

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)	Psychological, social, and financial impacts of COVID-19 on culturally and linguistically diverse communities in Sydney, Australia
AUTHORS	Muscat, Danielle; Ayre, Julie; Mac, Olivia; Batcup, Carys; Cvejic, Erin; Pickles, Kristen; Dolan, Hankiz; Bonner, Carissa; Mouwad, Dana; Zachariah, Dipti; Turalic, Una; Santalucia, Yvonne; Chen, Tingting; Vasic, Gordana; McCaffery, Kirsten

VERSION 1 – REVIEW

REVIEWER	Mude, William CQ University - Cairns Campus
REVIEW RETURNED	31-Oct-2021

GENERAL COMMENTS	<p>Peer Review: Psychological, social and financial impact of COVID-19 on culturally and linguistically diverse communities: a cross-sectional Australian study</p> <p>(BMJ Manuscript ID: bmjopen-2021-058323)</p> <p>I would like to thank the authors for this work. This is an important study that can contribute to effective support of CALD communities during and after COVID-19. However, there are major issues to be addressed which I believe will make the paper stronger. I am particularly concerned with the regression findings presented in this study. I have presented these issues below, and you will also find them commented on your paper. I hope you will find these helpful.</p> <p>Introduction:</p> <p>Page 6 of 36 line 7 - The word 'ethnic' is considered inappropriate these days because of its colonial history, but also because of its ambiguity (Does it refer to race, religion, tribes, culture, country of origin, language groups etc?) I would suggest using "people from racial and cultural minority groups"</p> <p>Page 7 of 36 lines 9-17. A further clarification is required for what the author means by "migrants". Did they mean people who came to Australia for economic opportunities other than being forced to</p>
-------------------------	--

flee their home country? One can argue that "migrants" can also include refugees and asylum seekers if not specified by the author.

Page 7 of 36 line 22 is not clear, rephrase. "... differences in financial and psychological impacts of COVID-19 those for who spoke a language other than English at home ...". The author seems to have missed a word between "COVID-19" and "those".

Methods:

Page 8 of 36 line 21-22 – The author wrote that their study was informed by the framework for culturally competent health research, but no further information is provided about this framework. Mentioning a framework without saying how it informs the study is not sufficient. The author needs to elaborate on how the framework used has informed the study and its analysis.

Page 8 of 36 line 58-60 - The author stated in their introduction that their study focuses on CALD community because the experience of this community is different from refugees who have come to Australia within the last 10 years. In lines 9-17 on page 7 of 36, the author's rationale for the study is to focus on "those who speak a language other than English at home who have not been forced to flee their home country". However, all most all Dinka (a tribe from South Sudan), Khmer, Assyrian and some Arabic speaking people came to Australia as refugees because they have been forced to flee their home countries. Including these communities as participants contradict the rationale provided in the introduction in lines 9-17. The author needs to elaborate how they excluded refugees and asylum seekers from their study? As it stands, the information about the participants and the rationale for their study do not match.

Page 10 of 36 line 13 – The author said they 'adapted' a validated tool. What were the reported KMO and Cronbach's alpha for the validated instrument? Did the author conduct a revalidation of the instrument after "adapting" it?

The referenced validated tool in line 13 on page 10 of 36 was developed to measure "financial toxicity of cancer within a public healthcare system" in Italy. Was this a relevant tool to adapt to measure the financial and psychological impacts of COVID-19? I doubt. It is not exactly clear which part of the validated instrument was 'adapted' for this manuscript. A look at the referenced work indicate that it was an abstract not a fully published paper. A further search found a publication in BMJ and the reported questionnaire did not match the items reported under the financial and psychological items. You can see the published 'validated tool' from Riva and colleagues' paper below.

Riva S, Arenare L, Di Maio M, et al. Cross-sectional study to develop and describe psychometric characteristics of a patient reported instrument (PROFFIT) for measuring financial toxicity of cancer within a public healthcare system. *BMJ Open* 2021;11:e049128. doi:10.1136/bmjopen-2021-049128

I would suggest the author reference kessler psychological distress scale here instead of the Riva's paper.

Page 10 of 36 line 13 - Reference (14) is an abstract. Author needs to reference the published version in *BMJ*.

Riva S, Arenare L, Di Maio M, et al. Cross-sectional study to develop and describe psychometric characteristics of a patient reported instrument (PROFFIT) for measuring financial toxicity of cancer within a public healthcare system. *BMJ Open*

2021;11:e049128. doi:10.1136/

bmjopen-2021-049128

Page 10 of 36 line 13-16 - Was the 'co-designed' questions pilot tested? Did the author perform factor analysis or principal component analysis to determine the validity of these items?

Page 12 of 36 line 12-16 – Is it appropriate for the author to use Likert scale to measure “mean”? I doubt this is the right approach.

Page 12 of 36 line 16-18: Any objective evidence to support this measurement? OR they were arrived at subjectively? If subjective and without the validation of the tools, how can a reader rely on the findings?

Page 12 of 36 line 18-22: Ditto

Page 12 of 36 lines 25-31: What is the significance level used to determine significant differences?

Results

Page 13 of 36 line 7: Is advanced diploma or diploma level not a tertiary qualification? If this is the case, the author needs to make it clear. I would suggest instead of reporting the negative (no tertiary qualification), the author needs to be specific and use a positive tone, e.g. 29.7% had bachelor degree level or higher. OR 70% had advanced diploma level education or lower.

Table 2 on page 14 of 36: Assyrian, Arabic speaking people, Dinka, Khmer have come to Australia as refugees. As previously pointed out, including these population groups conflict with the study's rationale at the introduction section.

Table 2 include "Years of living in Australia" categorised into "5 years or less", "6 to 10 years" and "more than 10 years". See the previous comments on participants inclusion and exclusion. Given that there are participants who lived in Australia for 5 years or less are included in the study, how did the author exclude those who are refugees from the included communities?

Page 16 of 36 line 26: How did the author determine whether participants had two or more chronic illnesses? The survey instrument in Table 1 does not contain question on participant disclosing their comorbidities.

The following issues were observed relating to all the regression analysis reported on page 16 of 36 line 22-29; Page 16 of 36 line 47-53; Page 17 of 36 line 10-25; Page 17 of 36 line 40-49, and Page 18 of 36 line 25-31:

Regression analysis finding provided relating to supplementary table 1. "Lowest ISRAD quintile" and "Years living in Australia" are not significant at 0.05 level for unadjusted analysis (Nervous/stressed and Alone/Lonely). Why did the author include

these factors in the adjusted analysis when they don't make any contribution into improving the model?

Supplementary table 2: All the highlighted factors (unadjusted analysis) were not statistically significant at 0.05 significance level and therefore are not making any contribution to improve the regression model. They shouldn't be included in the final model (adjusted analysis) because they are not significant explanatory variables in this case.

Supplementary table 3: The variable "Lowest IRSAD quintile" is not significant at 0.05. Why did the author include them in adjusted analysis when clearly it is not contributing to the improvement of the model without the other covariates? The same issue is observed for Adequate health literacy, English-language proficiency, and gender. I would suggest the author remove all these factors from the adjusted model. I am concerned with the regression finding presented in this study, which in my opinion requires major revision.

Page 17 of 36 Line 38-40 (including Table 3 and 4): How can the Likert score be used to measure the "mean financial burden" and the "mean negative impact on children"?

Likert scale data cannot use the mean measure central tendency because it has no meaning numerically. Instead of using the mean, I would suggest the author use % (combined agree and strongly agree) to report their finding in Table 3 and 4 for the two outcomes. Alternatively, they can remove the information on mean from Table 3 and 4

Page 18 of 36 lines 11-13: Ditto. I would suggest rewriting the whole of the remainder of this section (starting from the highlighted) with the comments in mind. Reporting a Likert data as a mean to measure central tendency is a major flaw.

Table 3: As reported in Table 3 and 4, it is useful for the author to describe how they grouped the survey items under psychological impacts, social impacts, financial impacts. Did they conduct PCA or factor analysis to determine these?

	<p>Discussion:</p> <p>I was a bit surprised by the discussions, which appeared limited in scope and nuances. I suggest the author thoroughly discuss their findings instead of resorting to making a general statement.</p> <p>Discuss the findings on page 16 of 36 line 16-20</p> <p>Page 21 of 36 line 43-45: Where are the evidence to support this statement?</p> <p>Page 21 of 36 line 51-55: I would suggest discussing what others discussed first before discussing your own previous study.</p>
--	---

REVIEWER	Michener, Lloyd Duke University, Family Medicine & Community Health
REVIEW RETURNED	30-Nov-2021

GENERAL COMMENTS	<p>Thank you for the opportunity to review “Psychological, social and financial impact of COVID-19 on culturally and linguistically diverse communities: a cross-sectional Australian study” which adds to the growing global literature on the disproportionate impact of COVID-19 on minority communities. Strengths of this study include its use of inclusive recruitment methods, including co-design with community staff, translated surveys, availability of interpreters, and participation in community events. Weaknesses are noted appropriately, and include use of survey questions from multiple sources, plus a few new survey questions, without the opportunity to reassess validity; convenience sampling; and self-reporting. All of these are appropriate choices given the goals and timelines.</p> <p>There are, however, multiple places in which the manuscript could be significantly improved. These represent opportunities for better clarity of what was done and found, These are listed consecutively below.</p> <p>GENERAL COMMENTS</p>
-------------------------	--

The word "impact" is used extensively throughout the article. The issue is that "impact" is causal language and would not be appropriate in a cross-sectional study. "Correlates" or "links" might be better.

