Author’s response to reviews

Title: Nonsynostotic plagiocephaly: a child health care intervention in Skaraborg County in Sweden

Authors:
Freda Lennartsson (freda.lennartsson@gmail.com)
Per Nordin (per.nordin@skaraborg-institute.se)

Version: 3 Date: 29 Oct 2018

Author’s response to reviews:

THANK YOU BOTH KINDLY FOR REVIEWING THE MANUSCRIPT. YOUR COMMENTS HAVE HELPED TO STRENGTHEN THE MANUSCRIPT.

Reviewer reports:

Brent R. Collett (Reviewer 1): This study adds to the sparse literature on preventive interventions for non-synostotic plagioccephaly, a common condition in infancy and frequent reason for referral to craniofacial and other specialty clinics. Other interventions described in the literature (e.g., orthotic helmets) are costly with modest evidence to support their effectiveness. Utilizing nurses as interventionists, the approach described has good potential to improve the standard of care in primary care settings. The authors appear to have done a nice job addressing common in this type of study (i.e., blinded assessment, reliability estimates for head shape measures).

There are several limitations that make it difficult to interpret the findings, some of which might be addressed depending on existing data.

1. Treatment fidelity. Unless I missed it, there do not appear to have been procedures to assess and monitor treatment fidelity. This is important on two levels: (1) Given that there were 35 different nurses implementing the intervention, it seems likely that there were differences in the extent to which they followed the treatment protocol. It may also be that nurses in the control group were making similar recommendations, even without prior exposure to this intervention. Within interventionists, there may have been 'drift' over time. How was this assessed and monitored, and were there procedures for re-training nurses as needed? (2) Among parents, it would be ideal to include data on behavior change and implementation of treatment
recommendations. For example, it would be interesting to limit the analysis to parents who implemented the recommendations at some minimum threshold. Treatment fidelity is especially important, given the preliminary nature of this research

... i.e., showing that the intervention itself resulted in changes, determining which components are most important, how to replicate the intervention in future research, etc.

THIS IS VERY TRUE, AND INTERESTING THAT YOU WRITE THIS. IN THIS STUDY, WE FOCUSED ON CRANIAL SHAPE. TREATMENT FIDELITY CAN BE SEEN IN TWO ARTICLES WE PUBLISHED PREVIOUSLY USING THE SAME SAMPLE. PLEASE SEE: LENNARTSSON, TEACHING PARENTS HOW TO PREVENT ACQUIRED CRANIAL ASYMMETRY; AND LENNARTSSON, INTEGRATING NEW KNOWLEDGE INTO PRACTICE: AN EVALUATION STUDY ON A CONTINUING EDUCATION FOR SWEDISH CHILD HEALTH NURSES ON NON-SYNOSTOTIC PLAGIOCEPHALY.

2. Baseline characteristics. I understand the rationale for assigning participants to treatment vs. control groups, rather than using randomization. It is reassuring that the two groups of infants were similar in most respects (I would recommend including Table S2 in the article, and add any demographic information available ... this seems too important to include only as a supplemental file). It would be helpful to include some description of the clinics and nurses in the two groups. For example, do the clinics serve similar patient populations in terms of socioeconomic status, etc.? Are the nursing staffs comparable in the two groups, or might they differ in important ways (e.g., education level, years of experience, etc.)?

ALL CHILD HEALTH NURSES HAVE SPECIALIZED IN PUBLIC HEALTH AND/OR PEDIATRICS WHICH IS WRITTEN IN THE BACKGROUND. THEY HAVE DIFFERENT YEARS OF EXPERIENCE WHICH WE DID NOT INVESTIGATE. ALL WERE ACCEPTED AS PEERS. YES, THE CLINICS DO SERVE SIMILAR PATIENT POPULATIONS. WE HAVE NOW INCLUDED THIS IN THE MANUSCRIPT UNDER SETTING, SENTENCE 2. WE HAVE ALSO PLACED TABLE S2 IN THE MAIN MANUSCRIPT AS YOU SUGGESTED.

