PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Correlates of intended COVID-19 vaccine acceptance across time and countries: Results from a series of cross-sectional surveys
AUTHORS	Kerr, John; Schneider, Claudia; Recchia, Gabriel; Dryhurst, Sarah; Sahlin, Ullrika; Dufouil, Carole; Arwidson, Pierre; Freeman, Alexandra; van der Linden, Sander

VERSION 1 – REVIEW

REVIEWER	Williams, Lynn University of Strathclyde, School of Psychological Sciences & Health
REVIEW RETURNED	04-Feb-2021

GENERAL COMMENTS	The manuscript reports on a large data collection endeavour to understand intended acceptance of a COVID-19 vaccine. The sample size, multi-country focus, and broad range of demographic, social and psychological correlates that were included, are all strengths of the study. I have a few points of recommendation which mainly focus on the inclusion of further information at different points within the manuscript. 1. The title - I think the title may be a little misleading in the months to come, I would encourage the authors to specify that they are talking about intended vaccine acceptance, rather than actual acceptance. Soon papers will appear that examine actual uptake and think it will be important that there is a clear distinction between those that looked at intended acceptance and those that looked at uptake. I would also encourage the authors to reconsider the use of the word predictors here, and elsewhere in the manuscript. The word correlates may be more appropriate given the cross-sectional nature of the data (the word correlates is used in the abstract for example). Elsewhere talking about the factors associated with uptake, rather than predicting uptake would be more appropriate. 2. Abstract – in the results it would be useful to report the range of % uptake figures across countries (i.e. 62.6% - 88.1%). 3. The introduction provides a good overview of the literature in this area, both in terms of likely COVID-19 vaccine hesitancy and vaccine hesitancy in general. The literature is fast moving in this area, but I would not suggest any changes to the intro. 4. In the method it would be useful to add some additional contextual information. For example, an indication of the timeline of the COVID-19 vaccine development process during the data collection period of the surveys would useful. 5. In addition, an indication of the infection pressure and the restrictions in place in each country at the time of data collection would also be helpful, perhaps as an additional column in Table 1. The heterogeneity in the factors associated with intended uptake

	may reflect the heterogeneity in the contexts within which the surveys were carried out, and so having some indication of the context would be useful and aid interpretation. 6. Clearly there is a huge amount of data here, and the figures that have been developed are excellent. The inclusion of the survey items within the supplementary files is also very welcome. 7. I appreciate that the authors have included a really broad range of measures but there are a few salient omissions in terms of socio-demographic factors which I think should be acknowledged as limitations. E.g. Evidence would suggest that ethnicity and socio-economic status are important factors to consider, and I was surprised that these were not included here. Presence of an underlying health condition, (or 'shielding' in UK terms) would also have been useful to assess. 8. Linked to this point, the authors do point to several groups where vaccine hesitancy may be more of an issue. In terms of the practical considerations for public health messaging, the importance of targeted messaging could be considered more in the discussion. In general, the practical considerations section is relatively brief, and I would welcome a more in-depth consideration of this. How else can the results here inform the targets of, or content of messages designed to increase uptake, or where should future research focus to help us get to this stage?

REVIEWER	Attwell, Katie University of Western Australia School of Social Sciences, School of Social Sciences
REVIEW RETURNED	05-Mar-2021

GENERAL COMMENTS	This well-written and engaging article reports the findings of surveys in multiple countries regarding COVID-19 vaccine uptake intentions and the factors that contribute. It dives deeper in US and UK where there were multiple surveys over time with different respondents. The article will need updating to reflect the current status of vaccine rollout since it was submitted. The authors should also update their literature review to reflect new studies that have come out since they submitted, including those that having similar findings, so as better to contextualise the contribution of the piece when it comes out. The authors should also reconsider / update their 'herd immunity' framing in the introduction as the vaccines' capacity to prevent transmission is yet to be determined in rollout scenarios and may be less than thought at the time of submission. This is a minor formatting point but Table 1 needs to be addressed so that it is easier to read without words split across lines. Can the authors explain why they used more than one survey company in UK and US, and why they think these companies found different results even when collecting data at the same time? Might it have made more sense to combine datasets when collected on same date (as UK Figure 3 indicates?). What are the strengths and limitations in their decisions surrounding these dual surveys? The findings about social trust and trust in government not being influential on vaccine acceptance are interesting, and the lack of significance of prosocial measures. The authors need to consider the latter point in light of the 'herd immunity' point above – COVID-19 vaccines will still inevitably have societal benefits, but what happens if we can't talk about them generating herd immunity?

	The study's findings are well-contextualised and the discussion connects well to the vaccine uptake research and make salient recommendations for future research.

