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)	The differential impact of physical distancing strategies on social contacts relevant for the spread of SARS-CoV-2: Evidence from a cross-national online survey, March – April 2020
AUTHORS	Del Fava, Emanuele; Cimentada, Jorge; Perrotta, Daniela; Grow, André; Rampazzo, Francesco; Gil-Clavel, Sofia; Zagheni, Emilio

VERSION 1 – REVIEW

REVIEWER	Chaw, Liling University of Brunei, PAPRSB Institute of Health Sciences
REVIEW RETURNED	19-Apr-2021

| GENERAL COMMENTS | The following are my comments for the manuscript entitled "The differential impact of physical distancing strategies on social contacts relevant for the spread of SARS-CoV-2: Evidence from the Covid-19 Health Behaviour Survey". Overall, it is very well-written, with a detailed set of supporting information attached. However, I have some concerns which I hope to get feedback from the authors.

Major comments:
1. The method on the collection of behavioural factors were not sufficient, in my opinion. Although references were given, it will be easier for readers to get at least a summary of what was done to collect these data. One idea is to just provide the survey on this section in the supplementary.

2. Please provide the definitions used for washing their hands more often, wearing a face covering & avoiding social activities. Again, although references were given, it will be easier for readers to get at least a summary in this paper.

3. Although data was collected for 4 settings (home, school, work & general community), only the latter 2 (work & general community) were reported in the results. Is there a reason for this? As the authors have mentioned in the supplementary, the distribution of household contacts was quite consistent across countries (Figure S5). Despite this, there are still variation (for example, between European countries & US). Also, household setting plays an important role in COVID-19 transmission, hence I think should be mentioned in the main results.

4. Relating to point 3 above, does the "Overall" setting in Table 1 indicate all 4 settings or just work & general community? |
|---|---|

5. What are the implications of your study results, in terms of SARS-CoV2 transmission?

Minor comments:
1. Please re-order your reference numbering, as there is an inconsistency between reference numbers 14 & 15 (page 5 & 6 of manuscript).

2. Please mention the software used for these analyses

REVIEWER	Kelly, Dervla University of Limerick, Graduate Entry Medical School
REVIEW RETURNED	14-May-2021

| GENERAL COMMENTS | Thank you for the opportunity to review this paper. This study presents the results of an online survey of health behaviours and allows a cross country comparison of number of social contacts. While a large amount of data was collected, the study design and data collection focused on health behaviours and it is not clear that the outcome 'r number' is sufficient to comment on changes in virus transmission, or answer the research question.

Abstract: Reference to "highly granular data" seems out of place at the moment given the focus on country level. Suggest review abstract.
Line 29: link between social contacts and reduction in r seems overreaching

Introduction
The rationale of quantifying the relative impact of measures on a hypothetical r value is not currently clear. This could be included in the introduction to set the reader up for this. Do social contacts predict r? What are the previous studies reporting this

The objective of the study is not clear. Was it to measure uptake/compliance with social distancing measures or the relationship between policies and changes in behaviours? The outcome data on the spread of covid-19 is weak so perhaps best not to focus on this (line 81).

Methods
Regarding r number, did the authors consider incorporating weekly case numbers from each area? Its not clear how the assumptions used to calculate r in this study mean that the r value reflects transmission and allows statements around transmission. Did you consider including population density?
Please include strobe checklist or similar in appendix

Discussion:
Further consideration of confounders or unknowns is needed throughout the discussion. Generalisability with a focus on country level is a limitation of the study. Is your model applicable to rural and urban areas alike? |

	Line 275: essential workers would be a more likely explanation for having high counts. Was information on profession collected?

VERSION 1 – AUTHOR RESPONSE

<u>**Reviewer: 1**</u>
Dr. Liling Chaw, University of Brunei

<u>Comments to the Author:</u>
<u>The following are my comments for the manuscript entitled "The differential impact of physical distancing strategies on social contacts relevant for the spread of SARS-CoV-2: Evidence from the Covid-19 Health Behaviour Survey". Overall, it is very well-written, with a detailed set of supporting information attached. However, I have some concerns which I hope to get feedback from the authors.</u>
Dear Dr. Chaw, thank you very much for the time you took to review our manuscript. Please, find below our replies to your comments.

