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.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The prevalence of suspected child abuse in children with constipation: a case-control study
AUTHORS	Vriesman, Mana Vrolijk- Bosschaart, Thekla F. Lindauer, Ramón J.L. van der Lee, Johanna H. Brilleslijper-Kater, Sonja Teeuw, Rian Benninga, Marc A.

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Dr. Peter Flom Institution and Country: Peter Flom Consulting Competing interests: None
REVIEW RETURNED	05-Nov-2021

GENERAL COMMENTS	I confine my remarks to statistical aspects of this paper.Unfortunately, there are some fairly major issues that need to be addressed before I can recommend publication. General questions: Was any matching done between the two groups? How many parents and or children refused to participate? Were they different from those who did participate on the information that you have for them? The outcome variable cannot be "prevalence of suspected CAN" in the groups. It has to be constipation. And, then the proper analysis depends on how FC is scored. Is it yes/no? Then logistic reg. is appropriate. But if it is ordinal or categorical, you might need ordinal logistic or multinomial logistic, or something else. The variables listed on lines 206-209 need to be operationalized, not just listed. Age and gender are pretty clear (although maybe gender needs some clarification) but the others are not. Comparing the groups on single variables is OK, but there really should be a logistic regression, so that the variables can be controlled for. In particular, variables about the parents' education and history need to be controlled for. Line 83 - unless you also give an estimate of abuse in the people without FC, this number doesn't mean much. Also, the huge range is worrying. I'm guessing different articles gave different numbers. This sentence might be expanded or contracted, but as is, it's not helpful. Peter Flom

REVIEWER	Reviewer name: Dr. Elise M. van de Putte Institution and Country: Univ Med Ctr Utrecht, Pediatrics Competing interests: None
REVIEW RETURNED	10-Dec-2021

GENERAL COMMENTS	This is an important study because there is a need for more substantiation or disconfirmation of the notion that functional complaints such as constipation may be related to forms of child abuse. It is a hard topic to investigate because of the methodological problems associated with the subject at the time of conducting a classic case-control study, such as the present study. The methodological problems are: inclusion of sufficient patients; medical ethical consent for the performance of the necessary diagnostic investigations; classification problems (child abuse or not) due to the fact that child abuse is difficult to classify so there is a high risk of misclassification; strong selection bias in inclusion. The authors have made an attempt to overcome the methodological problems as much as possible and correctly reflect the limitations of their study. I suspect that it is not possible to answer this research question in a better way than the researchers have done and so, despite its limitations, would be inclined to recommend publication. Moreover, they rightly indicate that a prospective cohort study would be the only way to answer the research question and the question is whether such a huge investment is in proportion to the relevance of the research question. I have a few more comments that could improve the manuscript: Abstract: it is strange to get results in the results section in the abstract that are not part of the research question. I would omit the section on parents with lower education from the abstract. Introduction r 98: 'after CAN'; CAN is usually not limited to one moment. I would rather say 'related to CAN' Methods Please add the fille number of the Medical Ethical Review Board R 130: exclusive: any sign of organic or functional disorder. Were psychiatric patients included? If not, please rephrase this sentence: in 'any disorder' R 141: 'recorded the findings during the physical examination'. I do not read back in the manuscript what symptoms and behavioral characteristics were scored for. I do read in Table 3 that there are no differences in symptoms and behavioral characteristics between the CAN + group and the CAN - group. This surprises me greatly because another article from the same research group showed that the physical examination and observed responses during physical examination can be diagnostic for the diagnosis of sexual abuse. So: either I expect extensive reflection on this or it is wiser to delete this methodological section and also remove these results from the table. It raises questions. R 168: please skip the sentences 166-171. These questionnaires were only applied in just a few children (8 years and older). The results are not specified in tables or text. So please skip these sentences. Prevalence of CAN R 176: please replace the word 'positive' in 'abnormal results'. Elsewhere in the manuscript the word 'positive' is being used. Please replace the word also in the other sections. R 186: please skip the last sentence. However …. Study. This sentence is not relevant (any more) for this study. No one expects you to report on this. The number of patients referred is reported

	elsewhere in the manuscript and that is sufficient. Patient an public involvement R 212: patients or 'parents of patients'? (the patients are very young…) RESULTS I would like to be informed about the agreement in deviant score between the 2 interviewers. This would inform me about the reproducibility of the chosen method. DISCUSSION R 266: the study of Tam: is this performed in a younger age group? That would be a good explanation for the lack of an association. R 313: 'our findings on physical examination and psychological symptoms should therefore be interpreted with care'. I don't understand this sentence. Why should we doubt about the findings during physical examination if this was performed before the interview? Table 3: Isn't it strange that the child characteristics between CAN + and CAN - do not differ? (Table 3). This still needs to be reflected upon in the discussion. For this indeed raises the question of whether the classification (CAN + or CAN -) has been correct. In general: the discussion could be more compactly written.

