

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)	Risk factors for misdiagnosis in children with developmental dysplasia of the hip- a retrospective single centre study
AUTHORS	Fan, Zong-Zhi Yan, Ya-Bo Sha, Jia Xu, Hui-Fa Li, Chao Liu, Zhi-Chen Liu, Jing Huang, Lu-Yu

VERSION 1 - REVIEW

REVIEWER	Reviewer Name: Dr. Peter Flom Institution and Country: Peter Flom Consulting, US
REVIEW RETURNED	25-Feb-2023

GENERAL COMMENTS	<p>I mostly confine my remarks to statistical and methodological aspects of this paper. The general approach (logistic reg.) is correct, but I have one relatively major issue and a few minor ones that need to be addressed before I can recommend publication.</p> <p>The huge range of DDH across countries (page 3, line 48 to 50) makes me want to know why it varies so much. Is this a variation in false negatives?</p> <p>p. 6 line 33-43. This is known as bivariate screening and it is not a good method. All the results of the multiple regression will be incorrect. P values will be too low, standard errors too small, and parameter estimates biased away from 0. For details, examples, and proofs, see Harrell, Regression Modeling Strategies.</p> <p>It is best to use substantive knowledge to build the model. If the authors insist on using an automatic method, then LOESS is not too bad.</p> <p>Various places: "Univariate logistic regression" should either be "bivariate" or "simple" logistic regression. (Bivariate logistic regression, technically, means two dependent variables).</p> <p>Also, wherever a p value is given (even if not significant) the effect size (odds ratio or whatever) should also be given.</p> <p>Table 1 Age should not be categorized. Doing so increases type 1 and type 2 error. Instead, leave age as years (or maybe months) and use splines to investigate nonlinearity.</p> <p>Figure 2 is fine, but it is very small (maybe this is a publication</p>
-------------------------	--

	issue). Could it use more of the page? Maybe use landscape instead of portrait orientation.
--	---

REVIEWER	Reviewer Name: Dr. Justine St George Institution and Country: The Children's Hospital at Westmead, Australia
REVIEW RETURNED	07-Mar-2023

GENERAL COMMENTS	<p>For better clarity, the institutional screening structure could be clearer.</p> <p>The biggest bias of this paper is that the population are already a select population - i.e. referred to a tertiary centre for DDH. Am I right in understanding that this is a major underrepresentation of the missed DDH which remains in the community?</p> <p>pg 9 line 43 - increases the risk of a poor prognosis pg 12 line 22 - "Shenton's" pg 12 line 22 - this is a paediatric journal. The importance of x-ray over USS over the age of 6 months should be highlighted further pg 12 line 36 - in fact, a missed diagnosis of DDH has a greater association with NO risk factors. Patients with risk factors should not be "slipping through the cracks" but those without still are a major contribution to those with a late diagnosis (hence the argument for universal ultrasound screening)</p> <p>The physician performing the clinical screening examination varies by country. Is it fair to say that further training in clinical screening could potentially reduce the discrepancy between physician type?</p>
-------------------------	--

VERSION 1 – AUTHOR RESPONSE

Dear editors and reviewers,

We thank you very much for giving us an opportunity to revise our manuscript, and thank you for the editors and reviewers' comments concerning our manuscript entitled "Risk Factors for Misdiagnosis at the First Visit of DDH in Children without Ultrasonographic Hip Screening". These constructive comments are all valuable and very helpful for revising and improving our work. We have studied the comments carefully and have made corrections which we hope meet with approval. Revised portions are marked in different colors in the revised manuscript. The main corrections in the paper and the responses to the reviewers' comments are as follows.

Responds to the reviewer's comments:

To editor in Chief:

Comment 1: Title amend to "Risk Factors for Misdiagnosis in Children with developmental dysplasia of the hip- a retrospective single centre study"

Response: Thank you for your valuable advice. The amended title is more appropriate. According to your suggestion, we have amended the title to: Risk Factors for Misdiagnosis in Children with developmental dysplasia of the hip- a retrospective single centre study.

Comment 2: What this study adds section. 1st sentence amend to "One in four children with DDH without hip ultrasound screening were misdiagnosed at their first visit".

Response: Thank you for your valuable advice. We have revised it according to your suggestion.