<u>Major comments:</u>
<u>1. The method on the collection of behavioural factors were not sufficient, in my opinion. Although references were given, it will be easier for readers to get at least a summary of what was done to collect these data. One idea is to just provide the survey on this section in the supplementary.</u>
Following the reviewer's suggestion, we moved part of the material from the supplementary file to the main part of the manuscript, providing more details on the recruitment of respondents using targeted Facebook ads, the construction of the post-stratification weights, and the definition of social contacts. The corresponding section now reads as follows:

"We designed the CHBS as a cross-country, cross-sectional, and voluntary opt-in online survey to collect key information on people's health and behaviour in eight high-income countries – Belgium, France, Germany, Italy, the Netherlands, Spain, the United Kingdom (UK), and the United States (US). The first wave of data collection ran from March 13 to April 12, 2020, a key period during which the global pandemic was well underway, even though at different stages across countries (Perrotta et al. 2021). To quickly reach large numbers of survey participants across several countries, recruitment occurred via targeted advertisements implemented through the Facebook Ads Manager, a tool usually employed by advertisers to create marketing campaigns for their products. The ads were stratified by sex, age group, and subnational level, in order to ensure that a minimum number of respondents could be reached in each stratum (Pötzschke e Braun 2016; B. Zhang et al. 2018). Post-stratification weights by sex, age group, and subnational level were used in all analyses. These were obtained by dividing the true population proportion in each stratum (based on nationally representative census data (Eurostat 2019; US Census Bureau 2019)) by the sample proportion from the same stratum. The effect of applying these weights to the sample is shown in the Supplementary Tables S1-S8.
       …
We asked respondents to report the number of social contacts (defined as any interaction involving either physical contact or a conversation of three or more words in the physical presence of another person) on the day prior to the survey in four different settings, i.e., at home, at school, at work, and in the general community (e.g., during commuting or leisure activities) (Supplementary Fig. S1)."

<u>2. Please provide the definitions used for washing their hands more often, wearing a face covering & avoiding social activities. Again, although references were given, it will be easier for readers to get at least a summary in this paper.</u>

Following this suggestion, we now provide the wording of the respective questionnaire items (indicated by question marks).

The corresponding section now reads as follows:

"We also asked participants to report, among others, their perceived level of threat (Q13: "What level of threat do you think the coronavirus poses to… you personally? … your family?") and their adoption of any protective behaviours that might have an impact on disease transmissibility (Teslya et al. 2020) (Q18: "Which of the following actions, if any, have you already taken to protect yourself from the coronavirus?... (i) washed hands more often?... (ii) worn a face mask?... (iii) avoided social activities (e.g., meeting friends)?")."

3. Although data was collected for 4 settings (home, school, work & general community), only the latter 2 (work & general community) were reported in the results. Is there a reason for this? As the authors have mentioned in the supplementary, the distribution of household contacts was quite consistent across countries (Figure S5). Despite this, there are still variation (for example, between European countries & US). Also, household setting plays an important role in COVID-19 transmission, hence I think should be mentioned in the main results.

Thank you for the comment. At the beginning of the observation period, school contacts were negligible given that all secondary schools and universities had already suspended in-person activities. Furthermore, home contacts remained rather stable during the survey period (Supplementary Figs. S3-S4). We, therefore, decided to report only contacts at work and in the general community (next to the overall number of contacts), given that these contacts showed more variability over time and were more directly affected by the implemented physical distancing guidelines than home contacts.

However, as remarked by the reviewer, we recognize the importance of home contacts for the spread of COVID-19 and hence the necessity of including information on home contact patterns in the mathematical models for the spread of disease infection. Hence, we reran the analysis and included the predicted number of home contacts (with SE) in Table 1 in the manuscript and in Supplementary Table S17.