3. Continuous vs. categorical measures. Given the preliminary nature of the study and the relatively small sample size, it seems that it might have been better to use continuous measures (e.g., standardized effect size estimates) vs. categorical outcomes. I understand that this might raise questions about clinical relevance, but it would also provide more robust data about treatment effects and help to justify larger scale replications of the intervention.
YES, WE CERTAINLY AGREE IT WOULD HAVE BEEN BETTER. THE RATIONALE FOR USING THIS METHOD WAS TWO FOLD: WE WANTED TO HAVE A NON-INTRUSIVE METHOD SO THAT PARENTS, INFANTS AND EVEN NURSES WOULD NOT BE SCARED OFF; ALSO, THE SEVERITY ASSESSMENT FOR PLAGIOCEPHALY AND THE SEVERITY ASSESSMENT FOR BRACHYCEPHALY SEEM TO BE USEFUL IN THE CLINICS. SO, WE INTRODUCED THEM TO THE NURSES IN THIS WAY. FOLLOW-THROUGH WAS VERY GOOD, SO THE METHOD HAD SOME MERIT.

4. Parent report data. More information is needed regarding parent report measures (time spent in a bouncer, car seat, tummy time, etc.). It is not all that surprising that these measures are highly variable, and I question the reliability/validity. For example, how recent was the interval for their reporting (i.e., last week, last day, etc.)? Am I correct in assuming that these were open-ended questions? Further, it seems likely that responses would be influenced by social desirability (i.e., parents might know that they are supposed to encourage tummy time/limit time in positioning devices, and this may affect their responses).

YES, THE QUESTIONS ABOUT POSITIONING DEVICES WERE OPEN-ENDED QUESTIONS. THE DATA WAS RECENT BECAUSE PARENTS WERE ASKED TO FILL IN THIS QUESTIONNAIRE AT THE TWO-MONTH WELL-CHILD VISIT. PARENTS DID NOT SEEM RELUCTANT ABOUT REPORTING LONG PERIODS OF TIME THEIR INFANTS SPENT DAILY IN POSITIONAL DEVICES AT THIS AGE BECAUSE THIS KNOWLEDGE IS NOT COMMONLY GIVEN TO PARENTS OF INFANTS IN SWEDEN. THIS IS KNOWLEDGE INTRODUCED IN THE EDUCATION FOR INTERVENTION GROUP NURSES.

WE HAVE MOVED TABLE S2 TO THE MAIN MANUSCRIPT AS YOU SUGGESTED. THE INFORMATION ON POSITIONAL DEVICES IS MORE ACCESSIBLE THAT WAY.

Deborah A. McNeil, PhD, RN (Reviewer 2): Abstract

Aim: the sentence needs a period at the end.

YES.
Results: unclear if the asymmetry reversal OR at 12 months for parent information was for all infants and what impact the intervention had on that OR.

NO, THIS REFERS TO THE SUBGROUP OF INFANTS WITH ASYMMETRY AT 2 MONTHS. THE IMPACT WAS FOUR TIMES MORE COMMON. BOTH ARE NOW INCLUDED IN THE ABSTRACT.

Unclear who were the non cases; non cases of what? Was it the infants who did not have brachycephaly at birth?

NON-CASES WERE INFANTS WITH NO ASYMMETRY. WE HAVE NOW RE-WRITTEN THIS IN THE ABSTRACT.

Conclusion what does benefitted reversal mean? Do the authors mean that those in the intervention group were more likely to have reversal of asymmetry?

YES, THAT IS WHAT IT MEANS. WE HAVE RE-PHRASED WITH “CONTRIBUTED TO “.

Background

Nice definition of NSP. As I understand it the intervention was education of the nurses to follow the Guidelines that were developed and the key to that was educating the parents in prevention.

YES. FIRST THERE IS AN EDUCATIONAL INTERVENTION AND THEN A CLINICAL INTERVENTION.

Methods

Page 4 lines 13 and 14 - Good attempt to address issues of "spill over" for nurses not providing the intervention by using nurses who did not attend the education program.