VERSION 1 – AUTHOR RESPONSE

Dear Dr. Shanti Raman, Prof. Imti Choonara, Dr. Flom, Dr. van de Putte and the editorial team,

Thank you for reviewing our paper and the opportunity to revise our manuscript. Below you will find the point-by-point response to the reviewers' comments. All changes in the manuscript are visible through 'track changes'.

Response to Editors:

Editor in Chief Comments to Author :
Discussion 2nd sentence delete "Surprisingly"
Answer the points raised by the reviewers who will see the revision
Response: Changed wording accordingly (line 255).

Associate Editor
Comments to the Author:
The reviewers have made useful suggestions and edits that would improve the quality of the final paper and I invite the authors to respond to these suggestions. I agree this is an important topic and not an easy one to investigate, and the authors have tried the best possible strategy.

Response to Reviewer #1:

Reviewer: 1
Dr. Peter Flom, Peter Flom Consulting
Comments to the Author
I confine my remarks to statistical aspects of this paper. Unfortunately, there are some fairly major issues that need to be addressed before I can recommend publication.

General questions:

R1C1 Was any matching done between the two groups?
Response: Our study protocol and recruitment methods were aimed at including a comparable group of children with different backgrounds and history. As described in the methods, all children presented at our specialised outpatient clinic during the time span of our study and were asked to participate in our study in the index group. This large city based clinic includes a wide variety of patients of different social-economic backgrounds and ethnicities, and includes both secondary and tertiary care for children with functional constipation (FC). For the healthy control group, we randomly selected 295 schools throughout different parts of the Netherlands, including both rural and urban areas. Per our study protocol, slow inclusion rates and our relatively small cohort, we did not perform any statistical matching between the two study groups.

As described in our methods and Table 1, demographic characteristics between the children with FC and healthy controls were compared and no statistical differences were found with respect to gender and age. Unfortunately, as mentioned in the discussion (lines 325-340), our study was at risk for selection bias, as we found important differences in demographic characteristics between the two study groups in terms of socio-economic status. Ideally, future studies should recruit a larger and more heterogeneous cohort of children in order to include a more representative patient and control group (including the possibility of statistical matching), see our discussion lines 371-373 “These observational studies…and control group”.

R1C2 How many parents and or children refused to participate? Were they different from those who did participate on the information that you have for them?
Response: As described in Figure 2, for the index group, 603 parents of children with FC were contacted of which 267 (44%) refused to participate. Unfortunately, the questionnaires/interview was only performed in children that agreed to participate and was not a standard procedure at the outpatient clinic. It is therefore impossible to compare the results/and or demographic information of both participating and non-participating children. For the healthy control group, 295 schools were contacted to recruit children without symptoms of FC. Of these 295 schools, 27 agreed to publish our study information letter in their school newsletter. According to our data, 5 children were excluded after initiation of the study due to refusal of parents (see Figure 2). However, since we contacted schools via the school newsletter and not on an individual basis, we do not have the exact number of children that were invited to participate for the control group and it is therefore impossible to compare the results of both participating and non-participating children.

R1C3 The outcome variable cannot be "prevalence of suspected CAN" in the groups. It has to be constipation. And, then the proper analysis depends on how FC is scored. Is it yes/no? Then logistic reg. is appropriate. But if it is ordinal or categorical, you might need ordinal logistic or multinomial logistic, or something else.
Response: We respectfully disagree with the reviewer on his suggestion on changing the main outcome variable, mainly because of our study design was developed together with our local Clinical Research Office. If this study included a cohort of children with a known history of CAN (child abuse and neglect), then it would indeed be appropriate to use constipation as the main outcome measure and calculate the difference on how many children had signs of FC and not. However, due to ethical objections and only small known cohorts of children with abuse, we performed a different strategy to determine a possible association between FC and CAN. In our study design, vice versa, we determined the prevalence of (suspected) abuse in a cohort of children with FC as compared to a cohort of healthy children. This way, the prevalence of suspected CAN as the main outcome variable seems to be in place.