Given the relatively lower variability in the home contacts, we found it problematic, due to convergence issues, to use the same set of variables that we used for modelling the overall contacts, as well as the contacts at work and in the general community. Hence, we changed the negative binomial regression model that we originally defined by removing the interactions between the calendar week and the age group and between the calendar week and the behavioural factors. In this way, we were able to use the same model for all dependent variables and to interpret in a more straightforward way the determinants of the reported number of contacts. Although this modification of the model led to small changes in the average number of contacts, the estimates of the reproduction number were not much affected.

Considering the last comment regarding the importance of the household contacts in the transmission of SARS-CoV-2, we added a paragraph in the Results section describing how the household became the main source of contacts during the survey period:

"This variation [in contacts] was mostly due to a decrease of contact numbers in the general community and, to a lower extent, at work. Indeed, while the former setting was the main contributor to the overall number of daily contacts in the pre-COVID-19 period, at the end of the study period (week 2020-15) and under the 90% quantile threshold, home contacts had become the largest component, representing up to 50% (in Germany and the Netherlands) and around 70% (in Italy, Spain, and the UK) of all the contacts (Supplementary Fig. S3); this result was slightly attenuated under the ≤29 threshold, as in Germany and the Netherlands, both the work and the general community represented more important contributors than the household (Supplementary Fig. S4), likely due to the smaller average household size in these two countries (Supplementary Fig. S5)."

4. Relating to point 3 above, does the "Overall" setting in Table 1 indicate all 4 settings or just work & general community?

The "Overall" setting in Table 1 indicates all four settings of contacts. Note that contacts at school usually represent a tiny fraction of the contacts reported by a population aged 18 years or more, not only during the study period, but also in the baseline period. To make it clearer, we added the following sentence to the caption of both Table 1 and Supplementary Table S1:

> "The "Overall" category encompasses contacts reported in all four surveyed settings, e.g., home, school, work, and general community."

5. What are the implications of your study results, in terms of SARS-CoV2 transmission?
Thanks for this comment. In the revised version of the manuscript, we clarified the objectives of the study, both in the abstract and the in the introduction. These objectives included (i) the assessment of the changes in social contact patterns across countries and over time, and (ii) the investigation of the implications of such changes for the spread of SARS-CoV-2. The latter objective was achieved by estimating the net reproduction number , an index that depends on the social contact patterns of individuals, as well as on the infection transmissibility and the duration of the infectious period. We found that the observed reduction in social contacts, compared to the numbers observed prior to the pandemic, resulted in the weekly estimates of smaller than 1 (95% confidence interval included) in all surveyed countries, except for Germany. In terms of infection transmission, keeping fixed factors such as the infectionransmissibility, the duration of the infectious period, and the initial estimate of the basic reproduction number , which we assumed to be consistent with the pre-pandemic contact patterns, we showed how the observed reduction in contact numbers was compatible with a reduced infection transmission, especially in a phase in which still few people used a protective face mask when having contacts outside the home.

To make this point clearer in the Discussion, we amended the part where we discuss the results of the epidemiological model:

> "Third, in almost all countries, we found that the reduction in contact numbers, while keeping fixed other epidemiological factors, resulted in a reduction of the net reproduction number below one, and was therefore associated with a lower infection transmission, especially in a period in which the adoption rate of a protective face mask outside the home was still low, albeit gradually increasing."

Minor comments:
1. Please re-order your reference numbering, as there is an inconsistency between reference numbers 14 & 15 (page 5 & 6 of manuscript).
Thanks. All references were checked, updated (converting preprints to published articles when that occurred), and re-ordered.

2. Please mention the software used for these analyses.
Thanks. We added a sentence at the end of the Methods>Data analysis section (created by merging the former sections on the statistical analysis and the epidemiological approach), describing which software was used for the analysis:

> "All data preparation and visualization were performed in R software [29], while the negative binomial regression models were fitted in Stata 16."

