PEER REVIEW HISTORY

BMJ Paediatrics Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

This paper was submitted to a another journal from Archives of Disease in Childhood but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Paediatrics Open. The paper was subsequently accepted for publication at BMJ Paediatrics Open.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Impact of stimulation among non-crying neonates with intact cord versus clamped cord on birth outcomes- observation study
AUTHORS	KC, Ashish Budhathoki, Shyam Thapa, Jeevan Niermeyer, Susan Gurung, Rejina Singhal, Nalini

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Dr. J Dekker Institution and Country: Leiden University Medical Center, Netherlands Competing interests: None
REVIEW RETURNED	28-Apr-2021

GENERAL COMMENTS	I would like to congratulate the authors with this interesting study. It is very important to study tactile stimulation, especially in this clinical scenario of intact cord management. These results will provide health care workers with more evidence on the need for intact cord stabilization at birth, which will likely guide and change clinical practice. When reviewing this paper, one major issue arose. In table 1, the proportions are not calculated correctly. For example, labour was induced in 482/1892 infants who were stimulated after cord clamping, which is in my opinion 25.5% instead of the 79.5% which is noted in the Table. It looks like the authors have calculated the proportions based on the total number of infants born after induction of labour (482+124=606). However, since the group of infants stimulated with an intact cord is much smaller than the group of infants stimulated after cord clamping, this calculation of proportions is in my opinion not correct, since you would always get lower percentages in the smallest group. I would like to suggest adaption of this entire table. In addition to this issue, statistics have to be recalculated, since there might be no significant differences with regard to these obstetrical and neonatal characteristics. Then, there would be no need to adjust for these variables.

	One other comment about this study is about the performance of stimulation. How was tactile stimulation defined? Was this intervention standardized? Or did the execution of this intervention differ between caregivers/centers/clinical scenarios? Maybe the authors could add a little bit more detail about this intervention in their methods section.

REVIEWER	Reviewer name: Not applicable Institution and Country: Not applicable Competing interests: None
REVIEW RETURNED	03-May-2021

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. The authors should be commended for their important work in establishing a quality improvement network in Nepal and the data generated from attending a large number of births across many sites. In this observational study, the authors studied vaginally born infants who did not cry immediately after birth. They conclude that stimulation with the cord intact may be more effective in triggering spontaneous breathing and results in better Apgar scores than stimulation provided after cord clamping. However, I do not think that this conclusion is valid given the following. 1. It is possible (and arguably probable), that infants born in poor condition (floppy, pale) triggered greatest concern for the attending health worker who clamped the cord early to commence bag mask ventilation. Similarly, infants born following difficult labour (for example, prolonged bradycardia) and/or difficult birth (for example shoulder dystocia) may have led to greater anxiety for the attending health worker leading to early cord clamping. The predictor variable “cord clamping prior to stimulation” is therefore likely to be a reflection of the sicker infants in the cohort. Unsurprisingly, these infants had worse outcomes than infants who were stimulated with intact cord, where the health worker is likely to have been less concerned regarding the clinical status of the infant. This is supported by the finding that the proportion of infants with 1-minute Apgar score <4 (as reported in the text- see my comment below regarding data in Table 2) was greater in the cord clamped group. It may theoretically be possible to try to answer the study question by statistically adjusting for how unwell the infant is at birth (for example, in a regression model). However, this would require detailed assessment of the infant’s condition immediately after birth, in addition to the antenatal/delivery factors. For example, tone, colour, heart rate. This data is difficult to measure, and for tone/colour, difficult to have standardised definitions that are not at risk of bias in an unblinded study. Fundamentally, I do not think the question of whether intact cord stimulation is associated with better outcomes (spontaneous breathing

