PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Parents' experiences regarding neonatal care during the COVID-19 pandemic – country-specific findings of a multi-national survey
AUTHORS	Kostenzer, Johanna; von Rosenstiel-Pulver, Charlotte; Hoffmann, Julia; Walsh, Aisling; Mader, Silke; Zimmermann, Luc; COVID-19 Zero Separation Collaborative Group, n.a.

VERSION 1 – REVIEW

REVIEWER	Miyawaki, Atsushi The University of Tokyo, Department of Public Health, Graduate School of Medicine
REVIEW RETURNED	17-Sep-2021

GENERAL COMMENTS	This study investigated and described how infant- and family-centered developmental care (IFCDC) was applied for newborns who received special/intensive care immediately after birth and their parents during the first year of the COVID-19 pandemic. Authors focused on 1148 participants in 12 geographically different countries. The article is simple cross-country description for IFCDC implementation and can be a basic data for discussion of global policy to advance IFCDC implementation. The manuscript is well-written overall but have some concerns to be improved. Methods: L167 "For this sub-analysis, countries having a minimum of at least 30 answers..... A subsequent country selection depending on pre-defined criteria, such as geographical variation and COVID-19 situation was conducted by main authors...using a consensus approach..." As long as I read the article 16 listed in the reference, authors excluded some countries with relatively large numbers of respondents (mainly in Europe), inlcluding UK, Netherlands, Portugal, and Norway. I can understand authors try to select only one country from one region (e.g. Scandinavian area), but it still remains a concern that they might intentionally select some countries so that the result looks better. Authors should explain more details about the criteria in selecting these countries, or adding the results for all the countries with 30+ respondents in a sensitivity analysis.

	Methods: This study is a sub-analysis of a paper in Reference 16. In Reference 16, they tested how different the implementation of IFCDC was depending on the strength of the non-pharmaceutical interventions (NPI) as perceived by the subject. This was limited in that they used subjective measures of NPI strength, and it was unknown whether or not the IFCDC implementation was associated with more objective NPI measure. In this study, authors may be able to bridge this knowledge gap by testing the correlation between country-level IFCDC implementation rate with country-level more objective NPI measure, such as COVID-19 GOVERNMENT RESPONSE TRACKER (https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker). Discussion: Discussion may become better if the authors can discuss more about the situation of the IFCDC implementation before the COVID-19 pandemic. As authors suggested, even before the COVID, IFCDC implementation had been very different across countries, depending on the policy history, healthcare systems, and healthcare resources in the neonatal care settings. For example, the low rate of IFCDC implementation rate in China would be observed even before the COVID (https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-020-09337-6). I acknowledge that authors are already addressing this point only in the limitations, but authors may be able to address pre-pandemic situation for some countries in the Discussion.

REVIEWER	Moinuddin, Md Edge Hill University, Medical School
REVIEW RETURNED	27-Oct-2021

GENERAL COMMENTS	Reviewer comments I thank the authors for their interesting and nice piece of work on investigating the parents' experiences regarding neonatal care during the COVID19 pandemic. A good project with enough sample however the work needs to be improved in several aspects such as data analysis, presenting results, and discussion. I would therefore recommend resubmitting the updated version of the article. Specific comments Comment 1: The authors described the situation in selected countries. However, the sample is not representative since the deliveries are mostly (72%) by C-section. Therefore, it is hard to say that the results obtained are would be reproducible. So how the conclusion would be useful is in doubt. The bad experience can be confounded by the delivery mode. Comment 2: Reporting overall percentage is not meaningful unless weighted. And if I am not wrong it will be almost impossible to the weighting. Alternatively, the overall estimate can be weighted using a meta-analysis model considering each country as an individual study. Comment 3: In table 3, the significance has been shown using colour code and the significance is generated using the confidence

	interval of the mean difference. The first point is that what is the reason for this testing while the authors compare country-specific mean with overall mean. Instead, between country comparison can be meaningful. Secondly, most of the variables are categorical so how did they calculate a mean and their difference from a categorical variable. Would you please make it clear? Alternatively, the CI for the difference in proportion can be estimated and between-country comparison can be useful. Comment 4: In the result section, the authors used a lot of adjectives. I would recommend avoiding them rather describe what they have observed in plain neutral language. Comment 5: In my opinion, the discussion is not strong enough. For every point of discussion, it would be good to answer why and how, the context, what could be the implications, by analysing the other literature (not just citing them rather discussing them). Secondly, without exploring the health systems of the country it is really hard to make conclusions. Therefore, I would recommend adding another table providing the country's sociodemographic profile such as per capita GDP, maternal education, maternal and child mortality, sanitation and hygiene etc. This will be helpful to understand the variation of the findings. Comment 6: High variation among the countries creates confusion about the results if it is something believable or found by a random chance. Also, if the parents' experience reported here is too problematic, reasonable? Please discuss Comment 7: Based on the findings, what recommendations or suggestions do the authors do? The article needs to address the 'so what' question. Comment 8: Line 246, the word 'answered' can be replaced by 'responded'