Pag4 lines 17 and 18 - Were the nurses who had not attended the education program but were in the intervention group provided with education?

ALL INTERVENTION GROUP NURSES PARTICIPATED IN THE 1½ HOUR CONTINUING EDUCATION ON NSP AT THEIR WORK PLACE. NO CONTROL NURSE DID (UNTIL AFTER THE STUDY WAS OVER. THEN CONTROL GROUP NURSES WERE
GIVEN THE SAME OPPORTUNITY). THE CONFUSION IS DUE TO THE MISTAKE IN THE CONSORT FLOW CHART WHICH IS NOW CORRECTED. GOOD YOU CAUGHT THAT!

Was there anything done to assess intervention integrity?

ETHICAL CONSIDERATIONS ARE FOUND ON PAGE 6 LINE 53 - PAGE 7 LINE 8. REGIONAL ETHICAL REVIEW BOARD APPROVAL WAS OBTAINED. PARENTS PROVIDED WRITTEN INFORMED CONSENT.

It would have been helpful to have more information about the distribution of the education and how many of the population of nurses attended the education program so we would know who many nurses by group were eligible to participate in the study.

ALL 35 INTERVENTION GROUP NURSES PARTICIPATED IN THE 1½ HOUR CONTINUING EDUCATION ON NSP. NO CONTROL NURSE DID. THE CONFUSION IS DUE TO THE MISTAKE IN THE CONSORT FLOW CHART AS EXPLAINED ABOVE. THIS IS CORRECTED NOW.

In looking at the supplemental material Consort chart for nurses, it identifies that all control group nurses received the intervention. This is at odds with what is written in the manuscript in the lines identified above. My sense is that the control intervention was routine care; it would help to explain this in the chart or change the wording to make it clearer. YOU PICKED UP A MISTAKE IN THE FLOW CHART. THE CONTROL GROUP NURSES DID NOT RECEIVE ALLOCATED INTERVENTION. THIS IS NOW CORRECTED.

It would also have been helpful to know the infant/parent population eligible to participate.

BIRTH OF INFANTS WITHIN THE COUNTY DURING RECRUITMENT TIME WERE THE ELIGIBILITY CRITERIA. RECRUITMENT TIME IS NOW INCLUDED UNDER PARTICIPANTS. PARENTS FOLLOWED ALONG WITH THEIR INFANTS.

There is a brief description of drop outs but at this point it is unclear.
ON PAGE 4 LINES 30-31 WE EXPLAIN THAT THREE NURSES DROPPED OUT WHEN THEY TERMINATED EMPLOYMENT AND SIX FAMILIES MOVED.

what the response rate was.

ALL BUT ONE PARENT AGREED TO PARTICIPATE. SEE LINE 30.

I would have like to have seen a calculated percent of the drop outs in the manuscript.

OK. THIS IS INCLUDED NOW.

The consort flow chart is also difficult to understand as it appears that all infants received the intervention.

THERE WERE MISTAKES IN THE CONSORT FLOW CHART WHICH ARE NOW CORRECTED. INTERVENTION GROUP INFANTS RECEIVED THE INTERVENTION AND CONTROL GROUP INFANTS RECEIVED ROUTINE CARE.

If the CONSORT flow charts remain as supplemental material there needs to be more in the text to clarify the above questions.

MORE INFORMATION IS NOW INCLUDED IN THE PARAGRAPH SETTING, PARTICIPANTS, AND GROUP ALLOCATION. ONE FLOW CHART WOULD HAVE BEEN FINE IN THE MANUSCRIPT, BUT WE HAVE TWO – ONE FOR THE NURSES AND ONE FOR THE PARENTS.

Page 4 Line 42- Please define what is meant by regularly. How often on average did nurses see infants?

REGULARLY IS NOW EXPLAINED IN THE MANUSCRIPT.

Page 4 Line 56 - how frequent in the first few months on average were visits? WEEKLY THE
FIRST MONTHS, MONTHLY UNTIL SIX MONTHS. THIS IS NOW INCLUDED IN THE MANUSCRIPT.