TITLE:

Somewhat misleading. The reported study is of Greater Western Sydney in New South Wales, and is not a survey of a cross section of the country of Australia. Renaming the manuscript to reflect the scope of the work more accurately would be helpful. Perhaps "Cultural and Linguistic Predictors of COVID-19-related Psychological, Social, and Financial Outcomes in Diverse Communities" may be a clearer title for readers

ABSTRACT

Objective: It is unclear what the predictors are of the psychological, social, and financial outcomes. Maybe refer to them as "sociodemographic predictors"?

The fact that this study was conducted primarily among "people who primarily speak a language other than English" should be included in the Abstract

Outcome measures: Again, the use of "correlates" rather than "impacts" and being clearer by saying "sociodemographic factors" associated with the psychological, social, and financial outcomes" is suggested.

Results:

Line 57-58 "Even prior to the COVID-19 outbreak in Sydney, 25% of the sample reported..." is unclear

Where did these values come from?

Is this paper a trend study? Were comparisons before and after the pandemic a substantial part of the analysis?

Presenting this statement may be going beyond the scope of what the authors indicated they were analyzing based on what's been written in the Abstract so far.

It took a while to find the Table with the numbers corresponding with "...most parents reported that their children were less active (64%), had more screen time (63%), were finding school harder (45%)." These numbers are buried in Table 4 in the All column but would probably be best positioned in either Table 2 (Descriptives) or Table 3.

Authors should be cautious when it comes to saying "distinct impact patterns". Such wording cues readers to think that the authors had conducted some type of cluster analysis (e.g., latent class/profile analysis), which is not done in this study. The patterns were mere observations by the authors. It would be better to temper down the wording here (e.g., "there may be patterns..." or "there appeared to be potential patterns...").

INTRODUCTION

The statement is made that Australia is one of the most culturally diverse nations worldwide, but without further discussion, explanation nor reference. In the next paragraph, Australian refugees are defined by language, but not by location. It can be difficult to determine, particularly in the Introduction, how groups are being defined, and which are national, regional, defined by language or other.

"Culturally and linguistically diverse groups" needs to be more clearly defined and the literature review really needs to do more to justify the inclusion of the sociodemographic factors that the authors chose to include in their analyses and models.

Language spoken and language spoken at home are covered, but what about English proficiency? Health literacy?

Why was an examination of demographic predictors (gender, age group, IRSAD quintile, and comorbidities) carried out? These are featured in the tables but not addressed in the Introduction.

Some justification of the particular psychological, social, and financial outcomes for this study may be warranted

Why nervousness and loneliness?

Why impact on relationships, time looking after children, screen time, physical activity, time with friends, school difficulties

Why financial status and need?

Was there a framework or tool that guided the selection of these particular outcomes for this study?

METHODS

Page 7 - It would be helpful to readers if the authors could explain in greater detail what the Framework for Culturally Competent Health Research is and whether it also informed the authors' selection of the some of the predictors in their tables/models.

Why were 11 language groups were selected? What was the rationale behind it? With N=708 participants, not much is able to be said about stratifying results by language due to limited samples and the many metrics/items they are including

Line 48 –what does “up to 39% of residents” mean? Up to 39% of the residents of one of the three listed regions? Up to 39% of the immigrants from one country within one region?

RECRUITMENT

Page 8, line 24. The survey was online, web based, and could be completed by participants, staff, or an interpreter. What was the required reading level for completion of the survey, and did that vary across language groups? This is of particular concern given the finding that 59% had low health literacy.

Page 9 of 36, lines 34-35: “Translations were completed by translators with National Accreditation Authority for Translators and Interpreters (NAATI) accreditation where possible.” Even though this seems to be appropriate, it would be helpful if, in addition to translations there were also adaptations made to the specific participants (the World Health Organization distinguishes between translations and adaptations) Translators and interpreters may not be familiar with specific terminology in some fields and

may do literal translations. It would be good to include this as a limitation.

MEASURES

Page 10 - Is Table 1 the full survey? 15 questions are described in the text, but only 13 seem to be listed.

Page 12 – what is a “tertiary qualification”?

If the authors are going to focus on gender, age, IRSAD quintile, and comorbidities in their results, this needs to be addressed.

Authors need to elaborate more on the Mean Negative Impact on Children as well as the Mean Financial Burden measures in the METHODS.

If they are presenting the n (%) for each of these individual indicators, then why are they computing means? They should choose to do one or the other.

What is the scientific justification for computing these means?

Is there a validated approach to computing a mean negative impact on children score from screen time, physical activity, time with friends, and schooling? What is the Cronbach's alpha for these items?

Is there a validated approach to computing a mean financial burden score from changing employment status, concern about financial problems, and inability to meet weekly expenses? What is the Cronbach's alpha for these items?

ANALYSES

The authors talk about weighting for their frequencies, but the results being presented look like they are unweighted. Typically, when weighted estimates are being presented, they are accompanied by a confidence interval for the estimated prevalence. This does not appear to be the case here.

Or is it really the case that N=708 is a weighted sample size derived from a smaller sample?

The Mean Perceived Financial Burden and Mean Impact on Children do not seem to represent validated measures. This is problematic.

It is unclear what the difference is between unadjusted and adjusted analysis

RESULTS

The way the Results are presented, the cultural and linguistic diversity aspect seems to take a lower priority over the demographic predictors (i.e., age group and gender). There should be a greater emphasis on the disparities according to health literacy and language spoken, since these pieces are what are most compelling about the study.

The results should highlight what was found with regard to health literacy, language group, English language proficiency, years living in Australia.

Table 2. Descriptive Statistics of Analysed Sample

	<p>This table needs to include the proportion of participants with children aged less than 18 years, since the analysis on negative impacts of children only pertains to this subgroup</p> <p>The descriptive statistics for the outcomes presented in the ALL column from Table 4 should actually be here, so readers can see the prevalence of the psychological, social, and financial correlates from the full sample.</p> <p>Whether participants received assistance with completing the survey may be worth including in this table as well.</p> <p>Table 3 and Table 4</p> <p>It might be easier for the readers if Tables 3 and 4 were reorganized into three separate tables, one for psychological, one for social, and one for financial COVID-19 correlates, with the demographic predictors (gender, age group, health literacy, IRSAD quintile, comorbidities, and language group on the side)</p> <p>These tables should also prioritize the presentation of results based on the health literacy and language group predictors, since these are supposed to be the star of the of study.</p> <p>Supplemental Tables 1, 2, & 3</p> <p>It is highly unclear what the difference is between the unadjusted and adjusted models.</p> <p>This applies to all Supplemental Tables</p> <p>There are some p-values that are misplaced</p> <p>What do the bolded p-values refer to? Are these results from Wald tests?</p> <p>There are some p-values that are misplaced.</p> <p>For the Alone/Lonely unadjusted results (Supplemental Table 1), why is there a P-value for the Reference group for co-morbidity? That does not belong there.</p> <p>Authors should report the analytic sample size for these regression models, as some models (i.e., negative impacts on children) may actually be a subsample of the 704 participants. This analysis would only pertain to respondents with children younger than 18 years old.</p> <p>IMPACTS</p> <p>page 15 on, plus Tables 3 and 4. While there are very useful comparisons noted, it is not clear to what extent the patterns noted predated COVID, reflecting broader societal shortfalls; were exacerbated by them, reflecting inadequate COVID response, likely on top of systemic failings; or even were relatively unaffected, reflecting community resilience. Part of the issue is that the (very comprehensive) data displays make it hard to identify fields for which comparable data is available.</p> <p>A major strength of the study is its focus on health literacy, English-language proficiency, language group, and years living in Australia predictors. These predictors are what make this study compelling, yet greater emphasis has been placed on the demographic predictors (e.g., gender and age) and just overall descriptives. This is somewhat of a missed opportunity that could really inform targeted intervention efforts.</p> <p>DISCUSSION</p>
--	--

	<p>The discussion should highlight what was found with regard to health literacy, language group, English language proficiency, years living in Australia.</p> <p>There needs to be a deeper discussion relating the children/families outcome as well as the financial burden outcomes to the cultural and linguistic needs of diverse communities.</p> <p>IMPLICATIONS</p> <p>The solutions that the authors propose need to be closely tied to the specific findings from their analyses.</p> <p>What is it about the authors' results that tell us about the need for system readiness or culturally safe support packages?</p> <p>What is it about the authors' results that tell us about communal ways of coping?</p> <p>What is it about the authors' results that tell us about the need for migrant-inclusive public information strategies?</p> <p>There is an interesting pattern, not clearly spelled out, which seems to tell a story of resilience across these communities:</p> <p>Page 20, line 56 – 25% of participants reported feeling nervous or stressed most or all of the time (national, April 2020)</p> <p>Page 21, line 5 – 20% of Australians experienced high or very high levels of psychological distress; 28% reported feeling nervous (national, June 2021)</p> <p>This study – 25.3% reported feeling nervous or stressed most or all of the time (NSW, March- July 2021)</p> <p>If this pattern of community resilience is correct, then it may be an important finding to highlight, in addition to the multiple findings of negative impacts.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Thank you for the opportunity to review “Psychological, social and financial impact of COVID- 19 on culturally and linguistically diverse communities: a cross-sectional Australian study” which adds to the growing global literature on the disproportionate impact of COVID-19 on minority communities.