R1C4 The variables listed on lines 206-209 need to be operationalized, not just listed. Age and gender are pretty clear (although maybe gender needs some clarification) but the others are not.
Response: To further clarify the statistical test that was performed per listed variable, we added this information to the Table 1. Please see our tracked changes to Table 1.

R1C5 Comparing the groups on single variables is OK, but there really should be a logistic regression, so that the variables can be controlled for. In particular, variables about the parents' education and history need to be controlled for.
Response: As the reviewer suggests, we did find parental variables (i.e. history of CAN and substance abuse) in children with a suspicion of CAN. However, when interpreting the results on parental variables in children with suspected CAN, it is important to note that our data was at risk for selection bias since important differences in demographic characteristics were found between the two study groups with respect to socio-economic status. As previously described in the literature, social factors such as socio-economic status are correlated with the risk of CAN [1]. Therefore, after consultation with our Clinical Research Office, our results with respect to parental variables should be interpreted with care and detailed elaboration including a logistic regression seems to be not in place (discussion, lines 336-339 “Owing to these differences…with suspected CAN)”.

[1]: World Health Organisation. Child Maltreatment Fact sheet No 150. https://www.who.int/news-room/fact-sheets/detail/child-maltreatment.

R1C6 Line 83 - unless you also give an estimate of abuse in the people without FC, this number doesn't mean much. Also, the huge range is worrying. I'm guessing different articles gave different numbers. This sentence might be expanded or contracted, but as is, it's not helpful.
Response: As the reviewer suggests, there is a big range of previously reported prevalence of abuse between studies in adults with functional defecation disorders. This highlights the methodological challenges on determining the prevalence of CAN. However, we agree with the reviewer and contracted the sentence (see changes line 83).

Response to Reviewer #1:
Reviewer: 2
Dr. Elise M. van de Putte, Univ Med Ctr Utrecht Comments to the Author This is an important study because there is a need for more substantiation or disconfirmation of the notion that functional complaints such as constipation may be related to forms of child abuse.
It is a hard topic to investigate because of the methodological problems associated with the subject at the time of conducting a classic case-control study, such as the present study.
The methodological problems are: inclusion of sufficient patients; medical ethical consent for the performance of the necessary diagnostic investigations; classification problems (child abuse or not) due to the fact that child abuse is difficult to classify so there is a high risk of misclassification; strong selection bias in inclusion.
The authors have made an attempt to overcome the methodological problems as much as possible and correctly reflect the limitations of their study.
I suspect that it is not possible to answer this research question in a better way than the researchers have done and so, despite its limitations, would be inclined to recommend publication. Moreover, they rightly indicate that a prospective cohort study would be the only way to answer the research question and the question is whether such a huge investment is in proportion to the relevance of the research question.
I have a few more comments that could improve the manuscript:

R2C1 Abstract: it is strange to get results in the results section in the abstract that are not part of the research question. I would omit the section on parents with lower education from the abstract.
Response: We agree with the reviewer and deleted the sentence on education level from the abstract (see changes lines 43-45).

R2C2 Introduction: r 98: 'after CAN'; CAN is usually not limited to one moment. I would rather say 'related to CAN'
Response: We concur with the reviewer and changed the wording accordingly (line 98).

R2C3 Methods: Please add the file number of the Medical Ethical Review Board
Response: The file number of the Medical Ethical Committee is now added (line 119).

R2C4 R 130: exclusive: any sign of organic or functional disorder. Were psychiatric patients included? If not, please rephrase this sentence: in 'any disorder'
Response: As shown in Table 1, behavioural comorbidity was not excluded and present in 17 children with FC, including ADHD and autism. To further clarify this, we rephrased line 130 accordingly.

R2C5 R 141: 'recorded the findings during the physical examination'. I do not read back in the manuscript what symptoms and behavioral characteristics were scored for. I do read in Table 3 that there are no differences in symptoms and behavioral characteristics between the CAN + group and the CAN - group. This surprises me greatly because another article from the same research group showed that the physical examination and observed responses during physical examination can be diagnostic for the diagnosis of sexual abuse. So: either I expect extensive reflection on this or it is wiser to delete this methodological section and also remove these results from the table. It raises questions.
Response: For this study, we used a combination of instruments providing information on both age-inappropriate sexual knowledge (including the SKPI) and sexual behaviour (including the CSBI) to encounter children with suspected sexual abuse. As previously described, age-inappropriate sexual behaviour is considered the most specific marker of sexual abuse in children and is observed in about one third of children who have been sexually abused [1, 2]. Other data of our research group supports the use of sexual knowledge as an important marker discriminating young children with and without a history of sexual abuse [3-5].