Page 5 lines 51-53 - Was the rating scale developed by the authors’?

CRANIAL TECHNOLOGIES DEVELOPED THE “SCALES” AND THE AUTHORS DEVELOPED A RATING SYSTEM.

In what way was this validated and reliability and validity assessed?

THE “SCALES” AND THE RATING SYSTEM WERE NOT VALIDATED. THE ASSESSORS WERE RELIABILITY-TESTED USING THE “SCALES” BEFORE THE CLINICAL INTERVENTION WAS LAUNCHED. PLEASE SEE OUR RELIABILITY STUDY: LENNARTSSON, RELIABLE ASSESSORS OF INFANT CRANIAL SHAPE IN CHILD HEALTH CARE

I do not believe that the next section on validation for 12-month measurements addresses the above question as this was for the blinded assessors and not for the nurses.

YOU ARE RIGHT. THIS IS SOMETHING DIFFERENT. THIS SECTION EXPLAINS HOW WE RELIABILITY TESTED THE BLINDED ASSESSORS’ 12 MONTHS ASSESSMENTS BY ALSO TAKING CRANIAL MEASUREMENTS WHICH WERE THEN MATCHED TO STANDARDS FOR SEX AND AGE.
Page 6 lines 31 and 32 - References for the following statement is needed. "The natural course, parents' habits, and national recommendations can influence cranial shape of all infants".

THIS IS THE AUTHORS’ ASSUMPTION, SO WE ADDED THE WORD CONCEIVABLY.

Page 6 line 48 should be in past tense. CORRECTED.

Page 6 - line 49 please spell out AC2.

THE ABBREVIATION AC2 WAS ALREADY EXPLAINED ON PAGE 6 LINE14

Page 6 line 51 I believe there is an error in the P value and it should be < 0.05 not 0.50.

YES. CORRECTED.

The entire Statistical methods section requires more information. For instance what was compared for the Chi2 analysis. What odds were being modeled in the regression analysis. What approach was taken for the multivariable analysis. How were co-variates or independent variables chosen? How was the modeling carried out? Was there any model validation undertaken?
THE USE OF REGRESSION MODELS CAN HAVE SEVERAL DIFFERENT APPROACHES. IN THIS SETTING, WE ARE NOT TRYING TO CREATE A MODEL WHERE THE TARGET IS A HIGH DEGREE OF EXPLANATION AND THE REASON FOR THIS IS QUITE STRAIGHTFORWARD. WE ARE SEEKING TO UNDERSTAND WHAT VARIABLES CAN BE OF VALUE IN RELATION TO THE OUTCOME. WE ARE NOT TRYING TO ESTABLISH AN EXHAUSTIVE MODEL AND WE ARE NOT MAKING SUCH CLAIMS. THEREFORE, THE CONCEPT OF MODEL VALIDATION BECOMES NOT AS IMPORTANT AS TRYING TO IDENTIFY A SET OF VARIABLES, POSSIBLE CANDIDATES FOR CONTINUED RESEARCH. WITHIN OUR SLIGHTLY MORE RESTRICTIVE FRAMEWORK WE USE THE MODEL WITH THE BEST FIT ALBEIT NOT CLAIMING THIS TO BE COMPLETE.

Results

In general the results section requires a significant amount of re-working. If there was statistical testing done then the p values should be displayed along with identifying the test that was used. I see that some results are reported in Tables 1 and 2 but it feels that there is a disconnect between what is in the text and the tables. Since there was statistical testing there should not be differences ascribed between groups unless there was a statistical difference. The number of sub-analyses with small samples is a problem and should be justified as more exploratory and addressed as a limitation.

P-VALUES AND TESTS ARE ALL SEEN IN THE TABLES. ONLY THE MOST INTERESTING ONES ARE EXPLAINED IN THE TEXT.