Strengths of this study include its use of inclusive recruitment methods, including co-design with community staff, translated surveys, availability of interpreters, and participation in community events. Weaknesses are noted appropriately, and include use of survey questions from multiple sources, plus a few new survey questions, without the opportunity to reassess validity; convenience sampling; and self-reporting. All of these are appropriate choices given the goals and timelines.

There are, however, multiple places in which the manuscript could be significantly improved. These represent opportunities for better clarity of what was done and found. These are listed consecutively below.

GENERAL COMMENTS

Comment: The word "impact" is used extensively throughout the article. The issue is that "impact" is causal language and would not be appropriate in a cross-sectional study. "Correlates" or "links" might be better.

- Response: Questions in our survey specifically asked participants about the impacts of COVID-19 on psychological, social and financial outcomes (e.g. “I worry about the financial problems I will have in the future as a result of the COVID-19 pandemic”). As such, we feel

that it is acceptable to refer to the impacts of COVID-19 throughout the manuscript. We have, however, amended wording throughout to refer more commonly to 'outcomes' and 'sociodemographic correlates'. As an example of one of several changes, the first point in the 'Strengths and limitations of this study' box has been revised: "This is the largest Australian survey exploring COVID-19-related psychological, social, and financial outcomes, and the sociodemographic correlates of these outcomes, among people who primarily speak a language other than English..."

TITLE:

Comment: Somewhat misleading. The reported study is of Greater Western Sydney in New South Wales, and is not a survey of a cross section of the country of Australia. Renaming the manuscript to reflect the scope of the work more accurately would be helpful. Perhaps "Cultural and Linguistic Predictors of COVID-19-related Psychological, Social, and Financial Outcomes in Diverse Communities" may be a clearer title for readers.

- Response: We have now changed the title to: "Psychological, social, and financial impacts of COVID-19 on culturally and linguistically diverse communities in Sydney, Australia"

ABSTRACT

Comment: Objective - It is unclear what the predictors are of the psychological, social, and financial outcomes. Maybe refer to them as "sociodemographic predictors"?

- Response: The objectives section of the manuscript has now been changed to: "To explore the socio-demographic predictors of COVID-19-related psychological, social, and financial outcomes in culturally and linguistically diverse communities in Sydney, Australia."

Comment: The fact that this study was conducted primarily among "people who primarily speak a language other than English" should be included in the Abstract.

- Response: The setting section of the abstract has now been revised to include this information: "Setting: Participants who primarily speak a language other than English at home were recruited from Greater Western Sydney, New South Wales, Australia."

Comment: Outcome measures - Again, the use of "correlates" rather than "impacts" and being clearer by saying "sociodemographic factors" associated with the psychological, social, and financial outcomes" is suggested.

- Response: The outcome measures section of the abstract has been revised as follows: "Thirteen items regarding COVID-19-related psychological, social, and financial outcomes were adapted from validated scales, previous surveys or co-designed in partnership with Multicultural Health and interpreter service staff. Logistic regression models (using post-stratification weighted frequencies) were used to identify socio-demographic predictors of outcomes. Surveys were available in English or translated (11 languages)."

Comment: Results - Line 57-58 "Even prior to the COVID-19 outbreak in Sydney, 25% of the sample reported..." is unclear. Where did these values come from?

Is this paper a trend study? Were comparisons before and after the pandemic a substantial part of the analysis? Presenting this statement may be going beyond the scope of what the authors indicated they were analyzing based on what's been written in the Abstract so far.

- Response: We apologise for the confusion created by this statement. Although not a formal part of the analysis, we believe this is important contextual information. We have re-written the sentence as follows: "This analysis, conducted prior to the 2021 COVID-19 outbreak in Sydney, showed that 25% of the sample..."

Comment: It took a while to find the Table with the numbers corresponding with "...most parents reported that their children were less active (64%), had more screen time (63%), were finding school harder (45%)." These numbers are buried in Table 4 in the All column but would probably be best positioned in either Table 2 (Descriptives) or Table 3.

- Response: We have significantly revised the tables in this manuscript, such that the social, psychological and financial outcomes are presented in separate tables. Each table

now also includes a 'Total' column at the top to allow readers to easily identify the proportion of people in the entire sample experiencing each outcome.

Comment: Authors should be cautious when it comes to saying "distinct impact patterns". Such wording cues readers to think that the authors had conducted some type of cluster analysis (e.g., latent class/profile analysis), which is not done in this study. The patterns were mere observations by the authors. It would be better to temper down the wording here (e.g., "there may be patterns..." or "there appeared to be potential patterns...").

- Response: We have changed all sections of the abstract which previously referred to 'distinct impact patterns'. This includes the results section ("Regression analyses consistently showed more negative outcomes for those with comorbidities and differences across language groups.") and the conclusion ("Culturally and linguistically-diverse communities experience significant psychological, social and financial impacts of COVID-19. A whole-of-government approach with policy and sustainable infrastructure is needed to support rapid co-design of culturally-safe support packages in response to COVID-19 and other national health emergencies, tailored appropriately to specific language groups and accounting for pre-existing health disparities.").

INTRODUCTION

Comment: The statement is made that Australia is one of the most culturally diverse nations worldwide, but without further discussion, explanation nor reference. In the next paragraph, Australian refugees are defined by language, but not by location. It can be difficult to determine, particularly in the Introduction, how groups are being defined, and which are national, regional, defined by language or other.

- Response: We have added the following sentence to qualify the statement about cultural diversity in Australia, referencing the Australian Human Rights Commission and Australian Bureau of Statistics: "Currently, people living in Australia identify with more than 270 ancestries, with almost seven million people migrating to Australia since 1945. In 2020, 29.8% of Australia's population were born overseas, a level that is higher than most countries within the Organisation for Economic Co-operation and Development (OECD)."
- In the following paragraph, we have replaced references to the language spoken by study participants which the most common countries from which refugees had arrived: "In a study of 656 refugees and asylum seekers who had arrived in Australia within the last 10 years (most commonly from Iraq (58.7%, n=385) and Syria (16.9%, n=111))..."

Comment: "Culturally and linguistically diverse groups" needs to be more clearly defined and the literature review really needs to do more to justify the inclusion of the sociodemographic factors that the authors chose to include in their analyses and models. Language spoken and language spoken at home are covered, but what about English proficiency? Health literacy?

Why was an examination of demographic predictors (gender, age group, IRSAD quintile, and comorbidities) carried out? These are featured in the tables but not addressed in the Introduction.

- Response: Thank-you for highlighting this important point. In response, we have included the following paragraph in the Introduction of the manuscript:

"There also remains limited data about the socio-demographic predictors of COVID-19-related psychological, social, and financial outcomes in culturally and linguistically-diverse communities. A myriad of socio-demographic factors put communities at increased risk for worsened COVID-19 outcomes. Language barriers, for example, are a well-established driver of inequitable outcomes in health care, often arising from worsened patient experience, unmet informational needs and discrimination (19). Further, the population whose main language is not English are also at greater likelihood of having lower socioeconomic status (20) and lower health literacy (21) among other socio-demographic risk factors which can compound the impact of health emergencies including COVID-19."

Comment: Some justification of the particular psychological, social, and financial outcomes for this study may be warranted.

- Why nervousness and loneliness?
- Why impact on relationships, time looking after children, screen time, physical activity, time with friends, school difficulties
- Why financial status and need?
- Was there a framework or tool that guided the selection of these particular outcomes for this study?

- Response: Outcomes for this study were developed in partnership with Multicultural Health and Health Care Interpreter Service staff. They directly reflected local information needs and priorities, as recommended by the Framework for Culturally Competent Health Research (Woodland, Blignault, O'Callaghan & Harris-Roxas, 2021). We initially started with a broad and comprehensive range of questions that closely mirrored those used in our national survey (McCaffery et al., 2021) and US sister survey (Wolf et al., 2020). From there we removed items with input from Multicultural Health and Health Care Interpreter Service staff collaborators and added some additional questions that they considered pertinent to their communities' experience. For example, this included questions about the impact on children informed by work that they had done previously with focus groups with community leaders.