As described in the methods, we also included findings of the CARE-NL questionnaire, on which the physician described his findings on abnormal behaviour and findings during routine physical examination of children with FC during the clinic visit. Please note that this was not part of the interview by our trained research members and simply included findings of the physician. According to the physician, children with FC showed abnormal behaviour during examination (n=34), refusal of examination (n=17), hematomas (n=10), anal fissures (n=8), abnormal interaction between parent and child (n=5), see Tables 1 and 3. No differences with respect to abnormal behaviour and symptoms during physical examination were found between children with FC with and without a suspicion of CAN (Table 3). However, due to ethical objections, no physical examination was performed in the healthy control children (lines 152-153). We were therefore unable to include the child's behaviour during physical examination as a sign of suspected CAN. Differences in observers (trained vs. not-trained physician) and challenges of diagnosing CAN in children with FC (see discussion lines 306-315) may have contributed to our results in comparison with findings of our previous research studies [3-5].

In agreement with the reviewer, we understand that without this reflection these results raises questions. Also in accordance with comment R2C14, we deleted these results from the manuscript. Please see our changes to Table 3.

[1]: Kendall-Tackett KE, Williams LM, Finkelhor D. The impact of sexual abuse on children: A review and synthesis of recent empirical studies. Psychological Bulletin. 1993;113:164–180.
[2]: Friedrich WN, Fisher J, Dittner C, et al. Child Sexual Behavior Inventory: Normative, psychiatric and sexual abuse comparisons. Child Maltreatment. 2001;6:37–49.
[3]: Vrolijk-Bosschaart TF, Verlinden E, Langendam MW, et al. The Diagnostic Utility of the Child Sexual Behavior Inventory for Sexual Abuse: A Systematic Review. J Child Sex Abus. 2018;27(7):729-751.
[4]: Brilleslijper-Kater SN, Friedrich WN, Corwin DL. Sexual knowledge and emotional reaction as indicators of sexual abuse in young children: theory and research challenges. Child Abuse Negl. 2004;28(10):1007-1017. doi:10.1016/j.chiabu.2004.06.005
[5]: Brilleslijper-Kater SN. Beyond Words : Between-Group Differences in the Ways Sexually Abused and Nonabused Preschool Children Reveal Sexual Knowledge. Enschede: Febodruk; 2005.

R2C6 R 168: please skip the sentences 166-171. These questionnaires were only applied in just a few children (8 years and older). The results are not specified in tables or text. So please skip these sentences.
Response: We respectfully disagree with the reviewer on this comment. According to our study protocol, our secondary objective was to determine clinical characteristics and psychological symptoms of children with suspected CAN. For this secondary objective, we used the RCADS-P/ RCADS and TSCYC/TSCC questionnaires.
All tools were used for the whole age range. However, due to the age-restrictions of specific instruments used in this study, some were used for parents of included children while others were also used for child-report. As described in the Appendix and shown in Figure 1, the TSCC and RCADS were only used in children ≥8 years old. Whereas the TSCYC and RCADS-P were used in children<8 years of age. The results of this secondary outcome are mentioned in the results, lines 247 -250 “According to the RCADS/RCADS-P…suspicion of CAN (n=282)” and in Table 1. To avoid confusion, we deleted the sentence “The TSCC and RCADS were only suitable for children >8years of age (line 168-169)”.

R2C7 Prevalence of CAN
R 176: please replace the word ‘positive’ in ‘abnormal results’. Elsewhere in the manuscript the word ‘positive’ is being used. Please replace the word also in the other sections.
Response: Changed wording accordingly (see lines 176, 180, 181, 348, Table 2) and clarified this in the methods (lines 179-180 “The prevalence of…the following abnormal results.”)

R2C8 R 186: please skip the last sentence. However …. Study. This sentence is not relevant (any more) for this study. No one expects you to report on this. The number of patients referred is reported elsewhere in the manuscript and that is sufficient.
Response: We deleted the sentences from the methods, this is now only reported in the discussion (lines 277-279).

R2C9 Patient an public involvement
R 212: patients or ‘parents of patients’? (the patients are very young…)
Response: The reviewer is correct, this was discussed with parents of included children. See changes to line 214.