REVIEWER	Wang, Wei Food and Drug Administration, Division of Biostatistics, Center for Devices and Radiological Health
REVIEW RETURNED	09-Mar-2021

GENERAL COMMENTS	In this manuscript, the authors presented a more comprehensive international analysis of the role of key social, political, and psychological predictors of COVID-19 vaccine acceptance across 12 countries, with multiple national surveys in some countries. From the statistical point of view, the analysis methods applied in this manuscript is generally acceptable and the results presentation is clear. I still have the following minor comments for the authors' consideration, (1) It is recommended that the authors include separate sections in the METHODS to describe the analysis methods used in the whole manuscript which will make the METHODS section more complete and the readers understand the statistical analysis methods used in the whole manuscript easily. Specifically, the authors need to elaborate the Table S2 metric invariance analysis a little bit, for example, what does CFI mean? (2) The authors compared the vaccine acceptance among respondents in Mexico between March and May, in US, between May and September. It is not clear whether the respondents are the same cohort or different cohorts. If they are the same cohorts, the two sample proportion test used is not correct. The authors may need consider data analysis methods to handle this correlated binary outcome, like GEE, or Generalized linear mixed models (or GLMMs). Even they are from different cohorts, it is also recommended to use Chi-Squared or Fisher's Exact Test to compare independent proportions. Although the z test used by the authors are not wrong, it does need additional normal approximation. Please clarify correct the manuscript accordingly. (3) The authors need to clearly specify using multiple logistic regression model instead of logistic regression model in the main analysis of the manuscript to avoid confusion.

REVIEWER	Wasfi, Rania Centre de recherche du CHUM
REVIEW RETURNED	28-Mar-2021

GENERAL COMMENTS	Peer Review: Predictors of COVID-19 vaccine acceptance across time and countries: Results from a series of cross-sectional surveys Thank you for conducting this research, it is an important topic that need to be addressed to help in understanding strategies governments can use to increase vaccination acceptance among people. The objective of the paper is clear, and important, however the way the methods are used and results presented are confusing and I think if modified will make an important contribution. In the current form now it is difficult to makes clear inference, without many unadjusted confounding.

	Please see below specific comments and suggestions: • The decision of stratifying the logistic regressions models by country is not explained and justified. Can you please elaborate on that • In Table S6: A pooled model is constructed, was the country and time of data collection adjusted for in the pooled model. I cannot see that in the Tables. If not, why is this the case? • Why is UK and USA samples excluded then analysed separately? There is no clear justification for that. • In Table S7. Why did not pool the data for the UK (for all months, and adjust for the month in one pooled the authors regression models. Again, it is very hard to interpret the results in this form, and understand the effect of each predictor independent from time. If authors think the predictor change over time, why did not they try an interaction variable with time? • In Table S8. The time indicates: “Full logistic regression results from model predicting vaccine recommendation to vulnerable other”- what is vulnerable other? Please elaborate. • In the logistic regression models, why did not the authors adjust for the company of recruitment, given that there were differences shown and presented by the authors in other Tables? • The authors presented a Table of missing data, but did not discuss how this missing data influence the generalizability of the results. I understand the sample was not a random sample but was matched by age and gender. Did this differ with missing values? • It is hard in the current format of the logit models to make between countries comparison. Why did not the authors use a nested model, where cases are nested within countries and adjusted for the country of survey? I suggest authors to try using multi-level logits to separate the random effects, and minimize bias, and be able to compare between countries. • It could be useful to indicate when the first cases of Covid-19 started in each country or indicate the number of cases at the time the survey was administered to get some context of what this time (Month) means during the pandemic. Moreover, use a multi-level logistic regression with pooled data (from all countries) adjusting for Country of survey, Time-from first covid-19 cases in the country, # of cases, and month, in addition to other covariates in the models. Minor comments: • On page 9 line 8 : (example: I will probably get sick with the coronavirus/COVID-19; αs .71-.89). For : “αs .71-.89” Why is findings reported in Methods? • On page 9 line 42. Please put an explanation and spell out the (CFI) on line 42.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Lynn Williams, University of Strathclyde

Comments to the Author:

The manuscript reports on a large data collection endeavour to understand intended acceptance of a COVID-19 vaccine. The sample size, multi-country focus, and broad range of demographic, social and psychological correlates that were included, are all strengths of the study. I have a few points of recommendation which mainly focus on the inclusion of further information at different points within the manuscript.