We have now provided justification for the selection study outcomes in Box 1: "Outcome measures for this survey study were developed in partnership with Multicultural Health and Health Care Interpreter Service staff. This included the selection of broad outcome domains (psychological, social and financial impacts) as well as individual questions."

We have also added additional information about the selection of items into the Measures section of the manuscript:

METHODS

Comment: Page 7 - It would be helpful to readers if the authors could explain in greater detail what the Framework for Culturally Competent Health Research is and whether it also informed the authors' selection of the some of the predictors in their tables/models.

- Response: Thank-you for this important comment. We value the opportunity to provide greater detail about the Framework for Culturally Competent Health Research and its application in this study. We have achieved this through the inclusion of Box 1 in the revised manuscript (reproduced below).

Box 1. Application of the Framework for Culturally Competent Health Research

a) Assemble a culturally competent team: The research team included Multicultural Health and Health Care Interpreter Service staff and bilingual community members from Western Sydney who have extensive experience working with culturally and linguistically diverse communities. Many share the language skills and cultural background of community members in western Sydney.

b) Address community need: Outcome measures for this survey study were developed in partnership with Multicultural Health and Health Care Interpreter Service staff. This included the selection of broad outcome domains (psychological, social and financial impacts) as well as individual questions. The survey was reviewed by the entire study team before implementation to ensure relevance, readability, and clarity of items for community members. Multicultural Health and Health Care Interpreter Service staff also played a key role in the selection of language groups for this study. The goal was to select groups based on several variables including perceived need and size of the community in western Sydney, while allowing for diversity in

regard to time since migration and English-language proficiency.

c) Address health inequities: Multicultural Health and Health Care Interpreter Service staff worked in partnership with researchers to influence decisions about research questions and design as well as interpretation and dissemination of findings. Inclusiveness in recruitment was emphasised throughout training sessions and study updates to promote the recruitment of diverse community members. Findings were presented as 2-page infographics and disseminated to communities through local networks, as well as in the peer-reviewed literature.

d) Address differences in power: This study built on enduring partnerships between researchers, health services and multicultural community organizations that have spanned multiple research projects. The goal for this study and others has been to bring together a range of health staff, consumers and researchers to co-create value together from the outset, placing high value on different types of knowledge, particularly the lived experiences of community members and contextually specific knowledge of our health services partners. Wherever feasible, the goal has been to redistribute knowledge-based power and replace it with mutual learning between all team members. Team building included mentoring clinician researchers and training in research methods, including non-coercive recruitment methods.

Comment: Why were 11 language groups were selected? What was the rationale behind it? With N=708 participants, not much is able to be said about stratifying results by language due to limited samples and the many metrics/items they are including.

- Response: Language groups were selected in consultation with Multicultural Health and Health Care Interpreter Service staff in western Sydney based on a number of variables including perceived need and size of the community in western Sydney, while allowing for diversity between groups in regard to time since migration and English-language proficiency. This has now been clarified in Box 1 (see above). In summary, we wanted to collect data from a range of varied language groups to inform the Local Health District's response to COVID pandemic, ensuring that data could be disaggregated in local reports to inform targeted responses. This was perceived by Local Health District staff to be a more helpful approach than selecting a smaller number of (more dominant) language groups. Pre-specifying language groups ensured that Participant Information Statements and surveys could be translated, and bilingual community staff made available for data collection. We initially had intended to have n=100 for each language group but because of the lockdown we had to end recruitment prior to meeting this target.

We respectfully disagree that “not much is able to be said about stratifying results by language due to limited samples”. In our regression analyses, language was one of only two consistent outcome predictors, and data from this study directly informed targeted responses in Western Sydney Local Health District and at a state level.

Comment: Line 48 –what does “up to 39% of residents” mean? Up to 39% of the residents of one of the three listed regions? Up to 39% of the immigrants from one country within one region?

- Response: We have now revised this sentence as follows: “Participants were recruited from Greater Western Sydney in New South Wales, Australia from three adjoining regions with high cultural diversity: Western Sydney (47% of residents born overseas), South Western Sydney (43% of residents born overseas), and Nepean Blue Mountains (24% of residents born overseas).”

RECRUITMENT

Comment: Page 8, line 24. The survey was online, web based, and could be completed by participants, staff, or an interpreter. What was the required reading level for completion of the survey, and did that vary across language groups? This is of particular concern given the finding that 59% had low health literacy.

- Response: Readability formulas were originally created for use with the English language and are unavailable for many of the languages included in this study. Nonetheless, the readability of the 13 impact questions in the English-language survey was Grade 7 (assessed using the Hemingway Editor). In addition to this, a range of other strategies were used to simplify the survey, including through simplified definitions added as bracketed text, for example, ‘Has your employment status (work) changed because of COVID-19?’. All participants had the option to complete the survey either by themselves online (in English or other language of their choice) or with assistance from bilingual staff or an interpreter who read the questions to them and recorded their responses. In this way, we purposefully integrated methods which would help to ensure the survey was understandable for our diverse population.

We have added information about readability to the Measures section of the manuscript: “The readability of the thirteen items (excluding response options) in English was Grade 7 as assessed using the Hemingway Editor.”

We have also included additional information to the recruitment section to clarify the role of bilingual staff and interpreters in this study: “Potential participants were offered two means of taking part: completing the survey themselves online (available in English or translated), or with assistance from bilingual staff or an interpreter who read the questions to them and recorded their responses.”

Comment: Page 9 of 36, lines 34-35: “Translations were completed by translators with National Accreditation Authority for Translators and Interpreters (NAATI) accreditation where possible.” Even though this seems to be appropriate, it would be helpful if, in addition to translations there were also adaptations made to the specific participants (the World Health Organization distinguishes between translations and adaptations). Translators and interpreters may not be familiar with specific terminology in some fields and may do literal translations. It would be good to include this as a limitation.

- Response: We completely agree that the application of Equivalence Theory to translation is not sufficient and can cause complex and confusing phrases or sentence structures and/or may distort intended meaning. All translators in this study were health staff who were familiar with specific terminology. Under the guidance of the Functional Theory of healthcare translation, the emphasis of the translations was on the reader’s response; the original text was adjusted to make the translation understandable and acceptable to the target readers from different language groups. We did not apply strict word-for-word equivalence. In addition, the final survey was reviewed by the entire multilingual study team as well as bilingual community members before implementation to ensure relevance, readability, and clarity of items for community members. This has now been outlined in Box 1 (reproduced above). Finally, interpreters who administered the survey received training to undertake in-language data collection in a way which retained the intended meaning of the survey questions. Interpreters also used translated versions of the survey to enhance consistency.

MEASURES

Comment: Page 10 - Is Table 1 the full survey? 15 questions are described in the text, but only 13 seem to be listed.

- Response: Table 1 includes survey items related to study outcomes only (i.e. excludes socio-demographic questions). This has now been clarified in the revised table title: “Table 1. Survey items related to study outcomes, including response options”.

There are only 13 questions being reported in this manuscript. Two additional questions related to the positive impacts of COVID-19 have been reported elsewhere (Cornell et al., 2022). Reference to ‘fifteen questions’ has now been changed to ‘thirteen questions’ throughout the manuscript, including in the abstract and the methods section. We have now also noted that other survey items are reported on elsewhere in the methods section: “This survey formed part of a larger study that examined COVID-19-related behaviour and

intentions, information sources, and impacts. Survey items reported here are those which were included in the current analysis. All other items are reported elsewhere (30-32).”.

Comment: Page 12 – what is a ‘tertiary qualification’?

- Response: We have removed the reference to tertiary qualifications, and now refer to the percentage of participants with a University bachelor degree or higher: “29.7% had University bachelor degree level or higher”.

Comment: If the authors are going to focus on gender, age, IRSAD quintile, and comorbidities in their results, this needs to be addressed.

- Response: The ‘Measures’ section of the manuscript has been amended to include additional detail about these variables. We now also refer readers to Table 2 for response categories.

“Demographic survey items relevant to this study included age, gender, education, whether born in Australia, years living in Australia, main language spoken at home, self-reported English language proficiency and a single-item health literacy screener (33). See Table 2 for response categories. Chronic disease status was determined by asking participants to self-report if their doctor had ever told them they had had one or more of the following: respiratory disease, stroke, asthma, diabetes, chronic obstructive pulmonary disease, depression, anxiety, high blood pressure, cancer or heart disease. The socioeconomic status of the area of residence for each individual was defined based on the SEIFA Index of Relative Socioeconomic Advantage and Disadvantage (IRSAD (34)). IRSAD aligns the statistical local area with a decile ranking (1–10), with lower scores indicating greater socioeconomic disadvantage. The IRSAD decile was not available for some participants (n=5), for example, because they had entered digits that did not correspond to a valid Australian postcode. IRSAD decile for these participants was replaced with the median IRSAD decile for speakers of the same language in the sample. For the analysis, IRSAD deciles were recoded into quintiles, and dichotomised (lowest quintile vs other”).

Comment: Authors need to elaborate more on the Mean Negative Impact on Children as well as the Mean Financial Burden measures in the METHODS.