Comment 3: Introduction. Last paragraph: delete the sentence " To our knowledge, the misdiagnosis of children with DDH at the first visit has not been reported." Journal policy is for authors NOT to describe their study as the first.

Response: Thank you for your correction. This was our mistake and we have deleted this sentence in the revised version.

Comment 4: % use one decimal point only.

Response: Thank you for your correction, we have kept all % only one decimal point in the revised version.

Comment 5: Discussion avoid P values. They should have been presented in the results section.

Response: Thank you for your correction, We have removed all the P values from the section in the discussion.

Special thanks to you for your good comments and suggestions.

To: Associate Editor

Comment 1: The points raised will need to be fully responded to re. the statistical approaches and the clinical comments, especially about understanding the pathway for referrals, etc, in this healthcare system.

Response: Thank you for your reminder. We have recalculated some statistical data based on the comments of reviewer 1, but the main results and conclusions have not been affected. We have also provided detailed explanations on issues related to clinical comments in our responses to reviewer 2.

Comment 2: Please also proof read carefully any revised version prior to submission.

Response: Thank you for your reminder. We have checked the revised manuscript carefully before submission.

Special thanks to you for your good suggestions and reminder.

To Reviewer 1:

Comment 1: The huge range of DDH across countries (page 3, line 48 to 50) makes me want to know why it varies so much. Is this a variation in false negatives?

Response: Thanks for your kind question. Please find the following detailed answers to your question.

It is difficult to estimate the incidence of DDH.¹ Data ranging from 1‰ to 34‰ are widely used in literature.^{2, 3} The incidence may be affected by the differences in the definition and diagnosis of the disease in different regions. Therefore, the possibility of some false negative variants cannot be excluded. However, studies have shown significant differences in DDH incidence between ethnic groups and geographic locations. Some studies show a wider range of fluctuations, ranging from an incidence rate per 1000 live births of 0.06 in Africa to 76.1 in Native Americans.⁴ Therefore, we considered that differences in DDH incidence may be due to differences in ethnicity and geographic location rather than false negative results.

References:

1 Nguyen JC, Dorfman SR, Rigsby CK, Iyer RS, Alazraki AL, Anupindi SA, et al.. ACR Appropriateness Criteria® Developmental Dysplasia of the Hip-Child. J AM COLL RADIOL. 2019; 16: S94-103.

2 Tian F, Zhao D, Wang W, Guo L, Tian S, Feng A, et al.. Prevalence of Developmental Dysplasia of the Hip in Chinese Adults. CHINESE MED J-PEKING. 2017; 130: 1261-68 PubMed .

3 Garcia S, Demetri L, Starcevich A, Gatto A, Swarup I. Developmental Dysplasia of the Hip: Controversies in Management. CURR REV MUSCULOSKE. 2022; 15: 272-82 PubMed .

4 O Beirne JG, Chlapoutakis K, Alshryda S, Aydingoz U, Baumann T, Casini C, et al.. International Interdisciplinary Consensus Meeting on the Evaluation of Developmental Dysplasia of the Hip. Ultraschall in der Medizin - European Journal of Ultrasound. 2019; 40: 454-64.

Comment 2: p. 6 line 33-43. This is known as bivariate screening and it is not a good method. All the results of the multiple regression will be incorrect. P values will be too low, standard errors too small, and parameter estimates biased away from 0. For details, examples, and proofs, see Harrell, Regression Modeling Strategies. It is best to use substantive knowledge to build the model. If the authors insist on using an automatic method, then LOESS is not too bad. Various places: "Univariate logistic regression" should either be "bivariate" or "simple" logistic regression. (Bivariate logistic regression, technically, means two dependent variables).

Response: Thank you for your valuable advice. Our previous statistical methods did have some shortcomings. According to your suggestions, we have made some changes in grouping and statistical methods.

First, we would like to make it clear that there was only one dependent variable in our study and that it was dichotomous. Because the examination methods used to diagnose DDH differ between children younger than 6 months and those older than 6 months, there are some differences in the classifications used to assess the severity of DDH. We previously performed a univariate stratified analysis of these two age groups, which showed a P value of more than 0.1 in both groups. Therefore, we did not include these items in the multivariate logistic regression analysis. Although there was no impact on the results of the multivariate logistic regression statistics, it was incorrect that we only stratified the single factors. In the revised version, we unified the two methods to assess the severity of DDH, no longer stratified by age, and reperformed the univariate analysis. Finally, all variables with a P value of < 0.1 were subsequently included in a nonautomatic multivariable logistic regression stepwise model. We have described the modified statistical methods in detail in the article. (Methods: Statistical Analyses)

Because we are not professionals in statistics, we referred to several similar studies that also applied the same statistical methods. 1-4 Therefore, we believe that although there is the possibility of error in binary logistic regression, the conclusions obtained can preliminarily meet the purpose of our research. If there is still something wrong with the statistical method, we hope to get further guidance from you and give us a chance to correct it.