1. The title - I think the title may be a little misleading in the months to come, I would encourage the authors to specify that they are talking about intended vaccine acceptance, rather than actual acceptance. Soon papers will appear that examine actual uptake and think it will be important that there is a clear distinction between those that looked at intended acceptance and those that looked at uptake. I would also encourage the authors to reconsider the use of the word predictors here, and elsewhere in the manuscript. The word correlates may be more appropriate given the cross-sectional nature of the data (the word correlates is used in the abstract for example). Elsewhere talking about the factors associated with uptake, rather than predicting uptake would be more appropriate.

Thank you for raising this point, we have amended the title and text accordingly, and now avoid causal language, using the terms correlate or independent variable, or referring to associations when discussing the methods and results.

2. Abstract – in the results it would be useful to report the range of % uptake figures across countries (i.e. 62.6% - 88.1%).

We now include this in the abstract.

3. The introduction provides a good overview of the literature in this area, both in terms of likely COVID-19 vaccine hesitancy and vaccine hesitancy in general. The literature is fast moving in this area, but I would not suggest any changes to the intro.

Thank you for the positive feedback.

4. In the method it would be useful to add some additional contextual information. For example, an indication of the timeline of the COVID-19 vaccine development process during the data collection period of the surveys would useful.

We now include a very brief overview of the timing of vaccine developments (announcements of candidates, phase III trials; P9, L1):

The information about potential vaccines also changed over the data collection period. In February, 2020, the first major vaccine candidates, the Moderna and Oxford AstraZeneca vaccines, were announced [48,49]. In mid-2020 the launches of Phase III trials for several vaccines were announced: Moderna and Pfizer BioNTech in July [50], and AstraZeneca in August [51]. Results of Phase III clinical trials and estimates of efficacy were not announced during the data collection period (ending

in October, 2020). No vaccines were approved for use by local regulators at the time(s) the surveys were conducted in each country.

5. In addition, an indication of the infection pressure and the restrictions in place in each country at the time of data collection would also be helpful, perhaps as an additional column in Table 1. The heterogeneity in the factors associated with intended uptake may reflect the heterogeneity in the contexts within which the surveys were carried out, and so having some indication of the context would be useful and aid interpretation.

This is an excellent suggestion. We have now added four additional columns to table 1 for each country and survey time point: Days since first reported case, Total cumulative COVID-19 deaths, cases reported in the week prior to the survey, and the Oxford Stringency Index (a broad index of the level of government restrictions in place, based on a number of indicators). The methods now include the following text (P8, L15):

It is important to note that the surveys were conducted at various timepoints as the pandemic unfolded in each country. Table 1 reports the total number of COVID-19 deaths for each country at each survey timepoint, and the number of reported cases in the week prior to the survey (with the caveat that reporting practices vary between countries). We also provide the Stringency Index measure generated by the COVID-19 Government Response Tracker [45], which is a 0-100 index based on various restrictions put in place by governments to control the pandemic (e.g. closing schools, 'shelter in place' requirements). External data were sourced from the COVID-19 Government Response Tracker [45] and Ali et al. [46]

6. Clearly there is a huge amount of data here, and the figures that have been developed are excellent. The inclusion of the survey items within the supplementary files is also very welcome.

Thank you for the positive feedback. We experimented with a range of methods to present the findings in an easily interpretable format and are glad that the final format was found to be useful.

7. I appreciate that the authors have included a really broad range of measures but there are a few salient omissions in terms of socio-demographic factors which I think should be acknowledged as limitations. E.g. Evidence would suggest that ethnicity and socio-economic status are important factors to consider, and I was surprised that these were not included here. Presence of an underlying health condition, (or 'shielding' in UK terms) would also have been useful to assess.

An excellent point which we now address in the limitations (P19, L21). It was not possible to collect ethnicity information due to ethical restrictions in several of the countries surveyed:

We must note that our surveys did not examine several sociodemographic factors that could explain additional variance in vaccination intentions; ethnic minority status, socio-economic status and underlying health conditions have all been shown to be associated with COVID-19 vaccine hesitancy in some contexts [67,68]. Future research should examine how these factors relate to vaccine confidence and intentions as vaccine campaigns progress.

8. Linked to this point, the authors do point to several groups where vaccine hesitancy may be more of an issue. In terms of the practical considerations for public health messaging, the importance of targeted messaging could be considered more in the discussion. In general, the practical considerations section is relatively brief, and I would welcome a more in-depth consideration of this. How else can the results here inform the targets of, or content of messages designed to increase uptake, or where should future research focus to help us get to this stage?

We are more than willing to expand on this in the discussion, however with revisions we are severely limited in word count and have left only original comments on this. We are happy to take advice from the editor on this, and add more if there is some leniency on length.