	included) can be addressed with any validity outside a randomised trial. While my own bias aligns with the authors' suggestion that intact cord stimulation is likely to be beneficial for most infants in primary apnea, the resulting delay in bag-mask-ventilation for the sicker infants is a critical factor to take into account. The balance of benefit for infants in primary apnea and harm for infants in secondary apnea who need bag-mask ventilation can only really be evaluated in a well conducted randomised trial. I am concerned that findings from observational studies that cannot realistically adjust for critical confounders may be used to inform resuscitation guidelines. Some further points: 2. I'm afraid I do not understand how the variables for inclusion in the regression model were selected. The authors state, "The multivariable analysis using logistic regression was done to assess the association of cord status during stimulation on breathing by adjusting for the variables which had a significant level difference of p<=0.01. The variables which were significantly different in cord status during stimulation were complications during admission, mode of delivery and immediate drying." Firstly, how were variable selected for univariable comparison with the study outcome? For example, preterm and low birthweight are in the baseline characteristics table, but no data is provided on the univariable association with the study outcome. Were there any other variables collected and evaluated in the analysis, for example breech vaginal birth, shoulder dystocia, meconium stained liquor? Secondly, the text suggests that variables were selected based on the association with the variable "cord status", rather than the study outcome "breathing after stimulation". Is this correct? The authors also state a univariable threshold of <=0.01 for inclusion in the multivariable model, but mode of delivery had p=0.21. 3. The following definitions require clarification a) maternal complication during admission: any medical/obstetric complications during admission. It would be helpful to provide a list of complications. b) Breathing after stimulation. How was this determined by the attending researcher? Was it simply not requiring bag-mask ventilation? If yes, it would be simpler to have bag-valve-mask ventilation as the study outcome. 4. The percentages shown in Tables 1 and 3 are very difficult to

	follow. It would be more informative to show the values as in Table 2, where within each column, the total for a variable adds up to 100%. 5. Table 2, Apgar score in the 2 groups is inconsistent with the text. It appears here that the intact cord stimulation group had worse Apgar scores.

REVIEWER	Reviewer name: Dr. Hege Langli Ersdal Institution and Country: Stavanger University Hospital, Norway Competing interests: None
REVIEW RETURNED	26-May-2021

GENERAL COMMENTS	Thank you for the opportunity to review this work. This is said to be an observational study assessing the impact of intact versus clamped umbilical cord on start of spontaneous breathing after stimulation of non-crying neonates in Nepal. I commend the authors for the comprehensive work setting up these studies in Nepal. My main concern regarding this manuscript is that you do not report any specific variables related to the condition of the babies at birth, neither the timing of different actions/events from birth (i.e. how quickly did they start stimulation, when was cord clamping performed, initiation of breathing vs ventilation after birth). Due to the observational design and cord clamping on discretion of the midwives, one would suspect a great selection bias between the two groups (as long as you do not adjust for more specific variables as mentioned above). The midwives would likely tend to delay cord clamping and try stimulation among babies who seem less depressed, and thus more likely to respond to stimulation and start spontaneous breathing rather quickly. Therefore, I struggle with the conclusion, lines 43-57 page 13: “In conclusion, our study found offer non-crying neonates stimulated with the cord intact more like to breathe than those who are stimulated with cord clamped.” I think the language could be improved several places throughout the manuscript. Abstract Since this is an observational study and you don’t adjust for the babies’ condition at birth, I think you should reframe this conclusion as well: “This study provides evidence on the effect of the intact cord during stimulation.” I think you should be cautious with words like “evidence” and “effect”. Methods Study sites: Could you describe the resuscitation equipment/bay at the study sites? You state that the health care workers were trained to resuscitate with intact cord. Participants: In the abstract you say ≥ 34 weeks. Is it >33 or ≥ 34? Variables: You did not include any variables that could indicate the