Reviewer: 2
Dr. Dervla Kelly, University of Limerick

Comments to the Author:
Thank you for the opportunity to review this paper. This study presents the results of an online survey of health behaviours and allows a cross country comparison of number of social contacts. While a large amount of data was collected, the study design and data collection focused on health behaviours, and it is not clear that the outcome 'r number' is sufficient to comment on changes in virus transmission or answer the research question.
Dear Dr. Kelly, thank you very much for the time you took to review our manuscript.

The main idea of this paper is to see how the introduction of the physical distancing guidelines in each surveyed country influenced the number and the distribution of social contacts, and, in turn, infection transmission. For this purpose, we chose the net reproduction number and its calculation method to show the impact of the reduction in contact numbers on infection transmission compared to what we might have expected, had contacts not changed from the patterns observed prior to COVID-19.

We defined the impact of the physical distancing guidelines on contact patterns as 'differential' to communicate the idea of cross-country heterogeneity in the observed contact patterns and their association with several demographic and behavioural factors.

Hopefully, our amendments in response to your comments have made the implications of our findings clearer for readers.

Abstract
Reference to "highly granular data" seems out of place at the moment given the focus on country level. Suggest review abstract.
Thank you for this comment. In the main part of the manuscript, we indeed focus on the country level, but we think that the data still is "high granular" not only in reference to the spatial resolution, but also to the temporal level of granularity, as we were able to obtain daily data on social contact patterns and health behaviours from a sample of individuals stratified by sex, age group, and region of residence in each country.

Yet, to address your comment, we made two changes in the manuscript. First, regarding the country component of your comment, we complemented the information in Table 1 (for the 90% quantile threshold) and Supplementary Table S17 (for the ≤29 threshold) with the information in the Supplementary Data File, which contains the average number of social contacts stratified by threshold, country, region of residence, setting of contact, sex, and age group.

Second, to make the objectives of the study clearer, we removed the expression "highly granular" from the abstract and rewrote the Objectives section in the following way:
> "We investigate changes in social contact patterns following the gradual introduction of non-pharmaceutical interventions in the early phase of the pandemic and their implications for infection transmission."

Line 29: link between social contacts and reduction in r seems overreaching
After clarifying the relationship between social contacts and the net reproduction number in the methodological section, we only slightly changed the sentence by adding the interpretation of the values of equal or below one:
> "Such reductions, which were uniform across age groups, were compatible with equal or smaller than one in all countries, except Germany, indicating lower levels of infection transmission, especially in a period of gradual increase in the adoption rate of the face mask outside the home."

Introduction
The rationale of quantifying the relative impact of measures on a hypothetical r value is not currently clear. This could be included in the introduction to set the reader up for this. Do social contacts predict r? What are the previous studies reporting this?
Thanks for this comment. We rephrased the last paragraph of the Introduction to clarify that our study (i) is consistent with the literature on the relationship between social contacts and the reproduction number for close-contact infectious diseases, as well as with other surveys conducted during the same period in other countries (and here cited) and (ii) it seeks to quantify the impact of changes in social contact patterns on infection transmission, measured by changes in the net reproduction number. The paragraph now reads as follows:
> "Following the literature on the relationship between social contacts and their impact on the reproduction number for close-contact infectious diseases (Brooks-Pollock et al. 2021; Melegaro et al. 2011; Wallinga, Teunis, e Kretzschmar 2006), and consistently with other social contact surveys conducted in Asia (J. Zhang et al. 2020), Europe (Backer et al. 2021; Coletti et al. 2020;

Jarvis et al. 2020; Latsuzbaia et al. 2020; Sypsa et al. 2021), Africa (Quaife et al. 2020), and North America (Dorélien et al. 2020; Feehan e Mahmud 2021) to assess changes in contact patterns and their impact on infection transmission during the early phase of the pandemic, we quantified participants' social contacts, adoption of protective behaviour, and perceived threat, and estimated the relative impact of changes in social contact numbers on the net reproduction rate , taking also the effect of other behavioural changes into account."