Reviewer: 2

Dr. Katie Attwell, University of Western Australia School of Social Sciences

Comments to the Author:

This well-written and engaging article reports the findings of surveys in multiple countries regarding COVID-19 vaccine uptake intentions and the factors that contribute. It dives deeper in US and UK where there were multiple surveys over time with different respondents.

Thank you for the positive comments.

The article will need updating to reflect the current status of vaccine rollout since it was submitted. The authors should also update their literature review to reflect new studies that have come out since they submitted, including those that having similar findings, so as better to contextualise the contribution of the piece when it comes out.

We now include mention of the current status of vaccine roll out, citing a recent summary of uptake across countries (P4 L6):

While recent announcements of effective vaccines[5,6] and their rollout to certain demographics in some countries is promising[7], the wider impact of vaccines on preventing the spread of disease is also dependent on the broad uptake within a given population.

We also made minor adjustments to wording throughout the manuscript in this regard.

We have also cited a number of new studies which examine the role of trust in vaccine hesitancy, finding similar results (Jennings et al., 2021; Petravić et al., 2021; Thaker, 2021).

The authors should also reconsider / update their 'herd immunity' framing in the introduction as the vaccines' capacity to prevent transmission is yet to be determined in rollout scenarios and may be less than thought at the time of submission.

We now explicitly note the uncertainty around prevention of transmission and the changing state of information, in the introduction (P4, L14):

It must be noted that, while there is evidence that currently available vaccines can reduce SARS-CoV-2 infections [10], there is only limited preliminary evidence that vaccination can reduce transmission of the virus at the time of writing [11]. Thus, the net impact of vaccination campaigns on the spread of the virus remains uncertain until more research is conducted [12].

And discussion (P18, L14):

However, such strategies depend on vaccines preventing transmission of the virus, rather than just symptoms. There is now preliminary evidence that this is case for the Moderna and Pfizer BioNTech vaccines [11], but further studies are required to confirm these findings.

This is a minor formatting point but Table 1 needs to be addressed so that it is easier to read without words split across lines.

Apologies, this was a formatting issue arising from the automatic PDF conversion process. In response to Reviewer 1's requested revisions, Table 1 has now been expanded and is presented in landscape in the manuscript.

Can the authors explain why they used more than one survey company in UK and US, and why they think these companies found different results even when collecting data at the same time? Might it have made more sense to combine datasets when collected on same date (as UK Figure 3 indicates?). What are the strengths and limitations in their decisions surrounding these dual surveys?

Thank you for raising this point. As can be seen in the surveys (Table 1) our initial UK and US samples (March, 2020) were collected via Prolific. However, following from work with Respondi in other countries, we decided to also collect UK samples through Respondi. However, as our Prolific data represented the earliest stage of the pandemic, we opted to continue sampling this population as well. This is now stated in the methods section (P7, L20):

Our initial US and UK samples were recruited via Prolific (prolific.ac). Although some later samples from these countries were recruited via Respondi, we continued to also recruit Prolific samples to allow comparisons with our earliest data points in the pandemic. As we did not have matching Prolific and Respondi samples at all time points, and as results differed slightly between these providers, we report these samples separately for transparency.

We believe that collecting data through two separate providers is a strength, inasmuch that it provides insight into the consistency and variability in results across two very widely used participant recruitment platforms. We believe that reporting these samples separately is a more transparent approach as they do differ in terms of vaccine acceptance, and as at some time points we only report data collected via one company.

Analysing differences between panel providers was not an aim of the current study. However, we would note that recruitment procedures differ between the two platforms and this may lead to differences between the panels, particularly in terms of socio-economic factors. We would be happy to add more on this in the discussion but have omitted this for now, due to length.

The findings about social trust and trust in government not being influential on vaccine acceptance are interesting, and the lack of significance of prosocial measures. The authors need to consider the latter point in light of the 'herd immunity' point above – COVID-19 vaccines will still inevitably have societal benefits, but what happens if we can't talk about them generating herd immunity?

This is a good point. In the discussion we now acknowledge the uncertainty regarding prevention of transmission (see response to the reviewer's earlier point; P17, L12), noting that there is now some evidence that Moderna and Pfizer vaccines prevent transmission. We would like to offer more speculation on the implications if this does not hold up, but given the scope of the paper and limitation on space, we have not done so at this stage.

The study's findings are well-contextualised and the discussion connects well to the vaccine uptake research and make salient recommendations for future research.

Thank you for the positive comments.