- Response: We have revised page 12 of the methods section to include additional information about the Mean Negative Impact on Children and Mean Financial Burden scores, including ranges and reverse coding: “A mean ‘perceived financial burden’ score was also calculated by averaging the two questions about financial impacts: a) worry about financial problems and b) ability to meet weekly expenses (reverse coded). Higher scores indicate greater perceived financial burden (range: 1-5). Similarly, a mean score for the impact on children was calculated by averaging questions related to four impacts: physical activity, screen time, schooling and time with friends. Higher scores indicate more negative impacts on children (range: 1-5)” Cronbach’s Alpha for this scale was 0.805, indicating a high level of internal consistency.

Comment: If they are presenting the n (%) for each of these individual indicators, then why are they computing means? They should choose to do one or the other. What is the scientific justification for computing these means?

- Response: It is common practice in psychology, the social sciences and medical education to develop Likert-type items, group them into a survey scale, and then calculate a total score or mean score for the scale items (Sullivan & Artino, 2013). This practice has been recommended especially when researchers are attempting to measure less concrete concepts, such as those included in the current study —where a single survey item is unlikely to be capable of fully capturing the concept being assessed (Sullivan & Artino, 2013). Collapsing categorical responses on single items of a scale also effectively reduces the information available.

Questions relating to financial burden were taken from a validated scale; the COmprehensive Score for financial Toxicity (COST) scale (Souza et al., 2017). We

adapted two items (FT3 'I worry about the financial problems I will have in the future as a result of my illness or treatment' and FT7 'I am able to meet my monthly expenses') from the original scale to be relevant to the COVID-19 context. In line with the approach taken by the original authors, we then combined the COST scale questions and computed an overall mean.

Questions contributing to the mean negative impact on children were combined based on their conceptual similarity, Cronbach's Alpha for this scale was 0.805, indicating a high level of internal consistency.

References:

Souza JA, Yap BJ, Wroblewski K, Blinder V, Araújo FS, Hlubocky FJ, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: The validation of the COmprehensive Score for financial Toxicity (COST). *Cancer*. 2017;123(3):476-84.

Sullivan GM, Artino AR Jr. Analyzing and interpreting data from likert-type scales. *J Grad Med Educ*. 2013;5(4):541-542. doi:10.4300/JGME-5-4-18

Comment: Is there a validated approach to computing a mean negative impact on children score from screen time, physical activity, time with friends, and schooling? What is the Cronbach's alpha for these items?

- Response: We computed the mean negative impact on children by combining questions which were conceptually similar in their focus. As now reported on page 12 of the manuscript, Cronbach's Alpha for this scale was 0.805, indicating a high level of internal consistency.

Comment: Is there a validated approach to computing a mean financial burden score from changing employment status, concern about financial problems, and inability to meet weekly expenses? What is the Cronbach's alpha for these items?

- Response: As noted above, In addition, questions relating to financial burden were taken from a validated scale; the COmprehensive Score for financial Toxicity (COST) scale (Souza et al., 2017). We adapted two items (FT3 'I worry about the financial problems I will have in the future as a result of my illness or treatment' and FT7 'I am able to meet my monthly expenses') from the original scale to be relevant to the COVID-19 context. In line with the approach taken by the original authors, we then combined the COST scale questions and computed an overall mean.

The COST measure demonstrated excellent internal consistency, with a Cronbach α of .92 (Souza et al., 2017).

References:

Souza JA, Yap BJ, Wroblewski K, Blinder V, Araújo FS, Hlubocky FJ, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: The validation of the COmprehensive Score for financial Toxicity (COST). *Cancer*. 2017;123(3):476-84.

ANALYSES

Comment: The authors talk about weighting for their frequencies, but the results being presented look like they are unweighted. Typically, when weighted estimates are being presented, they are accompanied by a confidence interval for the estimated prevalence. This does not appear to be the case here. Or is it really the case that N=708 is a weighted sample size derived from a smaller sample?

- Response: The total sample was N=708, and the frequencies presented are weighted (as indicated on pg 11). We derived simple post-stratification weights, which is a post hoc procedure to adjust the sample to be more representative of the population distribution in terms of certain characteristics (here, age and sex within each language group). Weights are calculated and applied to each observation based on the relative frequency of the age / sex strata in the population (i.e., population relative frequency divided by sample relative frequency). This results in weights that sum to the obtained sample size.

We have now included the following text as a footnote in Tables 3-5: “n=1 respondent excluded from the weighted analyses included in this table; weighted frequencies have been rounded to whole numbers for clarity”.

Comment: The Mean Perceived Financial Burden and Mean Impact on Children do not seem to represent validated measures. This is problematic.

- Response: The ultimate goal of this study was to co-design a pragmatic measurement tool which included questions which were both easy to understand and could inform policy and support packages for culturally and linguistically diverse communities. As researchers, we were faced with a challenge in that validated instruments are often lengthy and written well above recommended grade-reading levels, with the potential to exclude the very communities which we sought to support. As such, each item/tool included in this study was carefully and purposefully selected to achieve our goal, while avoiding unnecessary repetition. Please also see our response to an earlier comment on Page 8 of this letter for further details about the calculation of means.

Comment: It is unclear what the difference is between unadjusted and adjusted analysis.

- Response: We have revised the ‘Analysis’ section of the manuscript to make the difference between adjusted and unadjusted analyses clearer: “Unadjusted and adjusted regression analyses were then conducted to explore the predictors of COVID-19-related psychological, social, and financial outcomes. Linear regression models were used to analyse perceived financial burden (averaged across two impacts) and impacts on children (averaged across four impacts). Logistic regression models were used to analyse psychological impacts (feeling lonely or alone; feeling nervous or stressed) and impact on relationships. Age group, gender, chronic illness, education, health literacy, English-language proficiency, years lived in Australia, language group and IRSAD quintile were included in each adjusted regression model. Models predicting impacts on relationships also controlled for perceived public health threat of COVID-19, perceived financial burden and psychological variables; models predicting psychological impacts controlled for perceived public health threat of COVID-19 and perceived financial burden. All regression models also controlled for whether participants completed the survey before or after 23rd June, when restrictions were announced for all of Greater Sydney (23). In line with recommendations, bivariable significance was not used as a criterion for variable selection in multivariable modelling (32, 33). The significance level of 0.05 was used to determine significant differences.”

RESULTS

Comment: The way the Results are presented, the cultural and linguistic diversity aspect seems to take a lower priority over the demographic predictors (i.e., age group and gender). There should be a greater emphasis on the disparities according to health literacy and language spoken, since these pieces are what are most compelling about the study. The results should highlight what was found with regard to health literacy, language group, English language proficiency, years living in Australia.

- Response: We have significantly revised the Results section of the manuscript so that cultural and linguistic diversity are given higher priority. An example of this in relation to the

“Psychological impacts

Overall, 25.3% of participants reported feeling nervous or stressed most or all of the time over the past week. This ranged across language groups from 6% (n=5) for Chinese speakers to 38% (n=24) for Dinka speakers. 30.7% (n=89) of participants with inadequate health literacy and 21.4% (n=89) of participants with adequate health literacy reported feeling nervous or stressed most or all of the time. This was XX% for those who self-reported that they speak English well or very well, compared to XX% of those who speak English not well or not at all. See Table 3, which also outlines further sociodemographic differences. In the multivariable regression model when sociodemographic factors were controlled for, language group ($p<0.001$), female gender ($p=0.04$) and having two or more chronic illnesses ($p<0.001$) remained significantly associated with increased nervousness or stress, as did higher perceived financial burden ($p<0.001$). See Supplementary Table 1.

Overall, 22.3% of participants reported feeling alone or lonely most or all of the time. In regards to language groups, the range was from 5.6% (n=2) for Hindi speakers to 51.2% (n=32) for Khmer speakers. 27.8% (n=81) of participants with inadequate health literacy reported feeling alone or lonely most or all of the time; this proportion was 18.5% for participants with adequate health literacy (n=77). This was XX% for those who self-reported that they speak English well or very well, compared to XX% of those who speak English not well or not at all. See Table 4. In the multivariable regression model, having two or more chronic illnesses ($p<0.001$) and university education ($p<0.001$) remained as significant correlates of feeling lonely or alone, with statistically significant differences also observed between language groups ($p<0.001$)."

Language group and English-language proficiency information have also been added to Tables 3-5. Comment: Table 2. Descriptive Statistics of Analysed Sample

This table needs to include the proportion of participants with children aged less than 18 years, since the analysis on negative impacts of children only pertains to this subgroup

- Response: We have now included the n(%) of people who had children less than 18 years in Table 2.

Comment: Table 2. Descriptive Statistics of Analysed Sample

The descriptive statistics for the outcomes presented in the ALL column from Table 4 should actually be here, so readers can see the prevalence of the psychological, social, and financial correlates from the full sample.

- Response: Thank-you for your suggestion. We agree that the prevalence of the psychological, social, and financial correlates from the full sample were difficult to locate in the previous version of the manuscript. To address this, have now included a 'Total' row at the top of Tables 3-5 to showcase the descriptive statistics for each outcome.