	state/condition of the baby at birth? For example the final fetal heart rate and/or immediate newborn heart rate? Did you not include a variable to describe whether the baby was ventilated or not? Data collection: Were all deliveries observed by a clinical researcher? 24/7? If not, is there a potential selection bias? How do you know if stimulation happened before or after cord clamping – this is not explained in detail? I also think the timelines are essential – when did cord clamping happen in the two groups (e.g. from time of birth) in relation to stimulation and start of breathing or ventilation. Do you have data on that? I think these data together with more information about the babies' status at birth are essential to better understand a potential impact of stimulation with intact cord. Results It took me a while to understand table 1. For me it was intuitive (because of your design) to summarize the rows under each subheading for the two groups/columns (clamped vs intact). Maybe you should consider explaining a bit more how to read it, e.g. insert the denominator (n=…) for each row? What is reported in row 1? It is a bit confusing because you state in the methods that you are comparing the two groups (intact vs clamped cord), but in the result section you report differences among those being stimulated with intact cord for different obstetric and neonatal characteristics? "Among the neonates whose mother had complications during admission, intact cord during stimulation was lower than if mother had no complication (20.6% vs 26.9%, p-value=0.022)" etc. Table 2, I think there are some mistakes in the cut-off definitions of Apgars? And it is not consistent with the text. When (after birth) and how is mortality recorded? How is a stillborn defined? It may be okey to state that there is a 94% higher odds of breathing if the cord was intact etc.. But I think it is more common (and precise/correct) to say that the odds of starting breathing versus not starting breathing increased 1,94 fold in the group with intact cord during stimulation. Discussion Lines 38-57 at page 11; I am not sure how "similar" (and/or relevant) the studies you report here are to your study. They seem to include premature babies who required ventilation (CPAP and stabilisation or PPV as part of resuscitation?). Gestational age is not reported in your study, but from table 1, it seems to be mostly term babies. Lines 3-13 page 12; "…effect on delayed versus cord clamping"? Please explain what you mean. What was defined as "resuscitation measures performed" in this reference 19? Again, how "similar" is this study to your present study? Lines 24-30 page 12; Reference 23 investigates the relationship between time to cord clamp and onset of breathing or initiation

	of PPV following stimulation/suction, AND 24-hour outcome. Time to start ventilation (among those who receive it) is critical. Reference 39 does not investigate if there is "an association between time to cord clamping and onset of breathing or initiation of PPV following stimulation/suction" as you state. Overall, I think the discussion section is difficult to follow and often unclear to me. Several places, I struggle to see/understand the relevance/relation to what your data potentially show/indicate. Hopefully you could explain this in more detail.

VERSION 1 – AUTHOR RESPONSE

27 June 2021

Professor Imti Choonara, MBChB, MD, FRCPCH, DTM&H
Derbyshire Children's Hospital
University of Nottingham
Derby, UK

Dear Editor-in-Chief,
Re- Impact of stimulation among non-crying neonates with intact cord versus clamped cord on birth outcomes- observation study
The International Liaison Committee on Resuscitation (ILCOR) 2020 recommends that a neonate who is not crying or breathing with poor tone and heart rate less than 100 bpm should receive stimulation and clearing of airways (as needed) with intact cord. There is very little robust evidence on the impact of stimulation in neonates receiving resuscitation with the cord intact. World Health Organization (WHO) identifies resuscitation with the cord intact as a key research area to generate a stronger evidence base for care. Experimental studies both in animals and humans have shown breathing prior to umbilical cord clamping has been shown to result in smoother cardiovascular transition at birth.
In this large scale multi-centric study, stimulation with the cord intact have shown to be more effective in triggering spontaneous breathing and result in better Apgar scores than stimulation provided after cord clamping. This approach will provide the added advantage that stimulation provided by a single provider can induce spontaneous respiration with the cord intact in primary apneic neonates and would not require ventilation and additional providers, especially in low-resource settings.
This study provides evidence on benefit of stimulating the non-crying babies with cord intact. You have submitted also the updated manuscript based on the comments from the reviewers.