R2C10 RESULTS

I would like to be informed about the agreement in deviant score between the 2 interviewers. This would inform me about the reproducibility of the chosen method.
Response: As described in the methods, the abnormal reactions to the SKPI were scored by two independent observers and discussed with a blinded and specialised member of our CAN team (SB) (lines 177-178, 184). This member is highly specialised in the use of the SKPI and published several papers on its use in clinical practice [4,5]. The input of this member was therefore leading and only abnormal results as scored by the two interviewers was used to minimize bias (see lines 349-352).
As scored by the first interviewer, a total of 19/381 included children had abnormal reaction to the SKPI (5.0%). As scored by the second interviewer, a total of 32/381 children had an abnormal reaction to the SKPI (8.4%). After consultation with the blinded member of our research team, a total of 15 children were scored as having abnormal reaction to the SKPI. Of these, 13 children were included in the index group and 2 children in the control group (see Table 2).
However, because this information on agreement of the SKPI is not part of our research question we feel that these results are beyond the scope of our manuscript. Further research on the validation and more information on the reproducibility of the SKPI is now in progress and will be published separately [6].

[6]: van Ham K, Brilleslijper-Kater S, van der Lee H, et al. Validation of the Sexual Knowledge Picture Instrument as a diagnostic instrument for child sexual abuse: study protocol. BMJ Paediatrics Open 2020;4:e000799. doi:10.1136/bmjpo-2020-00799

R2C11 DISCUSSION
R 266: the study of Tam: is this performed in a younger age group? That would be a good explanation for the lack of an association.
Response: The study of Tam et al. [7] was performed in 282 children with FC with a mean age of 9 years. The study was performed in Hong-Kong and used a parental survey to diagnose CAN. Therefore, this similar age group is not likely to explain the difference in prevalence of abuse in contrast with our study. As described in the discussion (lines 268-277), however, we hypothesise that differences in social factors and diagnostic instruments could have contributed to the different outcomes of studies.

[7] Tam YH, Li AM, So HK, et al. Socioenvironmental factors associated with constipation in Hong Kong children and Rome III criteria. J Pediatr Gastroenterol Nutr. 2012;55(1):56-61.

R2C12 R 313: 'our findings on physical examination and psychological symptoms should therefore be interpreted with care'. I don't understand this sentence. Why should we doubt about the findings during physical examination if this was performed before the interview?
Response: As hypothesised in the discussion (lines 306-315), interpretation of findings during physical examination in children with FC can be challenging as symptoms of constipation and CAN might overlap. Physical examination including a rectal exam in children with FC might evoke feelings of shame and fear and could lead to more knowledge of the human body. Therefore, because the interview was performed right after the outpatient visit, signs of abnormal behaviour and/or sexual knowledge during the iterview should be interpreted with care. Objective findings, such as anal fissures or fear of rectal examination, are of course not influenced by the interview, however, should be interpreted with care and not directly be attributed as signs of (sexual) abuse. This is now clarified in the discussion (lines 310, 314-315).

R2C13 Table 3: Isn't it strange that the child characteristics between CAN + and CAN - do not differ? (Table 3). This still needs to be reflected upon in the discussion. For this indeed raises the question of whether the classification (CAN + or CAN -) has been correct.

Response: According to Table 3, we compared child characteristics including gender, age and behavioural problems between children with and without CAN. In contrast with what the reviewer suggests, age was statistically different between the two groups; children with a suspicion of CAN were significantly older as compared to children without a suspicion of CAN. This is in line with the current literature, as it could be hypothesised that younger children have more difficulties to find the proper wording and knowledge to report abuse [4], possibly resulting in a lower prevalence (or recognition) in this age group.
With respect to gender, we found no difference between children with and without a suspicion of CAN. For this analysis, we did not look into the subtypes of CAN. This is also in line with previous literature, stating that rates of abuse are similar for both sexes in high-income countries [8].
Due to the small number of children with behavioural problems in our cohort, no strong assumptions can be made from our data. Therefore, because our results concur with the previous literature and the methodological challenges, these results should be interpreted with care and detailed elaboration in the discussion seems to be not in place.

[8] Gilbert R, Widom CS, Browne K. Burden and consequences of child maltreatment in high-income countries. Lancet 2009;373:68-81.

R2C14 In general: the discussion could be more compactly written.
Response: Please see our tracked changes throughout the discussion.