VERSION 1 – AUTHOR RESPONSE

Reviewer #1, Dr. Atsushi Miyawaki, The University of Tokyo	***Response***
This study investigated and described how infant- and family-centered developmental care (IFCDC) was applied for newborns who received special/intensive care immediately after birth and their parents during the first year of the COVID-19 pandemic. Authors focused on 1148 participants in 12 geographically different countries. The article is simple cross-country description for IFCDC implementation and can be a basic data for discussion of global policy to advance IFCDC implementation. The manuscript is well-written overall but have some concerns to be improved.	Thank you for your precise review and detailed feedback. We appreciate your constructive comments which we carefully considered as described in the following.
Methods: L167 "For this sub-analysis, countries having a minimum of at least 30 answers..... A subsequent country selection depending on pre-defined criteria, such as geographical variation and COVID-19 situation was conducted by main authors...using a consensus approach..." As long as I read the article 16 listed in the reference, authors excluded some countries with relatively large numbers of respondents (mainly in Europe), inlcluding UK, Netherlands, Portugal, and Norway. I can understand authors try to select only one country from one region (e.g. Scandinavian area), but it still remains a concern that they might intentionally select some countries so that the result looks better. Authors should explain more details about the criteria in selecting these countries, or adding the results for all the countries with 30+ respondents in a sensitivity analysis.	We appreciate your comment and agree with you that the criteria for our country selection were insufficiently addressed. Based on your comment and advice, we adapted the methods section (p. 5, line 174-178). Multiple criteria, including recently published scientific articles [1–5], and a consensus approach were used to select countries. The way in which the country-selection influences the interpretation of the results has been addressed in the limitations section of the paper (p. 14, line 501-505). Furthermore, we have now also referenced the recently published project report, which covers results of 30 countries [6]. Interesting findings during the research process for the country selection, such as data on the cumulative COVID-19 cases as of November 2020 and the average government response stringency index during the time of data collection were provided in Supplementary Table S2 to aid the explanation of the variation in the findings.
Methods: This study is a sub-analysis of a paper in Reference 16. In Reference 16, they tested how different the implementation of IFCDC was depending on the	We highly appreciate your feedback on this point and thank you for referring to the COVID-19 government response tracker. We agree with you that correlating our data with the respective indexes of the government

strength of the non-pharmaceutical interventions (NPI) as perceived by the subject. This was limited in that they used subjective measures of NPI strength, and it was unknown whether or not the IFCDC implementation was associated with more objective NPI measure. In this study, authors may be able to bridge this knowledge gap by testing the correlation between country-level IFCDC implementation rate with country-level more objective NPI measure, such as COVID-19 GOVERNMENT RESPONSE TRACKER (https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker).	response tracker would be valuable in order to estimate the clear effect of the stringency measures on our data. With this survey we aimed to collect data on the parent's perspective which is of course subjective. In our previous paper [7] we aimed to classify the perceptions based on the perceived restriction levels, while the aim of this paper was to describe experiences from parents in different countries in an exploratory way. The focus was not relating these experiences to different degrees of restriction levels. We agree with you that a correlation between the stringency index and IFCDC would be useful, however, there is no IFCDC implementation index available showing the degree of IFCDC implementation. This makes a correlation between the stringency index and IFCDC implementation impossible. What adds to this is that no data on the implementation of the different aspects of IFCDC on a national and international level in pre-pandemic times are available [8] (see also comment below). This has been acknowledged in the discussion section (page 13, line 460-462; page 14, line 520-528). Because we agree with you that the COVID-19 government response tracker of the respective countries is of interest to the reader, we added the mean values for the respective time frame together with other demographic and COVID-19 related data as Supplementary Table (Supplementary Table S2). Moreover, we referenced the tracker in various parts of the manuscript (see comment 5, reviewer #2) and discussed our results in relation to the mean stringency indexes (Discussion: page 13, lines 456-458; page 13, lines 484-486).
Discussion: Discussion may become better if the authors can discuss more about the situation of the IFCDC implementation before the COVID-19 pandemic. As authors suggested, even before the COVID, IFCDC implementation had been very different across countries, depending on the policy history, healthcare systems, and healthcare resources in the neonatal care settings. For example, the low rate of IFCDC implementation rate in China would be observed even before the COVID (https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-020-09337-6). I acknowledge that authors are already addressing this point only in the limitations, but authors may be able to address pre-pandemic situation for some countries in the Discussion.	As per your suggestion, we added a paragraph on the concept and different aspects of IFCDC and its implementation in the introduction (p. 4, line 97-102). Thank you also for the provision of a reference of IFCDC implementation in China. We added this to our paper. Moreover, we addressed the implementation of IFCDC again in the discussion (p. 13, line 456-462). Although international perspectives on family-centred care are described in the literature [9], comprehensive data on the implementation of the different aspects of IFCDC on a national and international level is lacking [8]. We decided not to discuss the pre-pandemic situation of IFCDC implementation more extensively as it would