Comment: Table 2. Descriptive Statistics of Analysed Sample

Whether participants received assistance with completing the survey may be worth including in this table as well.

- Response: We have now included the weighted n(%) of people who required assistance in completing the survey in Table 2.

Comment: Table 3 and Table 4

It might be easier for the readers if Tables 3 and 4 were reorganized into three separate tables, one for psychological, one for social, and one for financial COVID-19 correlates, with the demographic predictors (gender, age group, health literacy, IRSAD quintile, comorbidities, and language group on the side)

- Response: Tables 3 and 4 have now been reorganised into three separate tables (Tables 3, 4 and 5) in line with the comment above. All tables have been transposed so that sociodemographic predictors are presented as rows, and outcomes are presented as columns.

Comment: Table 3 and Table 4

These tables should also prioritize the presentation of results based on the health literacy and language group predictors, since these are supposed to be the star of the of study.

- Response: We originally decided not to include data for individual language groups as we felt direct comparisons between individual languages can be stigmatising and is not pragmatically relevant for the purpose of this manuscript over and above knowing that there are differences between language groups. However, in light of your comment, we have now included this information.

Comment: Supplemental Tables 1, 2, & 3

It is highly unclear what the difference is between the unadjusted and adjusted models. This applies to all Supplemental Tables

- Response: We have included the following footnotes in Supplementary Tables 1, 2 and 3 to highlight the difference between the unadjusted and adjusted models.

Supplementary Table 1:

^b Unadjusted analyses do not control for co-variables; statistics represent the regression of each predictor on psychological outcomes with no other co-variables included in the model.

^c Adjusted analyses control for all covariates listed in this table.

Supplementary Table 2:

^c Unadjusted analyses do not control for co-variables; statistics represent the regression of each predictor on social outcomes with no other co-variables included in the model.

^d Adjusted analyses exploring factors associated with negative impacts on relationships control for all covariates listed in this table.

^e Adjusted analyses exploring factors associated with negative impacts on children do not control for perceived public health threat, financial burden or psychological outcomes (lonely/alone; nervous/stressed).

Supplementary Table 3:

^b Unadjusted analyses do not control for co-variables; statistics represent the regression of each predictor on financial burden with no other co-variables included in the model.

^c Adjusted analyses control for all covariates listed in this table.

Comment: There are some p-values that are misplaced. What do the bolded p-values refer to? Are these results from Wald tests?

- Response: We apologise for any confusion caused by the bolded font in the supplementary tables. We now only bold statistically significant p-values in line with more standard convention.

Comment: There are some p-values that are misplaced. For the Alone/Lonely unadjusted results (Supplemental Table 1), why is there a P-value for the Reference group for co-morbidity? That does not belong there.

- Response: Apologies for this typographical error. The p-value for the Reference group for co-morbidity in the Alone/Lonely unadjusted results (Supplemental Table 1) has now been removed.

Comment: Authors should report the analytic sample size for these regression models, as some models (i.e., negative impacts on children) may actually be a subsample of the 704 participants. This analysis would only pertain to respondents with children younger than 18 years old.

- Response: That is correct; some analyses (e.g. impacts of COVID-19 on relationship with partner; impacts of COVID-19 on children) include only a subsample of the total 708 participants. N's for each analysis are included as footnotes in Tables 3 and 4. The number of participants responding to each question is also highlighted in the text, e.g.:

“Of the 399 participants who responded to the question regarding impacts of COVID-19 on their relationship with their partner, one quarter (25.5%) reported negative effects; 62.9% said that the pandemic had no effect and 11.7% said that it had had positive effects.”

“Of the two hundred and sixty-two participants who reported having children aged less than 18 years, 72.8% reported spending more time looking after their children as a result of the pandemic (n=191).”

IMPACTS

Comment: Page 15 on, plus Tables 3 and 4. While there are very useful comparisons noted, it is not clear to what extent the patterns noted predated COVID, reflecting broader societal shortfalls; were exacerbated by them, reflecting inadequate COVID response, likely on top of systemic failings; or even were relatively unaffected, reflecting community resilience. Part of the issue is that the (very comprehensive) data displays make it hard to identify fields for which comparable data is available.

- Response: Thank-you for your reflections. We agree and have added a statement about this in the discussion of the manuscript on page 23.

Comment: A major strength of the study is its focus on health literacy, English-language proficiency, language group, and years living in Australia predictors. These predictors are what make this study compelling, yet greater emphasis has been placed on the demographic predictors (e.g., gender and age) and just overall descriptives. This is somewhat of a missed opportunity that could really inform targeted intervention efforts.

The discussion should highlight what was found with regard to health literacy, language group, English language proficiency, years living in Australia.

- Response: As noted above, we have significantly restructured the results of our manuscript to focus less on demographic predictors such as age and gender and more on health literacy, English-language proficiency, language group. We have also significantly revised the discussion on Pages 22 to 25.

DISCUSSION

Comment: There needs to be a deeper discussion relating the children/families outcome as well as the financial burden outcomes to the cultural and linguistic needs of diverse communities.

- Response: Our revised discussion now includes addition reference to financial outcomes and outcomes for children in the second paragraph.

IMPLICATIONS

Comment: The solutions that the authors propose need to be closely tied to the specific findings from their analyses.

What is it about the authors' results that tell us about the need for system readiness or culturally safe support packages?

What is it about the authors' results that tell us about communal ways of coping?

What is it about the authors' results that tell us about the need for migrant-inclusive public information strategies?

- Response: We have significantly revised the 'Implications' section of the manuscript to be better aligned with our findings.

“Our findings showcase a broad range of impacts of COVID-19 among culturally and linguistically-diverse Australian communities. A multi-level, whole-of-government approach is needed to address these, with policy and sustainable infrastructure to disseminate timely, understandable and culturally-appropriate information about financial, social and mental health resources and services and to co-design tailored support packages for different language groups (40). Qualitative studies have highlighted a large number of community-driven initiatives and actions that have emerged as a response to COVID-19, as well as embodied and communal ways of coping (41). Using a strengths-based perspective, we must acknowledge the multiple capacities and resources of our culturally and linguistically diverse communities and provide properly-resourced opportunities to work directly with them to address unique challenges that they face, as identified in this study. Our findings reinforce the need to prioritise support for community members living with comorbidities who are likely to bear a disproportionate impact.”

Comment: There is an interesting pattern, not clearly spelled out, which seems to tell a story of resilience across these communities:

Page 20, line 56 – 25% of participants reported feeling nervous or stressed most or all of the time (national, April 2020)

Page 21, line 5 – 20% of Australians experienced high or very high levels of psychological distress; 28% reported feeling nervous (national, June 2021)

This study – 25.3% reported feeling nervous or stressed most or all of the time (NSW, March- July 2021)

If this pattern of community resilience is correct, then it may be an important finding to highlight, in addition to the multiple findings of negative impacts.

- Response: While differences across studies (including population groups, time of data collection, varying survey items, wide-ranging case numbers, morbidity and mortality from COVID-19) make definite conclusions about resilience difficult, we have now noted the possible pattern of community resilience in the discussion on Page 23.

Reviewer 2

I would like to thank the authors for this work. This is an important study that can contribute to effective support of CALD communities during and after COVID-19. However, there are major issues to be addressed which I believe will make the paper stronger. I am particularly concerned with the regression findings presented in this study. I have presented these issues below, and you will also find them commented on your paper. I hope you will find these helpful.

Introduction:

Comment: Page 6 of 36 line 7 - The word 'ethnic' is considered inappropriate these days because of its colonial history, but also because of its ambiguity (Does it refer to race, religion, tribes, culture, country of origin, language groups etc?) I would suggest using "people from racial and cultural minority groups"

- Response: This sentence has now been revised to "People from racial and cultural minority groups in countries across the globe have been disproportionately affected...".

Comment: Page 7 of 36 lines 9-17. A further clarification is required for what the author means by "migrants". Did they mean people who came to Australia for economic opportunities other than being forced to flee their home country? One can argue that "migrants" can also include refugees and asylum seekers if not specified by the author.

- Response: Thank-you for raising this important point. The introduction of the manuscript has been substantially revised based on reviewers' feedback.

Comment: Page 7 of 36 line 22 is not clear, rephrase. "... differences in financial and psychological impacts of COVID-19 those for who spoke a language other than English at home ...". The author seems to have missed a word between "COVID-19" and "those".

- Response: Thank-you for spotting this typographical error. This sentence now reads: "Our own Australian surveys (and others – see, (12, 13)) have shown some differences in financial and psychological impacts of COVID-19 among those who speak a language other than English at home compared to those for whom English is their primary language."

Methods:

Comment: Page 8 of 36 line 21-22 – The author wrote that their study was informed by the framework for culturally competent health research, but no further information is provided about this framework. Mentioning a framework without saying how it informs the study is not sufficient. The author needs to elaborate on how the framework used has informed the study and its analysis.

- Response: Thank-you for your comment. We have now included Box 1 which outlines in detail how the Framework was applied in the current study. Box 1 is reproduced on page 4 of this document.