References:

1 Studer K, Williams N, Antoniou G, Gibson C, Scott H, Scheil WK, et al.. Increase in late diagnosed developmental dysplasia of the hip in South Australia: risk factors, proposed solutions. MED J AUSTRALIA. 2016; 204: 240.

2 Woh PY, Yeung MPS, Nelson EAS, Goggins WBI. Risk factors of non-typhoidal Salmonella gastroenteritis in hospitalised young children: a case-control study. BMJ PAEDIATR OPEN. 2021; 5: e898.

3 Olabarri M, Vazquez P, Gonzalez-Posada A, Sanz N, Gonzalez-Peris S, Diez N, et al.. Risk Factors for Severe Anaphylaxis in Children. *The Journal of Pediatrics*. 2020; 225: 193-97 PubMed .

4 Hum S, Liu H, Shaikh N. Risk Factors for the Development of Febrile Recurrences in Children with a History of Urinary Tract Infection. *The Journal of Pediatrics*. 2022; 243: 152-57.

Comment 3: wherever a p value is given (even if not significant) the effect size (odds ratio or whatever) should also be given.

Response: Thanks for your kind suggestion, which is valuable for improving the

manuscript's accuracy. We have modified it according to your suggestion.

Comment 4: Table 1 Age should not be categorized. Doing so increases type 1 and type 2 error. Instead, leave age as years (or maybe months) and use splines to investigate nonlinearity.

Response: Thanks for your valuable advice, which is valuable for improving the

manuscript's accuracy. In the revised manuscript, we no longer categorized age and leave age as months. Because age was not normally distributed, we used the Mann-Whitney U test for comparison.

Comment 5: Figure 2 is fine, but it is very small (maybe this is a publication issue). Could it use more of the page? Maybe use landscape instead of portrait orientation.

Response: Thank you for your kind suggestion. According to your suggestion, we have enlarged picture 2, and if necessary, we can further modify it to the picture of landscape direction.

Special thanks to you for your good comments and suggestions.

To Reviewer 2:

Comment 1: For better clarity, the institutional screening structure could be clearer.

Response: Thank you for your kind suggestions. Since no effective DDH screening structure has been established in this region, we have made some modifications to the discussion section and made some suggestions for the screening structure of DDH.

Since the infant physical examination is mainly performed by community physicians, childcare physicians, and general practitioners. Therefore, it is necessary to strengthen the training of these physicians in infant hip physical examination, so as to strengthen DDH screening in routine physical examinations. Of course, some studies have shown that physical examination alone may still lead to a high risk of misdiagnosis.⁹ Therefore, selective or universal hip ultrasound screening should be added to the infant physical examination program whenever possible in areas where it is available. In this way, more children with DDH can be diagnosed early and the incidence of misdiagnosis can be reduced. (Line 80-88, Discussion)

Comment 2: The biggest bias of this paper is that the population are already a select population - i.e. referred to a tertiary centre for DDH. Am I right in understanding that this is a major underrepresentation of the missed DDH which remains in the community?

Response: Thanks for your kind question. Please find the following detailed answers to your question.

Although children with DDH from one tertiary hospital were selected for this study, they were referred from hospitals at different levels. In northwest China, children with DDH have not been treated with grading treatment. Most of them were treated in tertiary hospitals, especially for children with DDH

who needed surgical treatment, because only part of tertiary hospitals could perform surgical treatment of DDH. Therefore, no matter which level of the hospital the children with DDH first visit, after diagnosis, most of them are treated in larger tertiary hospitals. As a large pediatric orthopedic center in northwest China, our hospital can accept children with DDH referred from hospitals at all levels. Therefore, we can calculate the misdiagnosis of DDH children in different levels of hospitals at the first visit. We believe that the small sample size of community hospitals is not due to the selected population. This is mostly due to the incomplete division of departments in community hospitals, and children often choose a superior hospital for their first visit. This is also a difficulty in the allocation of medical resources in our country: large hospitals are often full of patients, and the number of patients in community hospitals is less. Of course, the statistical results of a single center may have some bias, which we have explained in the limitations of the study in the discussion section of this article. (Line 89-94, Discussion)

Comment 3: pg 9 line 43 - increases the risk of a poor prognosis.