Detailed response to reviewers

	require extensive further research into the different aspects of IFCDC in each country and exceed the scope of this paper.

Reviewer #2, Dr. Md Moinuddin, Edge Hill University	***Response***
I thank the authors for their interesting and nice piece of work on investigating the parents' experiences regarding neonatal care during the COVID19 pandemic. A good project with enough sample however the work needs to be improved in several aspects such as data analysis, presenting results, and discussion. I would therefore recommend resubmitting the updated version of the article.	We appreciate your positive feedback and thank you for your comments which we carefully considered. Your comments are very helpful and we addressed the respective parts and have added more details.
Comment 1: The authors described the situation in selected countries. However, the sample is not representative since the deliveries are mostly (72%) by C-section. Therefore, it is hard to say that the results obtained are would be reproducible. So how the conclusion would be useful is in doubt. The bad experience can be confounded by the delivery mode.	We appreciate your comment. C-section is recommended when vaginal delivery might pose significant risks to the mother or the baby [10], which is the case in a higher than normal frequency in our sample as it consists of preterm, sick and low birthweight infants who have been admitted to the NICU or special care unit. Furthermore, caesarean section (C-section) rates rise globally. According to the WHO, C-sections account for more than 1 in 5 (21%) of all childbirths as of today [11]. In Latin America and the Caribbean, rates are as high as 4 in 10 (43%) and in Brazil and Turkey C-sections have already outnumbered vaginal deliveries [11]. The literature suggests that in times of the pandemic, delivery via C-section is becoming even more common, both in infected and non-infected women [12]. We acknowledge that the C-section rate might have had an influence on attendance during delivery, but not necessarily on NICU visits and presence with the baby. The representativeness of the sample has been addressed as a limitation and we also added a sentence on the c-section rate in the limitations section (p. 14, line 512-515). The high c-section rate is furthermore highlighted in the results section (p. 6, line 222).
Comment 2: Reporting overall percentage is not meaningful unless weighted. And if I am not wrong it will be almost impossible to the weighting. Alternatively, the overall estimate can be weighted using a meta-analysis model considering each country as an individual study.	Thank you for your comment. We are aware that overall percentages without weighting need to be interpreted with caution and added this to the limitations section (p. 14, line 519-522). The percentages given in the column "total" were not taken as an average of individual country percentages but as the percentage of the total numbers, meaning that countries with a higher number of responses do weigh more as they contribute more to the overall number. However, as you already addressed, weighting is impossible due to the extensive amount of data. We decided against using a meta-analysis model because it is a study design mainly to assess two or more separate research studies to derive conclusion about a certain body of research [13] and exceeds the purpose of our descriptive study.
Comment 3: In table 3, the significance has been shown using colour code and the significance is generated using the confidence interval of the mean difference. The first point is that what is the reason for this testing while the authors compare country-specific mean with overall mean. Instead, between country comparison can be meaningful. Secondly, most of the variables are categorical so how did they calculate a mean and their difference from a categorical variable. Would you please make it clear? Alternatively, the CI for the difference in proportion can be estimated and between-country comparison can be useful.	Thank you for your comment. We used the colour code mainly to aid the reader through the data. Indeed, the term "confidence interval for the difference in mean" was not correct. We calculated the 95% CI for a single answer option (CI of proportion), not based on a mean. This has meanwhile been changed in the manuscript (p. 5, line 185-189; p. 10, line 318-319). The results of the calculation are shown in the added Supplementary Table S5 and allow for a comparison of the CI of the total sample and the included countries as well as between countries. Based on your suggestion, we also calculated the CI for difference in proportion. However, between-country comparison for every country combination (12 in total = 144 combinations) exceeded the purpose of this paper and was thereby not further elaborated. To minimise the risk of selection bias, we additionally calculated the confidence intervals based on all available, global data [6,7] and compared this with the CI of the individual countries. We did not find major differences in the results compared to our current analysis. We acknowledge that our calculation of confidence intervals has limitations, also as only one answer option per question was selected for further analysis which was addressed in the limitations section (p. 14, lines 501-503). We decided to only choose one answer option for each question to aid the readability of the table. The selection was made based on content choosing the most exclusive answer option of the question for calculation (mostly 'not at all').
Comment 4: In the result section, the authors used a lot of adjectives. I would recommend avoiding them rather describe what they have observed in plain neutral language.	Thank you for your comment. We have reviewed the results section accordingly. However, many of the adjectives used refer directly to the terminology used in the questionnaire and we therefore could not reduce them as much.