Ashish KC, on behalf of co-authors and senior authors
Researcher, Department of Women's and Children's Health,
Uppsala University, Dag Hammarskjölds väg 14B, Uppsala, Sweden

VERSION 2 – REVIEW

REVIEWER	Reviewer name: Not applicable Institution and Country: Not applicable Competing interests: None
REVIEW RETURNED	14-Jul-2021

GENERAL COMMENTS	Thanks for the opportunity to review this revised manuscript. I would like to reiterate that authors should be commended for their important work in establishing a quality improvement network in Nepal and systematically collecting data from a large number of births.

	The main improvement in the revised manuscript is that the tables are now a lot easier to follow. Attempts have been made to address the points raised by reviewers. However, there are still significant issues that need careful attention. Abstract The last sentence of the conclusion is still overstated and not supported by the data, given the methodology. The authors would probably be better off stating that intact cord stimulation may help establish spontaneous breathing, but residual confounding from unmeasured or incompletely adjusted variables are likely contributing to the findings. The authors point out the issue of confounding in the revised 1st paragraph of the Discussion, but readers may not get to this section. It would be better to have a consistent set of valid conclusions throughout the paper. The authors should feel that being clear about the limitations provides confidence and highlights the important knowledge gaps to readers. What this study adds: Same issue as above, the conclusion is not supported by the data. Methods: the authors should expand on whether the researchers entering the data were aware of the study hypothesis. The potential for confirmation bias during data entry from this unblinded study should also be addressed in the Discussion, particularly as 'stimulation' is subjective (see my point re Fig 1 below where clarity is needed). Details of what maternal conditions contributed to 'complications during admission' are now provided. The authors state in their response that breech, dystocia and meconium-stained liquor are included in the definition of complication during admission, but this is not included in the list in the revised manuscript. It is unclear why the authors chose to group these variables together into 1 variable, 'complications during admission'. The sample size is large so it would be more valid to perform the uni- & multivariable regressions taking advantage of the granularity available in the dataset. Grouping the variables together risks existing relationships being masked. For example, if breech birth was positively associated with both early cord clamping and no spontaneous breathing, it would be a confounder, but the relationship would be lost by grouping breech with all the other variables that are included under 'complications during admission'. There are other potential antenatal/intrapartum confounders not included in the grouped variable 'complications during admission'. For example, maternal diabetes, reduced fetal movements, obstructed labour, prolonged 2nd stage. If there are no data on these variables, the authors should list them as examples of potential sources of residual confounding. After birth, the major confounder is the infant's condition (tone, colour etc.) This is a critical confounder as stated in my previous comments. It is unlikely that adjusting for 'complications during admission' (particularly given the limitations above) would adequately control for its effect. For example, there are births where there are no recognised complications, but the infant is born in poor condition. This confounder should be given appropriate prominence in the manuscript. Given the methodological limitations detailed above, the Discussion generally overstates the implications of the study findings. Fig 1, What happened to infants who were not crying nor stimulated? What definition of stimulation was used to exclude this

	group? I'm not sure Figure 2 adds anything useful. This data is reported in the manuscript text and Table. Table 2, the Apgar scores are still overlapping suggesting data inconsistency. Table 4, the intercept is not useful and impedes readability of the table. P values for low birth weight in the biavarite analysis are shown as "0", whereas for other variables are shown as "<0.0001". This needs to be presented consistently.

REVIEWER	Reviewer name: Dr. Rachel Mary Hilliam Institution and Country: The Open University, Mathematics and Statistics Competing interests: None
REVIEW RETURNED	24-Jul-2021

GENERAL COMMENTS	This is an interesting study and the authors should be complemented on the set up of this design which will not have been easy. There are a few statistical points which would help the readers of the journal Firstly I'm assuming that the p-values quoted in Tables 1-3 are p-values associated with a chi-squared analysis of independence rather than odds ratios? This should be made clear in the table label or in the text. If the p-values are odds ratios then the results section needs to reflect this. I assume that by cOR you are referring to crude odds ratios (or unadjusted odds ratios) and aOR you are referring to adjusted odds ratios. This should be made clear somewhere in your text. Particularly as you have already stated that you are adjusting in the logistic regression. As a general point when giving the size of cohorts (such as in Figure 1 - but also other places) I would prefer to see n=671 rather than n-671, is at an initial glance in the text that could be read as the total number in the study 'n' minus 671 which is definitely now what you mean. Thank you for submitting your work which I'm sure will be of interest to the readers of the journal.