Reviewer: 3

Dr. Wei Wang, Food and Drug Administration

Comments to the Author:

In this manuscript, the authors presented a more comprehensive international analysis of the role of key social, political, and psychological predictors of COVID-19 vaccine acceptance across 12 countries, with multiple national surveys in some countries. From the statistical point of view, the analysis methods applied in this manuscript is generally acceptable and the results presentation is clear. I still have the following minor comments for the authors' consideration,

(1) It is recommended that the authors include separate sections in the METHODS to describe the analysis methods used in the whole manuscript which will make the METHODS section more complete and the readers understand the statistical analysis methods used in the whole manuscript easily. Specifically, the authors need to elaborate the Table S2 metric invariance analysis a little bit, for example, what does CFI mean?

We have now removed the metric invariance analyses due to length considerations. While these added some strength to our findings, such statistics are not considered standard to report and their inclusion may detract from the overall thrust of the paper.

(2) The authors compared the vaccine acceptance among respondents in Mexico between March and May, in US, between May and September. It is not clear whether the respondents are the same cohort or different cohorts. If they are the same cohorts, the two sample proportion test used is not correct. The authors may need consider data analysis methods to handle this correlated binary outcome, like GEE, or Generalized linear mixed models (or GLMMs). Even they are from different cohorts, it is also recommended to use Chi-Squared or Fisher's Exact Test to compare independent proportions. Although the z test used by the authors are not wrong, it does need additional normal approximation. Please clarify correct the manuscript accordingly.

Thank you for the opportunity to clarify, we state in the methods that surveys in the same countries were independent samples (i.e. not the same participants; P8, L2)):

Participants who had previously completed a survey were prevented from completing further surveys, so all our samples represent different individuals.

Following the Reviewer's suggestion we now report Chi squared tests of independence for the comparisons (P12, L3).

(3) The authors need to clearly specify using multiple logistic regression model instead of logistic regression model in the main analysis of the manuscript to avoid confusion.

Duly noted and corrected. We now refer to multivariate logistic regression to resolve any ambiguity. Thank you.

Reviewer: 4

Dr. Rania Wasfi, Centre de recherche du CHUM, Universite de Montreal Ecole de Sante Publique

Comments to the Author:

Thank you for conducting this analysis. It is an important topic. Please find attached a detailed review of your paper, with comments and suggestions.

Peer Review:

Predictors of COVID-19 vaccine acceptance across time and countries: Results from a series of cross-sectional surveys

Thank you for conducting this research, it is an important topic that need to be addressed to help in understanding strategies governments can use to increase vaccination acceptance among people.

Thank you for the positive comments.

The objective of the paper is clear, and important, however the way the methods are used and results presented are confusing and I think if modified will make an important contribution. In the current form now it is difficult to makes clear inference, without many unadjusted confounding.

Please see below specific comments and suggestions:

The decision of stratifying the logistic regressions models by country is not explained and justified. Can you please elaborate on that.

Our aim was to draw on the data collected to provide a broad examination of the consistency of vaccine acceptance correlates between countries. A single model would mask between-country differences and we do not have enough countries in the dataset to reliably estimate country level random effects (Bryan & Jenkins, 2016). However, we can, and now do, adjust for these in estimating individual-level effects (see next point).

We outline this in a new Analysis section added to the Methods (P11, L 4):

To allow for descriptive comparisons between countries and across time, we report these model results separately for each country, time point, and (in the UK) panel provider.

In Table S6: A pooled model is constructed, was the country and time of data collection adjusted for in the pooled model. I cannot see that in the Tables. If not, why is this the case?

We included the pooled sample as broad indicator of the overall effects. Given the relatively low number of groups (Bryan & Jenkins, 2016), we are hesitant to present a multilevel model as key result in the manuscript. However, to address the Reviewer's further comment on this below, we now report a multilevel model controlling for country and time (month of the year) random effects as suggested. We report this in the supplementary materials (Table S7), and note in the main text (P14, L20) that the fixed effects are essentially identical to those reported in the simpler model in the main text.

Why is UK and USA samples excluded then analysed separately? There is no clear justification for that.

The UK and USA data was included in the 'ALL' model in Table S6 (now Table S5). We apologise, we initially split a very wide table into two tables simply for space. We have revised this to be only one wide table, split across two pages, with labels and captions revised accordingly.

In Table S7. Why did not pool the data for the UK (for all months, and adjust for the month in one pooled the authors regression models. Again, it is very hard to interpret the results in this form, and understand the effect of each predictor independent from time. If authors think the predictor change over time, why did not they try an interaction variable with time?

This is a fair point but we believe that presenting the individual models for each survey is more informative and transparent. Given the very dynamic nature of the pandemic, it is reasonable to believe that certain beliefs or attitudes may have become more or less relevant to vaccine decision-making as the environment changed. As we note in the manuscript, in the UK, political ideology only became a significant correlate of vaccine acceptance in later surveys. While it is tempting to model this as a linear interaction between ideology and time (and indeed we tried), we note that the pattern of effects seen in the individual survey models suggests a nonlinear trend (with ideology becoming less and then more strongly associated with acceptance over time in the UK; see Fig 3). While certainly interesting, we believe more complex non-linear modelling of associations and their interactions with time to be beyond the scope and aims of the current paper.