The objective of the study is not clear. Was it to measure uptake/compliance with social distancing measures or the relationship between policies and changes in behaviours? The outcome data on the spread of covid-19 is weak so perhaps best not to focus on this (line 81).

The objectives of the study are twofold. First, we seek to quantify social contact patterns in the early phase of the pandemic, when the first physical distancing guidelines were gradually introduced, and describe how they changed with respect to patterns observed prior to the COVID-19 pandemic. This is achieved through a survey that targeted several countries at the same time and that collected daily data on social contacts, as well as on heath behaviour and threat perception.

Second, we seek to understand the consequences of changes in contact patterns on disease transmission by calculating an index (the net reproduction number,) that directly depends on social contact patterns. To make these goals more explicit to readers, we modified the Introduction in the following way:

"In this study, we leverage new opportunities for data collection, enabled by the digitalization of our lives, to assess changes in social contact patterns across countries and over time, and to examine their implications for the spread of SARS-CoV-2 by estimating the net reproduction number, an index that directly depends on social contact patterns."

Methods
Regarding r number, did the authors consider incorporating weekly case numbers from each area? It's not clear how the assumptions used to calculate r in this study mean that the r value reflects transmission and allows statements around transmission.

Thank you for this comment. We assume that the advice of incorporating weekly numbers of new cases refers to the method based on data such as the incidence of new cases and the generation time, i.e., the time interval between the infection time of the infector and the individuals that he or she infected. In the current epidemiological literature, there are, however, several methods used to estimate the reproduction number. This is especially the case that has emerged in the context of the COVID-19. Indeed, a report published by the Royal Society[1], which focuses in COVID-19 in the UK and that we also cite in the manuscript, describes alternative methods that are not based on incidence data to estimate the reproduction number.

The method that we followed assumes that there are several key factors for the derivation of , such as the type and frequency of contacts between individuals in the population (contact patterns), the probability that a contact between a susceptible and an infectious individual will result in infection (infection transmissibility), and the duration of the infectious period. All these parameters enter in the computation of the next generation matrix (NGM), whose dominant eigenvalue is the reproduction number (Anderson e May 1992; Farrington 2003; Goeyvaerts et al. 2010; Van Effelterre et al. 2009; Wallinga, Teunis, e Kretzschmar 2006). The NGM relates the numbers of newly infected individuals in various categories, e.g., sex or age group, in consecutive epidemiological generations (Diekmann, Heesterbeek, e Metz 1990).

To clarify our estimation method and the data it requires, we decided to move material from the supplementary file to the manuscript to show that our calculation of depends on the derivation of the NGM. Hence, the second paragraph of the section Methods>Data analysis (created by merging the former sections on the statistical analysis and the epidemiological approach) now looks like that:

"Next, we assessed the implications of the changes in social contact numbers for the net reproduction number (, i.e., the average number of new infections generated by an infectious person at time in a population, either partially or fully susceptible, considering the current interventions and the potential

spontaneous behavioural change in response to the risk of infection (Liu et al. 2018; The Royal Society 2020)), estimated as the dominant eigenvalue of the next generation matrix , which provides information on the numbers of newly infected individuals by age group at a given time (Diekmann, Heesterbeek, e Metz 1990). Under the "social contact hypothesis" (Goeyvaerts et al. 2010; Wallinga, Teunis, e Kretzschmar 2006), we have the relationship , where denotes the matrix containing the average number of contacts between age groups, is the infection transmissibility, and is the length of the infectious period. We used the proportional relationship between and , i.e., , to assess changes in due to changes in social contact numbers (Coletti et al. 2020; Feehan e Mahmud 2021; Jarvis et al. 2020)"

Did you consider including population density?

The construction of the contact matrices – which include the average number of contacts between a person in the age group with a person in the age group – requires the use of data on the underlying population by age group, but not the population density, which would also include information on the size of the area in which the population is located. As our focus is not on spatial transmission of SARS-CoV-2, we deemed it not necessary to include the information on the population density of the surveyed geographical areas in our models.