Comment: Page 8 of 36 line 58-60 - The author stated in their introduction that their study focuses on CALD community because the experience of this community is different from refugees who have come to Australia within the last 10 years. In lines 9-17 on page 7 of 36, the author's rationale for the study is to focus on "those who speak a language other than English at home who have not been forced to flee their home country". However, all most all Dinka (a tribe from South Sudan), Khmer,

Assyrian and some Arabic speaking people came to Australia as refugees because they have been forced to flee their home countries. Including these communities as participants contradict the rationale provided in the introduction in lines 9-17. The author needs to elaborate how they excluded refugees and asylum seekers from their study? As it stands, the information about the participants and the rationale for their study do not match.

- Response: Reference to “those who speak a language other than English at home who have not been forced to flee their home country” was not intended to provide a rationale for this study, but rather to highlight a point of difference between our work at that which has already been conducted. Although our primary aim was not to specifically explore the impacts of COVID-19 among refugees and asylum seekers, we did not systematically exclude these groups from our study.

Given the potential for confusion, the introduction of the manuscript has been substantially revised and no longer includes reference to “those who speak a language other than English at home who have not been forced to flee their home country”.

Comment: Page 10 of 36 line 13 – The author said they ‘adapted’ a validated tool. What were the reported KMO and Cronbach’s alpha for the validated instrument? Did the author conduct a revalidation of the instrument after “adapting” it?

- Response: We have now included additional information about the selection of survey items, and adaptation of validated instruments in the Measures section of the manuscript on Page 10. For example: “Items related to financial impacts were adapted from the COmprehensive Score for financial Toxicity (COST) scale (31). We adapted two items (FT3 ‘I worry about the financial problems I will have in the future as a result of my illness or treatment’ and FT7 ‘I am able to meet my monthly expenses’) to be relevant to the COVID-19 context.”. We did not conduct a revalidation of the instrument after adapting it.

Comment: The referenced validated tool in line 13 on page 10 of 36 was developed to measure “financial toxicity of cancer within a public healthcare system” in Italy. Was this a relevant tool to adapt to measure the financial and psychological impacts of COVID-19? I doubt. It is not exactly clear which part of the validated instrument was ‘adapted’ for this manuscript. A look at the referenced work indicate that it was an abstract not a fully published paper. A further search found a publication in BMJ and the reported questionnaire did not match the items reported under the financial and psychological items. You can see the published ‘validated tool’ from Riva and colleagues’ paper below.

Riva S, Arenare L, Di Maio M, et al. Cross-sectional study to develop and describe psychometric characteristics of a patient reported instrument (PROFFIT) for measuring financial toxicity of cancer within a public healthcare system. *BMJ Open* 2021;11:e049128. doi:10.1136/bmjopen-2021-049128
I would suggest the author reference kessler psychological distress scale here instead of the Riva’s paper.

- Response: Sincere apologies for including the incorrect reference. This has now been rectified:

Souza JA, Yap BJ, Wroblewski K, Blinder V, Araújo FS, Hlubocky FJ, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: The validation of the COmprehensive Score for financial Toxicity (COST). *Cancer*. 2017;123(3):476-84.

Comment: Page 10 of 36 line 13 - Reference (14) is an abstract. Author needs to reference the published version in BMJ.

Riva S, Arenare L, Di Maio M, et al. Cross-sectional study to develop and describe psychometric characteristics of a patient reported instrument (PROFFIT) for measuring financial toxicity of cancer within a public healthcare system. *BMJ Open* 2021;11:e049128. doi:10.1136/

- Response: Thank you this error has now been rectified as above.

Comment: Page 10 of 36 line 13-16 - Was the 'co-designed' questions pilot tested? Did the author perform factor analysis or principal component analysis to determine the validity of these items?

- Response: The co-designed questions were not pilot tested, but were reviewed by the entire study team (including Multicultural Health and Health Care Interpreter Service staff and bilingual community workers) as well as members of the general public before implementation to ensure relevance, readability, and clarity of items for community members. As stated elsewhere, this has now been included in Box 1 of the manuscript. We did not perform factor analysis or principal component analysis to determine the validity of these items given time constraints.

Comment: Page 12 of 36 line 12-16 – Is it appropriate for the author to use Likert scale to measure “mean”? I doubt this is the right approach.

- Response: As referenced in our response to Reviewer 1, it is common practice in psychology, the social sciences and medical education to develop Likert-type items, group them into a survey scale, and then calculate a total score or mean score for the scale items (Sullivan & Artino, 2013). This practice has been recommended especially when researchers are attempting to measure less concrete concepts, such as those included in the current study — where a single survey item is unlikely to be capable of fully capturing the concept being assessed (Sullivan & Artino, 2013). Combining responses on a scale also effectively reduces information available and assists with interpretation of large data sets.

Sullivan GM, Artino AR Jr. Analyzing and interpreting data from likert-type scales. *J Grad Med Educ.* 2013;5(4):541-542. doi:10.4300/JGME-5-4-18

Comment: Page 12 of 36 line 16-18: Any objective evidence to support this measurement? OR they were arrived at subjectively? If subjective and without the validation of the tools, how can a reader rely on the findings?

- Response: The ultimate goal of this study was to co-design a pragmatic measurement tool which included questions that were both easy to understand and could inform policy and support packages for culturally and linguistically diverse communities. As researchers, we were faced with a challenge in that validated instruments are often lengthy and written well above recommended grade-reading levels, with the potential to exclude the very communities which we sought to support. As such, each item/tool included in this study was carefully and purposefully selected to achieve our goal, while avoiding unnecessary repetition.

We have now included additional information about the selection of survey items of page 10 of the manuscript: “Thirteen items regarding the impacts of COVID-19 were selected for this survey study in partnership with Multicultural Health and Health Care Interpreter Service staff. See Table 1. Items related to financial impacts were adapted from the COmprehensive Score for financial Toxicity (COST) scale (31). We adapted two items (FT3 'I worry about the financial problems I will have in the future as a result of my illness or treatment' and FT7 'I am able to meet my monthly expenses') to be relevant to the COVID-19 context. Psychological items were taken verbatim from our previous COVID-19 work (18). Questions regarding social impacts (including impacts on relationships and children) were co-designed with Multicultural Health and Health Care Interpreter Service staff based on local information priorities.”

Finally, in addition, we calculated the Cronbach’s Alpha for the impact on children scale, which was found to be 0.805, indicating a high level of internal consistency. This provides support that readers can rely on the findings of this study.

Comment: Page 12 of 36 line 18-22: Ditto

- Response: Please see our response above.

Comment: Page 12 of 36 lines 25-31: What is the significance level used to determine significant differences?

- Response: The significance level of 0.05 used to determine significant differences was 0.05. This has now been included in the methods section of the manuscript: “The significance level used to determine significant differences was 0.05”.

Results

Comment: Page 13 of 36 line 7: Is advanced diploma or diploma level not a tertiary qualification? If this is the case, the author needs to make it clear. I would suggest instead of reporting the negative (no tertiary qualification), the author needs to be specific and use a positive tone, e.g. 29.7% had bachelor degree level or higher. OR 70% had advanced diploma level education or lower.

- Response: Thank-you. This has been revised as suggested: “29.7% had University bachelor degree level or higher”.

Comment: Table 2 on page 14 of 36: Assyrian, Arabic speaking people, Dinka, Khmer have come to Australia as refugees. As previously pointed out, including these population groups conflict with the study's rationale at the introduction section.

- Response: Thank you – this has now been rectified in the introduction.

Comment: Table 2 include “Years of living in Australia” categorised into “5 years or less”, “6 to 10 years” and “more than 10 years”. See the previous comments on participants inclusion and exclusion. Given that there are participants who lived in Australia for 5 years or less are included in the study, how did the author exclude those who are refugees from the included communities?

- Response: Refugees were not systematically excluded from this study. As outlined above, this has now been rectified in the introduction.

Comment: Page 16 of 36 line 26: How did the author determine whether participants had two or more chronic illnesses? The survey instrument in Table 1 does not contain question on participant disclosing their comorbidities.

- Response: Table 1 includes survey items related to study outcomes only. It does not include predictive variables. We have now amended the study title to reflect this: “Survey items related to study outcomes, including response options”.

Additional information about the determination of comorbidities has now been included in Measures section of the manuscript: “Chronic disease status was determined by asking participants to self-report if their doctor had ever told them they had had one or more of the following: respiratory disease, stroke, asthma, diabetes, chronic obstructive pulmonary disease, depression, anxiety, high blood pressure, cancer or heart disease.”

Comment: The following issues were observed relating to all the regression analysis reported on page 16 of 36 line 22-29; Page 16 of 36 line 47-53; Page 17 of 36 line 10-25; Page 17 of 36 line 40-49, and Page 18 of 36 line 25-31:

Regression analysis finding provided relating to supplementary table 1. “Lowest ISRAD quintile” and “Years living in Australia” are not significant at 0.05 level for unadjusted analysis (Nervous/stressed and Alone/Lonely). Why did the author include these factors in the adjusted analysis when they don't make any contribution into improving the model?