In Table S8. The time indicates: “Full logistic regression results from model predicting vaccine recommendation to vulnerable other”- what is vulnerable other? Please elaborate.

This was an additional outcome variable we included in the survey but did not report in the main text due to space (see P14, L18). In addition to their own vaccination choice, participants were asked if they “Would recommend getting the vaccine to vulnerable friends and family” (yes/no). We have revised the table caption to make this more explicit.

In the logistic regression models, why did not the authors adjust for the company of recruitment, given that there were differences shown and presented by the authors in other Tables?

This is a very reasonable point and one we had already considered. Ultimately we decided that it would be more transparent to report these separately, and maintain a consistent model fitted to all surveys.

The authors presented a Table of missing data, but did not discuss how this missing data influence the generalizability of the results. I understand the sample was not a random sample but was matched by age and gender. Did this differ with missing values?

We now acknowledge this directly in the footnote on missing data (P11):

In the remaining samples the average proportion of missing responses for vaccine intention and recommendation items was 1% (see supplementary Tables S2 and S33 for description of missing data and the age and gender distribution of those participants who answered the vaccine acceptance item). We acknowledge that in some cases estimates of vaccine acceptance may not be based on samples exactly matched to a country’s population age and gender distribution due to this missing data, but note that age and gender are controlled for in the models below.

For transparency, we now report the full age and gender distribution of participants in each sample who answered the vaccine acceptance item, but have not provided more detailed analyses as the main focus of the study is the correlates of vaccine acceptance (including age and gender).

It is hard in the current format of the logit models to make between countries comparison. Why did not the authors use a nested model, where cases are nested within countries and adjusted for the country of survey? I suggest authors to try using multi-level logits to separate the random effects, and minimize bias, and be able to compare between countries.

We have followed the Reviewer's suggestion and now report a multilevel model, incorporating additional survey-level variables in the supplementary materials (Table S7; incorporating suggestion below). The fixed effects are essentially the same as those produced by the simpler pooled model. Due to the low number of countries and time points we cannot report reliable random effects (Bryan & Jenkins, 2016), so include this model in the supplementary material, noting this caveat.

It could be useful to indicate when the first cases of Covid-19 started in each country or indicate the number of cases at the time the survey was administered to get some context of what this time (Month) means during the pandemic. Moreover, use a multi-level logistic regression with pooled data (from all countries) adjusting for Country of survey, Time-from first covid-19 cases in the country, # of cases, and month, in addition to other covariates in the models.

These are all good suggestions, in light of this comment and those of Reviewer 1, we have added additional columns to Table 1 in the manuscript, for each country and time point: number of deaths, cases reported in the week prior to the survey, and level of government restriction. We believe these offer a succinct insight into the state of the pandemic in each country at each time point.

Regarding multi-level models, see response to the previous comment.

Minor comments:

On page 9 line 8 : (example: I will probably get sick with the coronavirus/COVID-19; αs .71-.89). For : "αs .71-.89" Why is findings reported in Methods?

In our experience, scale reliability is usually reported as part of the methods, for brevity. Should editorial requirements dictate that these statistics be moved to the results, we are more than willing to do so, but would prefer to keep them in the methods alongside the scale descriptions.

On page 9 line 42. Please put an explanation and spell out the (CFI) on line 42.

Thank you for noting this oversight. We have now removed the metric invariance analyses due to length considerations. While these added some strength to our findings, such statistics are not considered standard to report and their inclusion may detract from the overall thrust of the paper.

References

Bryan, M. L., & Jenkins, S. P. (2016). Multilevel Modelling of Country Effects: A Cautionary Tale. European Sociological Review, 32(1), 3–22. https://doi.org/10.1093/esr/jcv059

Jennings, W., Stoker, G., Willis, H., Valgardsson, V., Gaskell, J., Devine, D., Mckay, L., & Mills, M. C. (2021). Lack of trust and social media echo chambers predict COVID-19 vaccine hesitancy. MedRxiv, 2021.01.26.21250246. https://doi.org/10.1101/2021.01.26.21250246

Petravić, L., Arh, R., Gabrovec, T., Jazbec, L., Rupčić, N., Starešinič, N., Zorman, L., Pretnar, A., Srakar, A., Zwitter, M., & Slavec, A. (2021). Factors Affecting Attitudes towards COVID-19 Vaccination: An Online Survey in Slovenia. Vaccines, 9(3), 247. https://doi.org/10.3390/vaccines9030247

Thaker, J. (2021). The Persistence of Vaccine Hesitancy: COVID-19 Vaccination Intention in New Zealand. Journal of Health Communication, 1–8. https://doi.org/10.1080/10810730.2021.1899346

VERSION 2 – REVIEW

REVIEWER	Williams, Lynn University of Strathclyde, School of Psychological Sciences & Health
REVIEW RETURNED	29-Apr-2021

GENERAL COMMENTS	Thank you for your detailed response to the considerations that were raised. I have no further suggestions for the manuscript.