However, to make it clear for the reader that the population age distribution, , contributes to the construction of the age-specific contact matrices, we updated the supplementary material by providing more information on how such matrices are derived:

"First, for each country, based on the data collected in the POLYMOD study between 2005 and 2006 (Mossong et al. 2008), and in the Comes-F studies in 2012 (Béraud et al. 2015), we constructed pre-COVID age-specific social contact matrices , using the same age groups of the survey (18 – 24 years, 25 – 44 years, 45 – 64 years, and 65 years or more), which contained the average number of contacts between participants in the th age group and their contactees in the th age group, using a constraint for the reciprocity of contact when considering the total number of contacts at the population level. This constraint implies that, in a closed population, the total number of contacts from age class to age class must be equal to the total number of contacts from age class to age class . Hence, if we had individuals in age class and individuals in age class , reciprocity would entail that . For this purpose, we started by dividing the total number of contacts, , for the number of participants in each age group, , i.e., . We then constructed the matrix with the total number of contacts between age classes at the population level, by multiplying the expected mean number of contacts reported by participants in the age group with contactees in the age group , , for the total number of individuals in the age group in the population, i.e., . At this point, we adjusted for reciprocity by averaging the total number of contacts in one direction, , with those in the other direction, , weighting for the sample size by age group, namely, (Arregui et al. 2018). We finally obtained the expected average number of contacts under reciprocity at the population level dividing again by the population , i.e., ."

Please include strobe checklist or similar in appendix.

Thanks for the comment. Also following the request of the editor, we now provide a STROBE checklist for reports of cross-sectional studies with the revised manuscript, including the page numbers where all the listed information can be found.

Discussion

Further consideration of confounders or unknowns is needed throughout the discussion.

Thank you for this comment, as it gives us the chance to better discuss our choices for the epidemiological modelling part of the manuscript. In the Discussion, we added a fourth limitation, mentioning the choice of the initial estimate of the basic reproduction number and the extent to which the age-specific contact matrices that we built based on the collected contact data differed from those presented in similar studies during the early phase of the pandemic:

"Finally, the calculation of the net reproduction number hinges on two factors characterized by large uncertainty, namely, the initial estimate of the basic reproduction number and the age-specific

social contact matrix. For , we employed a common value for all parameters coming from a meta-analysis based on data from China, South Korea, and Italy; although this estimate was not specific for each surveyed country, the attached uncertainty made it consistent with estimates used in other studies (Coletti et al. 2020; Feehan e Mahmud 2021; Jarvis et al. 2020) and allowed us to disentangle the impact of the reduction in contact numbers from that of other epidemiological factors. For the contact matrices, as we did not collect the age of the individuals contacted by respondents like other single-country studies did (Backer et al. 2021; Coletti et al. 2020; Feehan e Mahmud 2021; Jarvis et al. 2020; Latsuzbaia et al. 2020), we scaled down the pre-pandemic matrices by age-specific factors derived from the data collected in our study. Although our matrices might not reflect the true social mixing patterns during the early phase of the pandemic, we believe that the results on the reduced infection transmission entailed by the reduced contact numbers is still valid and consistent with the results of the other studies."

Generalisability with a focus on country level is a limitation of the study. Is your model applicable to rural and urban areas alike?

Thanks for this comment. Considering that we performed a cross-country data collection and analysis and we post-stratified our data to make them approximately representative of the whole country population, we assumed that our results generalise to the country level. Nonetheless, the Supplementary Data File provided along the manuscript contains estimates of the average number of contacts (with standard errors) at different subnational levels, as well as by age and sex. This is consistent with the other social contact surveys that were conducted before and during the pandemic, as the focus is usually on the country level.
Although in this work we did not look at the distribution of participants between urban and rural areas in each country, this is something that would be worth of investigation in the future. Such analysis could shed light on the bias introduced by the Facebook (or other social media) recruitment in terms of rural/urban residence, which is something that we did not consider in our poststratification adjustments.
We recognized this limitation raised by the reviewer by clarifying that we generalised our results only to the country level and that possible bias in terms of rural/urban residence cannot be explicitly considered in the design of the Facebook advertisements:

"First, we collected data using a sample of Facebook users. Although such nonprobability samples are somewhat less accurate than probability samples in obtaining statistics that are generalisable to the whole population, with the appropriate statistical adjustments, such as those we made, they offer a good approximation to results that could be obtained from probabilistic samples, at least at the country level.