VERSION 2 – AUTHOR RESPONSE

31 August 2021

Comment 1#. Abstract: The last sentence of the conclusion is still overstated and not supported by the data, given the methodology. The authors would probably be better off stating that intact cord stimulation may help establish spontaneous breathing, but residual confounding from unmeasured or incompletely adjusted variables are likely contributing to the findings. The authors point out the issue of confounding in the revised 1st paragraph of the Discussion, but readers may not get to this section. It would be better to have a consistent set of valid conclusions throughout the paper. The authors should feel that being clear about the limitations provides confidence and highlights the important knowledge

gaps to readers.

Response- Our revised conclusion is "Intact cord stimulation may help establish spontaneous breathing in apnoeic neonates, but residual confounding variables may be contributing to the findings."

Comment 2#. What this study adds: Same issue as above, the conclusion is not supported by the data.

Response- We have now revised it as "Intact cord stimulation to breathe with the cord intact may help deliver intervention more quickly than stimulation after cord clamping and may also avoid the reflex bradycardia that can occur when the cord is clamped before the lungs are aerated."

Comment 3#. Methods: the authors should expand on whether the researchers entering the data were aware of the study hypothesis. The potential for confirmation bias during data entry from this unblinded study should also be addressed in the Discussion, particularly as 'stimulation' is subjective (see my point re Fig 1 below where clarity is needed).
Response- The data collector involved in the study were trained to collect data on intrapartum care, immediate newborn care and neonatal resuscitation. Researchers were not aware of the hypothesis of this study. However, data collectors were aware that this research aimed to assess the quality of care.

Comment 4#. Details of what maternal conditions contributed to 'complications during admission' are now provided. The authors state in their response that breech, dystocia and meconium-stained liquor are included in the definition of complication during admission, but this is not included in the list in the revised manuscript.
Response- We have mentioned the list of complications during admission. Previously, we did not provide the whole list of complication during admission assessed by the data collectors. The list of complications during admission collected were pre-eclampsia, eclampsia, diabetes, fever, premature rupture of membranes, pre-term premature rupture of membranes, oligohydramnios, cephalo-pelvic disproportion, breech / transverse lie, prolonged labour, decreased fetal movements, ante-partum haemorrhage, chorioamnionitis, cord prolapse, cord around the neck and medical complication non-pregnancy related complication diabetic mellitus, pre-existing hypertension.
We have now done additional analysis on meconium stained amniotic fluid between two groups.

Comment 5#. It is unclear why the authors chose to group these variables together into 1 variable, 'complications during admission'. The sample size is large so it would be more valid to perform the uni- & multivariable regressions taking advantage of the granularity available in the dataset. Grouping the variables together risks existing relationships being masked. For example, if breech birth was positively associated with both early cord clamping and no spontaneous breathing, it would be a confounder, but the relationship would be lost by grouping breech with all the other variables that are included under 'complications during admission'.
Response- There is a list of complication during admission collected, we have now provided it We are attaching the form 1 as appendix.
We have conducted the additional analysis on the list of maternal complication during admission between two groups. We found that neonates who had cord intact had lower proportion of breech or transverse lie during labour than those who had the cord clamped (1.3% vs 3.5%, p-value =0.003). Neonates who had cord intact have had lower proportion of with meconium stained amniotic fluid during labour than those who had the cord clamped (22.2% vs 28.4%, p-value=0.002).
We have now mentioned it in the result section now.