REVIEWER	Wang, Wei Food and Drug Administration, Division of Biostatistics, Center for Devices and Radiological Health
REVIEW RETURNED	01-May-2021

GENERAL COMMENTS	All my concerns were addressed in the revised manuscript.

REVIEWER	Wasfi, Rania Centre de recherche du CHUM
REVIEW RETURNED	29-Jun-2021

GENERAL COMMENTS	Thank you for all the revisions and for addressing my comments/ suggestions, and the other reviewers suggestions in the first round of review. I find the paper much clear with the revisions and give useful insights on the vaccination acceptance. Please find few points I would like the authors to clarify and few suggestions below : I suggest in the multi level model, that data be nested within months, then put the months as a variable with a reference group (this will show the trend over time) , as well as the country, as another variable. That might solve the random effects problem. Although the variables days since first day in Covid and total confirmed cases in the nested model doesnot show statistical significance, I would suggest also adjusting for them in the separate models you are presenting (model per study). Have you tried to adjust for total death? Finally, I think it would be interesting to also adjust for the Stringency index, to understand whether measures put in place to decrease COVID-19 transmission, are correlated with vaccine acceptance. For example are countries with strict lock down measures correlated with more vaccine acceptance? Thank you

VERSION 2 – AUTHOR RESPONSE

Reviewer 4

R4: Thank you for all the revisions and for addressing my comments/ suggestions, and the other reviewers suggestions in the first round of review. I find the paper much clear with the revisions and give useful insights on the vaccination acceptance.

Response: Thank you for the positive comments.

R4: Please find few points I would like the authors to clarify and few suggestions below:

I suggest in the multi-level model, that data be nested within months, then put the months as a variable with a reference group (this will show the trend over time) , as well as the country, as another variable. That might solve the random effects problem.

Response: This is an excellent suggestion and would certainly be worthwhile in a dataset with more countries and timepoints. We do now nest the data in months and countries separately as suggested, rather than nesting months within countries. However, this does not resolve the issue arising from the very small number of groups in the multi-level model (now 10 countries and six months). We cannot reliably estimate the effect of, say, month, based on so few datapoints—particularly where those months are unevenly distributed across different countries. The same applies for country/month-level predictors (e.g. confirmed cases). While controlling for these in a multi-level model can improve estimates of individual-level parameters, we cannot reliably make inferences regarding country/month-level parameters.

We refer the reviewer to Bryan and Jenkins' (2016) simulation study which outlines the problems with attempting to estimate country-level parameters with a small number of countries (or any other higher level grouping).

With large sample sizes of individuals within each country but only a small number of countries, analysts can reliably estimate individual-level effects but estimates of parameters summarizing country effects are likely to be unreliable. Multilevel modelling methods are no panacea. (p3).

Additionally:

When there are few countries in a multi-country dataset, there is little information with which to estimate country effects, whether these effects refer to the fixed parameters on country-level predictors or the variances of random country intercepts. Multilevel model users need to be cautious in the claims they make about country effects of either type. (p19)

Bryan and Jenkins go on to note that issues arising from a small number of countries are more pronounced in multi-level logit models (such as ours), as compared with linear models, and they suggest a minimum of 30 countries is required to obtain reliable estimates.

Our Monte Carlo simulations suggest that users require 25 countries for linear models and 30 countries for logit models at the very minimum, and most likely more for models with a specification other than a relatively basic one. Otherwise, estimates of country-level fixed parameters are likely to be estimated imprecisely and this will not be adequately reflected in test statistics reported by commonly used software: users will conclude too often that a country effect exists when it does not. Country random variances will be biased downwards and have CIs that are too narrow. The only estimates that are unaffected by having a small number of countries are the fixed parameters on individual-level predictors (the number of individuals per country is typically large): provided there is not also a random component attached to the slope, these parameters are estimated without bias and with the correct SEs… (p20)

It is for this reason that we are reluctant to include the multi-level model in the main text, but include it as a supplementary analysis.