- Response: As suggested by Agresti (2002) and Harrell (2015), removal of non-significant variables can invalidate the estimate of the model variance, as well as the standard errors of the other coefficients in the regression model (which in turn, affects the p-values and the confidence interval width of the remaining coefficients); removal of the “non-significant” variables would also amount to imposing that the population regression coefficient is exactly zero, which is not likely to be the case. Further, as Sun et al (1996) indicate, the use of bivariable significance as a criteria for variable selection in multivariable modelling can be considered inappropriate as it risks incorrectly rejecting potentially important variables if the relationship between an explanatory variable and the outcome is confounded by any other confounder and when this confounder is not properly controlled for in the model. Therefore, while this approach may result in a less parsimonious model, the aim of these analyses is in effect estimation, with pre-selection and inclusion of theoretically / conceptually relevant (and available) factors irrespective of their statistical significance in univariate modelling.

We have now noted this decision in the Analysis section of the manuscript: "In line with recommendations, bivariable significance was not used as a criterion for variable selection in multivariable modelling (32, 33)."

References:

Agresti, Alan. 2002. *Categorical data analysis*. Hoboken, NJ: Wiley
Harrell, Frank E. 2015. *Regression modeling strategies*. Cham: Springer
Sun, G.-W., Shook, T.L., and Kay, G.L. 1996. Inappropriate use of bivariable analysis to screen risk factors for us in multivariable analysis. *J Clin Epidemiology* 49(8):907-916.

Comment: Supplementary table 2: All the highlighted factors (unadjusted analysis) were not statistically significant at 0.05 significance level and therefore are not making any contribution to improve the regression model. They shouldn't be included in the final model (adjusted analysis) because they are not significant explanatory variables in this case.

- Response: Please see our response to the comment above.

Comment: Supplementary table 3: The variable "Lowest IRSAD quintile" is not significant at 0.05. Why did the author include them in adjusted analysis when clearly it is not contributing to the improvement of the model without the other covariates? The same issue is observed for Adequate health literacy, English-language proficiency, and gender. I would suggest the author remove all these factors from the adjusted model. I am concerned with the regression finding presented in this study, which in my opinion requires major revision.

- Response: Please see our response to the comment above.

Comment: Page 17 of 36 Line 38-40 (including Table 3 and 4): How can the Likert score be used to measure the "mean financial burden" and the "mean negative impact on children"?

- Response: As noted above, it is common practice in psychology, the social sciences and medical education to develop Likert-type items, group them into a survey scale, and then calculate a total score or mean score for the scale items (Sullivan & Artino, 2013). This practice has been recommended especially when researchers are attempting to measure less concrete concepts, such as those included in the current study —where a single survey item is unlikely to be capable of fully capturing the concept being assessed (Sullivan & Artino, 2013). Combining responses on a scale also effectively reduces information available and assists with interpretation of large data sets.

In addition to this, financial burden items were adapted from a previously validated financial toxicity scale (the COmprehensiv Score for financial Toxicity (COST)) which used a similar approach whereby individual items were combined to create a mean score.

References

Sullivan GM, Artino AR Jr. Analyzing and interpreting data from likert-type scales. *J Grad Med Educ*. 2013;5(4):541-542. doi:10.4300/JGME-5-4-18

Comment: Likert scale data cannot use the mean measure central tendency because it has no meaning numerically. Instead of using the mean, I would suggest the author use % (combined agree and strongly agree) to report their finding in Table 3 and 4 for the two outcomes. Alternatively, they can remove the information on mean from Table 3 and 4.

- Response: Please see our response to the comment above.

Comment: Page 18 of 36 lines 11-13: Ditto. I would suggest rewriting the whole of the remainder of this section (starting from the highlighted) with the comments in mind. Reporting a Likert data as a mean to measure central tendency is a major flaw.

- Response: Please see our response to the comment above.

Comment: Table 3: As reported in Table 3 and 4, it is useful for the author to describe how they grouped the survey items under psychological impacts, social impacts, financial impacts. Did they conduct PCA or factor analysis to determine these?

- Response: As now noted in Box 1, survey items were developed in partnership with Multicultural Health and Health Care Interpreter Service staff. This included the selection of broad outcome domains (psychological, social and financial impacts) as well as individual questions. Questions were grouped into impact domains theoretically, based on their content. We did not conduct PCA or factor analysis to determine these.

Discussion:

Comment: I was a bit surprised by the discussions, which appeared limited in scope and nuances. I suggest the author thoroughly discuss their findings instead of resorting to making a general statement.

- Response: We have substantially revised the discussion on pages 22 to 25.

Comment: Discuss the findings on page 16 of 36 line 16-20

Page 21 of 36 line 43-45: Where are the evidence to support this statement?

- Response: We have substantially revised the discussion on pages 22 to 25.

Comment: Page 21 of 36 line 51-55: I would suggest discussing what others discussed first before discussing your own previous study.

- Response: Our discussion has been restructured such that we now discuss previous research conducted in the United States and Australia, followed by reference to our own nationally representative survey.

VERSION 2 – REVIEW

REVIEWER	Michener, Lloyd Duke University, Family Medicine & Community Health
REVIEW RETURNED	07-Feb-2022

GENERAL COMMENTS	<p>Thank you for the opportunity to review the revision of Manuscript ID bmjopen-2021-058323.R1 entitled "Psychological, social, and financial impacts of COVID-19 on culturally and linguistically diverse communities in Sydney, Australia" The authors have thoughtfully responded to the many questions and concerns raised, and the revised manuscript is much clearer, and much better described and documented. The only change to consider at this point is placing the study more clearly in the context of the NSW COVID outbreak and the 'lockdowns' imposed. Specifically:</p> <p>Current text (page 8)</p> <p>"Setting The survey was conducted from 21 March to 9 July, 2021. During this period, rollout of the COVID-19 vaccines had begun across Australia, and daily cases in New South Wales (NSW) were very low by international standards, ranging from 0 – 46 positive cases from a population of approximately 8 million people (23). A 'stay at home' order across Greater Sydney due to rising cases began on June 23rd (24). On the day the survey closed the NSW daily case count was 45, and 24% of the population had received one COVID-19 vaccination (25). "</p> <p>But this leaves some ambiguity about when the pandemic began, and what restrictions were in place during the study period. Something like the following would be helpful to those not familiar with the NSW time course:</p>
-------------------------	--

	<p>"The survey was conducted from 21 March to 9 July, 2021. The first case of COVID-19 in Australia was detected in January 2020; a national emergency with social distancing rules and closure of "non-essential" services followed in March 2020. Peaks of infection occurred in March 2020, and May-June 2020, and vaccinations begun in February 2021. During the study period, daily cases in New South Wales (NSW) were very low by international standards, ranging from 0 – 46 positive cases from a population of approximately 8 million people (23). A 'stay at home' order across Greater Sydney due to rising cases began on June 23rd 2021 (24). On the day the survey closed the NSW daily case count was 45, and 24% of the population had received one COVID-19 vaccination (25). "</p> <p>Thanks and congratulations to the authors for their thoughtful and helpful work.</p>
--	--

VERSION 2 – AUTHOR RESPONSE

Comment:

Thank you for the opportunity to review the revision of Manuscript ID bmjopen-2021-058323.R1 entitled "Psychological, social, and financial impacts of COVID-19 on culturally and linguistically diverse communities in Sydney, Australia". The authors have thoughtfully responded to the many questions and concerns raised, and the revised manuscript is much clearer, and much better described and documented. The only change to consider at this point is placing the study more clearly in the context of the NSW COVID outbreak and the 'lockdowns' imposed. Specifically:

Current text (page 8)

"Setting

The survey was conducted from 21 March to 9 July, 2021. During this period, rollout of the COVID-19 vaccines had begun across Australia, and daily cases in New South Wales (NSW) were very low by international standards, ranging from 0 – 46 positive cases from a population of approximately 8 million people (23). A 'stay at home' order across Greater Sydney due to rising cases began on June 23rd (24). On the day the survey closed the NSW daily case count was 45, and 24% of the population had received one COVID-19 vaccination (25). "

But this leaves some ambiguity about when the pandemic began, and what restrictions were in place during the study period. Something like the following would be helpful to those not familiar with the NSW time course:

"The survey was conducted from 21 March to 9 July, 2021. The first case of COVID-19 in Australia was detected in January 2020; a national emergency with social distancing rules and closure of "non-essential" services followed in March 2020. Peaks of infection occurred in March 2020, and May-June 2020, and vaccinations begun in February 2021. During the study period, daily cases in New South Wales (NSW) were very low by international standards, ranging from 0 – 46 positive cases from a population of approximately 8 million people (23). A 'stay at home' order across Greater Sydney due to rising cases began on June 23rd 2021 (24). On the day the survey closed the NSW daily case count was 45, and 24% of the population had received one COVID-19 vaccination (25). "

Thanks and congratulations to the authors for their thoughtful and helpful work.