Comment 5: In my opinion, the discussion is not strong enough. For every point of discussion, it would be good to answer why and how, the context, what could be the implications, by analysing the other literature (not just citing them rather discussing them). Secondly, without exploring the health systems of the country it is really hard to make conclusions. Therefore, I would recommend adding another table providing the country's sociodemographic profile such as per capita GDP,	We appreciate your comment and added a Supplementary Table (Supplementary Table S2) with each countries' sociodemographic profile, including GDP per capita, female education, maternal and child mortality and sanitation. Moreover, we included data on the cumulative COVID-19 cases as of November 2020 and the average government response stringency index during the time of data collection (COVID-19 Government Response Tracker,

maternal education, maternal and child mortality, sanitation and hygiene etc. This will be helpful to understand the variation of the findings.	https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker) to understand the variation of the findings. We put the individual countries' information in context with the results of our study in the discussion (page 13, line 456-462; page 14, line 484-486).
Comment 6: High variation among the countries creates confusion about the results if it is something believable or found by a random chance. Also, if the parents' experience reported here is too problematic, reasonable? Please discuss	In our methods and results section, we have explained in more detail now how the countries were selected (p.5 line 174-178) and have also provided some more contextual information (see comment 5) to make this point clearer (p.14, line 421-422). However, differences in country characteristics were also outlined as a limitation in the discussion (p.14, line 512-515). We have no reason to doubt the big differences between countries regarding implementation of aspects of IFCDC as we know from our own visits to many countries (for e.g. the big difference between Sweden and China), from parent experiences and from literature both before COVID (please also see ref 16, 61 in the manuscript) and during COVID (see ref 3-15).
Comment 7: Based on the findings, what recommendations or suggestions do the authors do? The article needs to address the 'so what' question.	A Lancet commentary [14] and extensive project report [6], both based on our data and including a call to action, have been referenced and elaborated in the conclusion (p. 14, line 545-547).
Comment 8: Line 246, the word 'answered' can be replaced by 'responded'.	This has been changed.