WE BELIEVE THAT STATISTICAL SIGNIFICANCE IS IMPORTANT, YES. BUT SOMETIMES CLINICAL INTEREST JUSTIFIES PRESENTING RESULTS THAT ARE NOT STATISTICALLY SIGNIFICANT. THE REASON FOR THIS IS THAT WHEN WE MOVE TOWARDS SUB ANALYSES THE SAMPLE SIZES ARE REDUCED AND THEREFORE PART OF THE ANALYSES WILL BECOME UNDERPOWERED MEANING THAT RELEVANT DIFFERENCES MIGHT NOT BE ASSOCIATED WITH STATISTICAL SIGNIFICANCES. THE OBSERVED DIFFERENCES PRESENTED CAN STILL BE ELUCIDATED AND INVESTIGATED IN FUTURE STUDIES.

THE SUB-ANALYSES ARE ADDRESSED AS EXPLORATORY IN THE LIMITATIONS.
WE DO NOT AGREE THERE IS ANY DISCONNECTION BETWEEN TEXT AND TABLES. TABLES 1 AND 2 (NOW 2 AND 3 AFTER CHANGES) ARE COMPLICATED TABLES. THE PREVENTION TEXT Follows TABLE 1 AND POINTS OUT THE RESULTS OF INTEREST. THE REVERSAL TEXT follows TABLE 2 AND POINTS OUT RESULTS OF INTEREST. THE TEXT IS ALL ABOUT WHAT IS IN THESE TABLES.

Page 7 lines 18 and 19 - It would be inappropriate to say that there was a greater proportion of infants with side preference as the p value was > 0.05. This needs to be reworded to reflect that there were not differences because they did not reach statistical significance. WE DELETED THE INFORMATION INSTEAD. IT DOES NOT NEED TO BE THERE.

Page 7 line 25. A more appropriate term than minimum - maximum is range. IN STATISTICS, THE RANGE OF A SET OF DATA IS THE DIFFERENCE BETWEEN THE LARGEST AND SMALLEST VALUES. THE MINIMUM- MAXIMUM VALUES ARE WHAT WE WANT TO PRESENT.

Page 7 line 31- please identify if this is the result of a Kappa analysis and the Kappa statistic of if it is just an agreement percent.

WE USED AC2 INSTEAD of KAPPA, SEE PAGE 6 LINE 14. WE HAVE ADDED A CLAUSE TO CLARIFY THIS.

Page 8 line 25 - a description of how prevention failure was categorized would be helpful here based on cases and non-cases.
THIS IS EXPLAINED IN THE BEGINNING OF THE SENTENCE – “THE OPPOSITE TO PREVENTION”. NOW WE HAVE ADDED A NEW CLAUSE TO CLARIFY FURTHER.

Page 8 line 29 - I am a bit concerned about analyzing the total NSF i.e. each group combined since there is a statement on page 8 line 42 that indicates that results were different among the groups which may indicate that the groups should not be combined and that the total analysis could attenuate the findings. There should be in the methods section or here the rationale for conducting a total of all three types analysis given the differences in trajectories found.

WE HAVE INCLUDED THE SENTENCE “WE ANALYZED GENERAL NSP AND THE THREE MAIN GROUPS – PLAGIOCEPHALY, BRACHYCEPHALY AND COMBINED PLAGIOCEPHALY-BRACHYCEPHALY IN TERMS OF PREVENTION FAILURE AND REVERSAL” IN THE METHODS SECTION PAGE 6. TOGETHER WITH THE DEFINITION IN THE BEGINNING OF THE INTRODUCTION, IT HELPS CLARIFY WE ARE INTERESTED IN BOTH GENERAL NSP AND EVEN THE 3 MAIN GROUPS. THE SENTENCE ON PAGE 8 LNE 42 IS MERELY SHOWING THAT BRACHYCEPHALY IS A PROBLEM IN BOTH GROUPS.

Page 8 line 33 and remainder of the paragraph - should be past tense: change "are" to "were".

YES. CORRECTED.

Page 8 line 34 - I am assuming that the statistic used was Chi square, however due to the small numbers a more appropriate statistic would be Fishers Exact test.