Comment 6#. There are other potential antenatal/intrapartum confounders not included in the grouped variable 'complications during admission'. For example, maternal diabetes, reduced fetal movements, obstructed labour, prolonged 2nd stage. If there are no data on these variables, the authors should list them as examples of potential sources of residual confounding.
Response- We have now provided the list of maternal complications during admission and conducted additional analysis on complication between two groups.

Comment 7#. After birth, the major confounder is the infant's condition (tone, colour etc.) This is a critical confounder as stated in my previous comments. It is unlikely that adjusting for 'complications during admission' (particularly given the limitations above) would adequately control for its effect. For

example, there are births where there are no recognised complications, but the infant is born in poor condition. This confounder should be given appropriate prominence in the manuscript.
Response- We highly appreciate the reviewer's explanation of the possible confounder which might have led to clinician's bias for determining the timing of cord clamping. We agree that since that is not a randomized clinical trial where strict criteria would have been used to clamp the cord among non-crying neonates as per the intervention group. This is an observation study, where the researcher did not intervene the clinician's performance on resuscitation as a result "residual confounder" might have created the bias. If the neonate was not crying, deeply cyanotic and profoundly hypotonic, the provider might have chosen to clamp and cut the cord immediately, knowing that the neonate is unlikely to respond to stimulation alone. If the neonate is not crying, but only little blue and has some tone, the provider most likely stimulated the neonate first with the cord intact. We have provided a thorough explanation to the bias in the methodological section. This study suggests to the further clinical controlled trial.

Comment 8#. Given the methodological limitations detailed above, the Discussion generally overstates the implications of the study findings.
Response- We have now added an explanation of the possible reason for health workers bias toward cord clamping based on the neonatal overall appearance i.e residual confounder as the methodological limitation. The provider's comfort to cord clamp neonates.

Comment 9#. Fig 1, What happened to infants who were not crying nor stimulated? What definition of stimulation was used to exclude this group?
Response- We did not include those infants who were not crying and were not stimulated in the analysis. Among those infants ventilation was done. We excluded based on whether the non-crying infants who were not stimulated. The definition of stimulation was "additional rubbing of back when the baby was assessed to be not-crying"

Comment 10#. I'm not sure Figure 2 adds anything useful. This data is reported in the manuscript text and Table.
Response- We have now removed it.

Comment 11#. Table 2, the Apgar scores are still overlapping suggesting data inconsistency.

Response- We have made up the cut off of Apgar score at 1 minute 3 or less and 4 or more and Apgar score at 5 minute as 6 or less and 7 or more.

Comment 12#. Table 4, the intercept is not useful and impedes readability of the table. P values for low birth weight in the biavarite analysis are shown as "0", whereas for other variables are shown as "<0.0001". This needs to be presented consistently.

Response- We have now ensured that if the p-value is kept as <0.0001 for value. Intercept is an important value in the logistic regression. In the bi-variate analysis, each comparing variable has an intercept. However, in multi-variate analysis, there is a single intercept for all the variable adjusted, rather than variable specific intercept.

Response to second reviewer's comment

Comment 13#. Firstly I'm assuming that the p-values quoted in Tables 1-3 are p-values associated with a chi-squared analysis of independence rather than odds ratios? This should be made clear in the table label or in the text. If the p-values are odds ratios then the results section needs to reflect this.
Response- The p-value mentioned in tables 1 to 3 are associated with a chi-square analysis, we have mentioned it in method section.

Comment 14#. I assume that by cOR you are referring to crude odds ratios (or unadjusted odds ratios) and aOR you are referring to adjusted odds ratios. This should be made clear somewhere in your text. Particularly as you have already stated that you are adjusting in the logistic regression.
Response- We have made it clarified cOR as crude odds ratio and aOR as adjusted odds ratio in the footnote of table 3.

Comment 15#. As a general point when giving the size of cohorts (such as in Figure 1 - but also other places) I would prefer to see n=671 rather than n-671, is at an initial glance in the text that could be read as the total number in the study 'n' minus 671 which is definitely now what you mean.

Response- We have kept it as n = 671 rather than n-671.