…

However, post-stratification weights cannot correct entirely for different behaviours or social inequalities (e.g., education level, ethnicity, type of profession, or urban/rural residence) among Facebook users, since such variables cannot be included in the ads stratification. Further improvements, which go beyond of the scope of this work, may come from using a multilevel regression and poststratification (MRP) approach, which has shown to work well in practice when making inference from nonprobability samples, especially when having at disposal data from census or a true probability sample on the demographic strata of interest in the population (Wang et al. 2015)."

Line 275: essential workers would be a more likely explanation for having high counts. Was information on profession collected?

Thanks for this comment. We collected information on the employment status of participants, adapting the question from the European Social Survey, but not on the profession of those who reported being in paid work. From the analysis of the determinants of the reported number of contacts, whose coefficients are now reported in the new Supplementary Fig. S6, we show that retired people reported on average less contacts than those in paid work, especially in countries like Germany and the Netherlands, where people

reported higher number of work contacts with respect to the other surveyed countries. This result is now made it explicit in the Results section:

"The analysis of the determinants of the contact numbers showed a pattern of cross-country heterogeneity, with a dependence on socio-demographic variables such as the household size and being in paid work – all positively associated – and, to a lower extent, being male, and reporting data during the weekend."

Moreover, to recognize the limitation due to the selection bias in terms of type of profession, we added, in the second limitation in the Discussion section, the type of profession to the list of social inequalities for which we could not adjust in the design of the Facebook recruitment campaigns:

"However, post-stratification weights cannot correct entirely for different behaviours or social inequalities (e.g., education level, ethnicity, type of profession, or urban/rural residence) among Facebook users, since such variables cannot be included in the ads stratification."

[1] "Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK: methods of estimation, data sources, causes of heterogeneity, and use as a guide in policy formulation"

VERSION 2 – REVIEW

REVIEWER	Kelly, Dervla University of Limerick, Graduate Entry Medical School
REVIEW RETURNED	30-Aug-2021

GENERAL COMMENTS	Many thanks to the authors for their detailed, well considered responses. One additional minor point: The lack of study data other important behaviours that affect transmission (frequency or compliance with mask wearing, physical distancing as well as environmental factors such as disinfections and ventilation) is a limitation and worth including discussion either as limitation or focus of future work.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Dr. Dervla Kelly, University of Limerick

Many thanks to the authors for their detailed, well considered responses.
One additional minor point: The lack of study data other important behaviours that affect transmission (frequency or compliance with mask wearing, physical distancing as well as environmental factors such as disinfections and ventilation) is a limitation and worth including discussion either as limitation or focus of future work.

To address the new comment, we clarified in the Discussion how the lack of data on some important behavioural (e.g., physical distancing) and environmental (e.g., ventilating indoor spaces) factors or the exclusion of the available behavioural data (i.e., wearing the face mask, increasing hand sanitation, and avoiding social activities) from the calculation of the reproduction number may be considered as a limitation of the study. We edited, therefore, the text in lines 288-295 of the clean revised manuscript:

"Changes in the are brought about not only by changes in social contact patterns, but also in infection transmissibility, which might be affected by other health behaviours (e.g., physical distancing in public spaces, increased hand washing or sanitizing, compliance with mask wearing [31,36]) or environmental factors (e.g., surface disinfection and indoor ventilation). Although our sample provided data on some of the behavioural factors (i.e., hand washing, face mask wearing, and avoiding social activities) – which we included in the estimation of the contact numbers – we could not explicitly account for their impact, as well as for the impact of other important factors for which we did not collect data, in the calculation of , which may be seen as a limitation of this study."