Further references

1. Yan B, Zhang X, Wu L, Zhu H, Chen B. Why Do Countries Respond Differently to COVID-19? A Comparative Study of Sweden, China, France, and Japan. The American Review of Public Administration. 2020 Aug;50(6–7):762–9.
2. Claeson M, Hanson S. COVID-19 and the Swedish enigma. The Lancet. 2021 Jan;397(10271):259–61.
3. Baral S, Chandler R, Prieto RG, Gupta S, Mishra S, Kulldorff M. Leveraging epidemiological principles to evaluate Sweden's COVID-19 response. Annals of Epidemiology. 2021 Feb;54:21–6.
4. Amer F, Hammoud S, Farran B, Boncz I, Endrei D. Assessment of Countries' Preparedness and Lockdown Effectiveness in Fighting COVID-19. Disaster med public health prep. 2021 Apr;15(2):e15–22.
5. Garcia PJ, Alarcón A, Bayer A, Buss P, Guerra G, Ribeiro H, et al. COVID-19 Response in Latin America. The American Journal of Tropical Medicine and Hygiene. 2020 Nov 4;103(5):1765–72.
6. EFCNI, Kostenzer J, von Rosenstiel-Pulver C, Hoffmann J, Walsh A, Fügenschuh S, et al. Zero separation. Together for better care! Infant and family-centred developmental care in times of COVID-19 – A global survey of parents' experiences Project Report. EFCNI; 2021.
7. Kostenzer J, Hoffmann J, von Rosenstiel-Pulver C, Walsh A, Zimmermann LJI, Mader S. Neonatal care during the COVID-19 pandemic - a global survey of parents' experiences regarding infant and family-centred developmental care. EClinicalMedicine. 2021 Sep;39:101056.
8. Al-Motlaq MA, Carter B, Neill S, Hallstrom IK, Foster M, Coyne I, et al. Toward developing consensus on family-centred care: An international descriptive study and discussion. J Child Health Care. 2019 Sep;23(3):458–67.
9. Vetcho S, Cooke M, Ullman AJ. Family-Centred Care in Dedicated Neonatal Units: An Integrative Review of International Perspectives. Journal of Neonatal Nursing. 2020 Apr;26(2):73–92.
10. Gao Y, Xue Q, Chen G, Stone P, Zhao M, Chen Q. An analysis of the indications for cesarean section in a teaching hospital in China. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2013 Oct;170(2):414–8.
11. World Health Organization. Caesarean section rates continue to rise, amid growing inequalities in access. WHO, Geneva; 2021.
12. Arab W, Atallah D. Cesarean section rates in the COVID-19 era: False alarms and the safety of the mother and child. Eur J Midwifery. 2021 May 20;5(May):1–2.
13. Haidich AB. Meta-analysis in medical research. Hippokratia. 2010 Dec;14(Suppl 1):29–37.
14. Kostenzer J, Zimmermann LJI, Mader S, Abenstein A, Daly M, Fügenschuh S, et al. Zero separation: infant and family-centred developmental care in times of COVID-19. The Lancet Child & Adolescent Health. 2021 Nov;S2352464221003400.

VERSION 2 – REVIEW

REVIEWER	Moinuddin, Md Edge Hill University, Medical School
REVIEW RETURNED	05-Feb-2022

GENERAL COMMENTS	R1 Reviewer comments I thank the authors for their hard work to improve the manuscript and I congratulate them. I found the article more clear, readable and insightful. They have addressed the comments in an intelligent way. I am looking forward to seeing this polished piece of work out soon with minor corrections. R1 Minor comments Comment 1: page 6, line 187: ". . . the CI of all countries" the word would be instead ". . . the proportion of all countries". We do not test if CI is significantly different rather we test if the proportion is different using the CI. Comment 2: Just wondering how the ethics is waived while the authors are using the data from human participants. Comment 3: page 9, line 272: a bit confusing, probably typo, please check "around 80–>90%". If not a typo, try a different way to present it.

VERSION 2 – AUTHOR RESPONSE

Reviewer #2	***Response***
I thank the authors for their hard work to improve the manuscript and I congratulate them. I found the article more clear, readable and insightful. They have addressed the comments in an intelligent way. I am looking forward to seeing this polished piece of work out soon with minor corrections.	We thank you for your feedback and your positive evaluation of the revised manuscript. We appreciate your remaining suggestions which we carefully considered and implemented in this adapted version of the manuscript.
Comment 1: page 6, line 187: ". . . the CI of all countries" the word would be instead ". . . the proportion of all countries". We do not test if CI is significantly different rather we test if the proportion is different using the CI.	We agree with you and changed the wording as per your suggestion (see page 5, line 187).
Comment 2: Just wondering how the ethics is waived while the authors are using the data from human participants.	The procedure was the following: We submitted a complete ethics application to the Ethics Committee of Maastricht UMC+ including all information about the survey and the final questionnaire. After having reviewed our application, the Ethics Committee concluded that our survey is not a medical-scientific research. As no patient sensitive data are collected, the 'Medical Research Involving Human Subjects Act' (WMO) does not apply to this study. Thus, the Committee stated

Detailed response to reviewers

	"*we are pleased to confirm that the Medical Research Involving Human Subjects Act' (WMO) does not apply to the above mentioned study and that an official approval of this study by our committee is not required*".
Comment 3: page 9, line 272: a bit confusing, probably typo, please check "around 80–>90%". If not a typo, try a different way to present it.	Thank you for pointing to this imprecision. We changed the wording to "Between 80% and more than 90% of participants…" (see page 8, line 272).