FISHER'S EXACT TEST ESSENTIALLY DOES NOT ALTER OUR CONCLUSIONS. FISHER'S Exact Test can be too conservative but in general our results remain intact regardless of test method.

Page 8 lines 35, 36 - The results are identified as lower but not statistically significant. The results should say there were no statistical differences not that the percent was lower in one group.

Here we are actually admitting that the intervention group performed somewhat worse than the control group. So, we believe the sentence should be kept.

Page 8 lines 42,34 - Were these results statistically significant?

These are descriptive statistics.

Page 8 lines 50,51 - Were the results statistically significant?

Descriptive statistics of interest to the reader.

Page 9 line 30-44 - there are many causal statements that are beyond the scope of the analysis used. At best associations were found.

Yes. We begin this paragraph saying factors that might explain reversal were investigated. We also admit that written information seemed to be more important than our intervention. We are careful not to draw conclusions on casuality by using the words might and seemed.
Page 9 line 31-32- The correct description of the OR should be that there were 4 x greater odds of NSP reversal from T1 to T2.

A STATISTICIAN HELPED ME WORD THIS SO THAT IT WOULD BE READER FRIENDLY. I WOULD PREFER TO KEEP THE SENTENCE AS IT IS.

Page 10 line 0 - another example of claiming differences when not statistically different.

WE ARE MERELY SHOWING TENDENCIES OF INTEREST.

Page 10 lines 4-16 - unclear if these results are statistically significant or not.

THESE ARE DESCRIPTIVE STATISTICS.

Page 10 lines 22, 23 - this statement requires rewording to reflect the analysis. Suggest something like: … the intervention was associated with reversal and reduced risk… The key point is association and not causation. Suggest the investigators read Bradford Hill and criteria for causation. One recent article describes them for modern times (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4589117/).

THANK YOU. THAT WAS HELPFUL.

Page 11 lines 6-11 - This analysis should be in the methods and results. It seems an afterthought which is should not have been.

TO PRESENT AN ALTERNATIVE RESULT IN THE DISCUSSION SECTION IS PERHAPS NOT THE MOST COMMON THING. WHILE CONSIDERING THE OUTCOME OF THIS STUDY WE WERE TRYING TO FIGURE OUT WHAT WENT WRONG IN THE INTERVENTION. SINCE THIS RESULT DOES NOT BELONG IN OUR RESULTS
SECTION WE FOUND IT APPROPRIATE TO INCLUDE IT IN THE DISCUSSION SINCE IT SHEDS LIGHT ON ONE OF OUR STRATEGIES THAT MIGHT NEED IMPROVEMENT – THE NURSES’ ABILITY TO DETECT NSP IN EARLY INFANCY.

Conclusion

The results are overstated see above comments. The conclusion has much information about future research needed with which I agree. I would have liked to have seen a "So What" statement as in examining graphs and data by 12 months there was little difference between the control and intervention group which might be explained by the brachycephaly persistence. However, if there was little difference at 12 months is the intervention truly needed. How many babies would be helped and what are the impacts of a misshapen head at 12 months.

GOOD POINTS. WE CORRECTED BY USING ASSOCIATION AS YOU SUGGEST ABOVE. THE INTERVENTION DID NOT HAVE ANY MAJOR EFFECT, SO WE DUG INTO THE DATA AND ANALYZED CAREFULLY TO FIND WHAT DID WORK. AND, IT RAISED THE QUESTION “WHAT MORE CAN WE DO TO IMPROVE OUR STRATEGIES?”.

Tables 1 and 2

Requires more in the legend including as explanation of case and non-case and intervention failure.

INCLUDED NOW.

Unclear why the RR is reported since it is not mentioned in the methods or results but OR is identified in the text.

WE DO HAVE RR MENTIONED IN THE STATISTICAL METHODS SECTION. IN THE TABLES WE ARE LOOKING AT RISK RATIOS. THE ORS PERTAIN TO REGRESSION ANALYSIS.