R4: Although the variables days since first day in Covid and total confirmed cases in the nested model does not show statistical significance, I would suggest also adjusting for them in the separate models you are presenting (model per study). Have you tried to adjust for total death?

Response: The first point is a worthwhile suggestion. But again the structure of the data collected limits the usefulness of including such variables in the models fitted to each separate sample. In short: there simply isn't enough variability in these predictors to draw meaningful inferences at the level of each survey.

Most surveys were fielded over a few days, with the bulk of participants completing the survey in the first day or two. Thus, there would be minimal differences between participants, in say, total cases or deaths reported in their country at the time of survey completion.

For example, we could include 'total deaths reported at the time of survey completion' for each individual participant, based on the date they completed the survey. However, examining the data for Germany collected in March 2020 (as an example), we find that 86% of participants completed the survey on the 23rd of March, when 123 deaths were officially recorded. The remainder of the sample completed the survey the following day (24th March) where the number of deaths was recorded as 157.

Notwithstanding the extent to which participants were actually aware of these daily changes in reported deaths, there is just too little variation in these numbers to include them as continuous covariates in the models for each sample.

Regarding the second point, we do now include total confirmed deaths as a country/month-level predictor in the multilevel model. The effect of total deaths is significant in the model. However, for the reasons outlined in Bryan and Jenkins (2016) we don't report this directly as a result and include a cautionary note in the table warning the reader not to take these country/month-level parameter estimates at face value.

R4: Finally, I think it would be interesting to also adjust for the Stringency index, to understand whether measures put in place to decrease COVID-19 transmission, are correlated with vaccine acceptance. For example are countries with strict lock down measures correlated with more vaccine acceptance?

Response: Another excellent suggestion, but again we are limited by our dataset. In the multilevel model, we do now adjust for the level of intervention in each country at each timepoint by including the stringency index, but cannot draw any conclusions about its effect for the reasons outlined above.

In light of the changes made in response to the reviewer's suggestions, the multilevel model (Table S7) and main text have now been amended (Page 13, Line 12):

We fitted an additional multi-level model to the pooled data, adjusting for country, month, days since first case, level of government intervention, total reported deaths, and number of cases reported in each country at each time point (Table S7).

As a final point, we reiterate that the effects of individual-level predictors in the updated model are still essentially unchanged from those in the pooled model reported in the main text.

We thank the reviewer again for these excellent suggestions and we regret that our dataset was not comprehensive enough to fully incorporate them in this instance.

References

Bryan, M. L., & Jenkins, S. P. (2016). Multilevel Modelling of Country Effects: A Cautionary Tale. *European Sociological Review*, *32*(1), 3–22. https://doi.org/10.1093/esr/jcv059

Freeman, D., Loe, B. S., Yu, L. M., Freeman, J., Chadwick, A., Vaccari, C., Shanyinde, M., Harris, V.,

Waite, F., Rosebrock, L., Petit, A., Vanderslott, S., Lewandowsky, S., Larkin, M., Innocenti, S., Pollard, A. J., McShane, H., & Lambe, S. (2021). Effects of different types of written vaccination information on COVID-19 vaccine hesitancy in the UK (OCEANS-III): a single-blind, parallel-group, randomised controlled trial. *The Lancet Public Health*, *6*(6), e416–e427. https://doi.org/10.1016/S2468-2667(21)00096-7

Lazarus, J. V., Ratzan, S. C., Palayew, A., Gostin, L. O., Larson, H. J., Rabin, K., Kimball, S., & El-Mohandes, A. (2020). A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine*. https://doi.org/10.1038/s41591-020-1124-9

Lindholt, M. F., Jørgensen, F., Bor, A., & Petersen, M. B. (2021). Public acceptance of COVID-19 vaccines: Cross-national evidence on levels and individual-level predictors using observational data. *BMJ Open*, *11*(6), 48172. https://doi.org/10.1136/bmjopen-2020-048172

Lucid. (2021). *Sample Sourcing FAQs*. https://support.lucidhq.com/s/article/Sample-Sourcing-FAQs

Parsons, S., Starling, B., Mullan-Jensen, C., Tham, S. G., Warner, K., & Wever, K. (2015). What the public knows and wants to know about medicines research and development: A survey of the general public in six European countries. *BMJ Open*, *5*(4), 6420. https://doi.org/10.1136/bmjopen-2014-006420

Smith, L. E., Potts, H. W. W., Amlôt, R., Fear, N. T., Michie, S., & Rubin, G. J. (2021). Adherence to the test, trace, and isolate system in the UK: Results from 37 nationally representative surveys. *The BMJ*, *372*. https://doi.org/10.1136/bmj.n608