child protection in the Australian context. This reiterates the need to clearly define these exposures/experiences in the</td>
</tr>
</tbody>
</table>

We do in fact include in the concept of witnessing any incident where the child either sees or hears the incident, and this is incorporated into our screener.
introduction, the rationale for using those definitions. A rationale for not considering awareness of IPV as a form of exposure should be provided or acknowledged in the limitations (work by Wathen, MacMillan, Hamby, and others discusses the impact of awareness of IPV).

questions. We have clarified this in the text in this second paragraph.

(As a further note: in the configuration and testing process (including conceptual testing) we determined it was not optimal or possible to include more distal manifestations of “witnessing” such as awareness or learning of an incident after the event. This approach is also congruent with the JVQ, especially in its core screener in this domain (which the ACMS in fact extends by including hearing), and its approach to both screeners and follow-ups. To employ the concept in the broadest possible way by including awareness without seeing or hearing (at any time in childhood) would introduce a risk in the ACMS of inflating the data and introducing bias into the measurement of associated health outcomes. In situations where we were confronted with a choice of limitations produced by two different courses of action, we consistently elected the conservative option to ensure we could reduce the likelihood of being accused of data inflation. Naturally, in our project outcomes we will fully detail our approach to conceptualisation of maltreatment types and we will acknowledge both the strengths and any limitations these have for analyses).

Pilot data processing and analysis

I appreciate that some of the analysis related to the pilot is ongoing, however, given that the likelihood of the protocol being routinely reviewed and cited by other jurisdictions in the future, there are a few details that I think would be helpful here. For example, can the proportion of data (range for the different modules in the survey) and whether the data was missing at random be added to the protocol?

We have added a sentence about missing data and refusal rates, and a statement in the first paragraph in this section to state that: “In forthcoming work we report full details of the process of developing and testing the modified instrument, which will include comprehensive psychometric data.”

We have adopted this approach because before submission of the protocol the research team discussed whether or not to include figures and specifics of the psychometric data. We concluded it was not appropriate in a protocol manuscript to include this data, because: (1) some journals expressly prohibit inclusion of this data in protocol manuscripts and encourage separate publication of it as a
research article; (2) we did not wish to publish pilot data about the experience of maltreatment types since it is not representative and there is a risk of misuse of such data; (3) we conducted an extensive process of instrument configuration and testing and we are writing a separate comprehensive article explaining why this was required, how it was done, why we conducted the range of analyses, the nature and significance of the pilot data, and a full discussion of its implications.

**Main Study**

**Second paragraph**

- Please provide the timeframe in which the advance text message to notify potential participants of the survey is provided before the phone call.
- Please also clarify if individuals could, after receiving the text message, can: (1) reply to the text and request to complete the survey right then; (2) click on a link in the text and complete the survey electronically (and if not, why not?)

In the second paragraph we have added the requested information about the timeframe.

We have also provided information about individuals’ capacity to reply to the text message, and the purpose of this reply (establishing age-eligibility and providing an opt out mechanism).

We trust it is clear from the protocol that administration is uniformly by CATI rather than self-completion online, and so we have not repeated this here.

**Planned Analysis section:**

- *Please provide an indication if weights will be generated prior to analysis to ensure estimates are representative of the population.*
- *Please provide clarity of what type of imputation will be utilized and when this will occur (will it occur for the entire data set or will imputation procedures be completed per analysis focus?)*
- *My sense is that having the expert’s complete imputation procedures PRIOR to releasing the data to other researchers would be important to ensuring consistency in estimates and analytical approaches.*

We have added information on the derivation of survey weights by comparison with Census data and other key national collections. Thorough analysis of sample representativeness will be undertaken and all analyses will incorporate survey weights to ensure estimates are representative of the Australian population.

As noted above, additional information has been provided on imputation procedures, which will be partially data-driven.

Weighting and imputation will be undertaken prior to releasing data to other researchers, and the data made available will include all derived variables, weights and imputation sets, allowing other researchers to replicate estimates from the study.
<table>
<thead>
<tr>
<th>Estimating Disease Burden Attributable to Child Maltreatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A citation is needed after the description of approach for generating YLLs, YLDs, and DALYS.</td>
</tr>
<tr>
<td>• Can the authors offer more information about how they will estimate the economic burden of CM?</td>
</tr>
</tbody>
</table>

We have added this citation, and have provided further detail of this with another citation.

We have added six lines of information about the estimation of economic burden to this paragraph.

<table>
<thead>
<tr>
<th>DISCUSSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel strongly that a brief paragraph, at the start of the discussion, needs to be added to articulate the methodological and substantive contributions of this protocol. This team’s work presents an incredible opportunity to present a ‘standard’ by which this important data can be collected in a robust and respectful way.</td>
</tr>
</tbody>
</table>

We have added a paragraph to the start of the Discussion, and the first half of it accommodates this suggestion by stating: “This protocol outlines the ACMS approach to measuring child maltreatment in a national population, and its associations with mental disorders and physical health, and burden of disease. Much international work has been conducted in this field, and the ACMS aims to make further contributions to inform subsequent research of the highest rigour.” |

Suggest that the Distress Protocol and the Risk of Harm protocol be attached as supplementary files. My apologies if I missed this. Again, this reiterates the need to support the translational impact of this work to other countries.

We do wish to support the translational impact of this work elsewhere, and for this reason we are currently in the process of writing separate articles about these protocols. These protocols require full analysis since they involve multiple complex questions, and hence require separate comprehensive treatment. We will be including the protocols in those articles.

To respond to this Reviewer’s request, we have added a statement to indicate this approach, in the final paragraph of the Participant Safety section.

Can you strengthen the mobile phone administration section of the work by providing an indication of the proportion of individuals in Australia who have access to a mobile phone? In addition (and to address potential concerns around missing Indigenous Australians), one could use geomapping methods to examine the extent to which mobile phone ownership attached to post codes (for example) are located areas with a high concentration of Indigenous peoples.

In our original draft, we had included information about mobile phone consumption to further support our approach, but we decided to omit this due to word count requirements. In this final revised manuscript we have added this requested material at the “Mobile phone administration” section.

In relation to the geomapping methods suggestion: it is not possible to include further material on this point. We have noted in the Limitations section that Indigenous Australians are likely to be under-represented. Geomapping methods would not be advantageous
given the small cell sizes of likely participation by Indigenous Australians, further limitations by location, and the limitations on analysis of this particular sub-population in this type of benchmark national study.

Suggest removing the use of “mainstream” populations in the paragraph right before the Ethics and Dissemination section.

We have replaced the words “surveys of mainstream populations” with “surveys of random samples of the Australian population”.

Ethics and Dissemination
I just want to note my appreciation to the authors for their plan to make the instrument available through a creative commons license.

Thank you.

---

**VERSION 2 – REVIEW**

| REVIEWER                              | Degli Esposti, Michelle  
University of Oxford, Social Policy & Intervention |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>REVIEW RETURNED</td>
<td>31-Mar-2021</td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

Thank you to the authors for thoughtfully and thoroughly addressing our review comments. The protocol has been strengthened and I would be pleased to see it published in BMJ Open.

| REVIEWER                              | Kimber, Melissa  
McMaster University, Department of Psychiatry and Behavioural Neurosciences |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>REVIEW RETURNED</td>
<td>20-Apr-2021</td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

Thank you for the opportunity to re-review the ACMS protocol. The authors responses to my own and the other reviewer’s revision requests are thoughtful, measured, and well justified. I look forward to the academic community having access to this important protocol. I have no other concerns to be addressed.