Response: Thank you for your kind suggestion. It will be more reasonable to modify it like this. We have modified it according to your suggestion.

Comment 4: pg 12 line 22 - "Shenton's".

Response: Thank you for your correction. This is our spelling error, which has been corrected in the revised manuscript.

Comment 5: pg 12 line 22 - this is a paediatric journal. The importance of x-ray over USS over the age of 6 months should be highlighted further.

Response: Thanks for your suggestion, we have further emphasized the importance of X-ray examination in children with DDH over 6 months of age in the corresponding position.

The ossified nucleus of the femoral head has appeared in most children over 4 to 6 months. Therefore, anteroposterior pelvic X-ray has become the first choice for diagnosing DDH in children over 4 to 6 months old. The diagnostic indicators include the acetabular index, center-edge angle, Perkin's quadrant, and Shenton's line.²⁴ Anteroposterior pelvic X-ray can be taken in almost all hospitals. It has a low radiation dose, low cost, and high diagnostic value. Therefore, we believe that it is necessary for physicians to perform an anteroposterior pelvic X-ray on children over 6 months who cannot completely exclude DDH. (Line 60-67, Discussion)

Comment 6: pg 12 line 36 - in fact, a missed diagnosis of DDH has a greater association with NO risk factors. Patients with risk factors should not be "slipping through the cracks" but those without still are a major contribution to those with a late diagnosis (hence the argument for universal ultrasound screening).

Response: Thank you for your suggestion. We agree with you very much, and we also think universal ultrasound screening is necessary, which has been highlighted in the discussion section of the article.

The results of the present study showed that the risk factors of DDH, such as sex, family history, parity, and fetal presentation, have no significant influence on the misdiagnosis of DDH. This is a mistake that should not be made.

Therefore, selective or universal hip ultrasound screening should be added to the infant physical examination program whenever possible in areas where it is available. In this way, more children with DDH can be diagnosed early and the incidence of misdiagnosis can be reduced. (Line 85-88, Discussion)

Comment 7: The physician performing the clinical screening examination varies by country. Is it fair to say that further training in clinical screening could potentially reduce the discrepancy between physician type?

Response: Thank you for your question. This is a mistake in our expression that has caused you to misunderstand. We have made some changes in the revised manuscript.

The routine infant physical examination has not played a due role in diagnosing DDH. Since the infant physical examination is mainly performed by community physicians, childcare physicians, and general practitioners. Therefore, it is necessary to strengthen the training of these physicians in infant hip physical examination, so as to strengthen DDH screening in routine physical examinations. (Line 79-83, Discussion)

Of course, training in clinical screening does not necessarily reduce differences among categories of physicians. We hope that through this study, doctors in relevant departments can pay attention to DDH. They can master the risk factors and diagnostic methods of DDH to reduce the occurrence of missed diagnoses.

Special thanks to you for your good comments and suggestions.

VERSION 2 – REVIEW

REVIEWER	Reviewer Name: Dr. Peter Flom Institution and Country: Peter Flom Consulting, USA
REVIEW RETURNED	10-Apr-2023

GENERAL COMMENTS	<p>The authors have responded very well to most of my comments. One issue remains. They continue to use bivariate screening to build the multiple logistic regression model. They note that they looked at other studies to see what they did.</p> <p>My response:</p> <p>It is certainly a good idea to look at other studies, and I will readily admit that bivariate screening is widely used. I see it used all the time in articles I review, for one thing.</p> <p>But it's still not a good method. As i noted in my first review, all the output will be incorrect (see that review for details). These issues apply to logistic regression, linear regression, and other forms of regression as well.</p> <p>I recommend two possible courses of action:</p> <p>1. (Preferable): Use substantive knowledge and interest to build the model. Include variables that others have found important, that you are particularly interested in, that make substantive sense, and so on. Try to include any covariates which readers would think it odd to omit. (I don't know what these are, but the authors probably do).</p> <p>Remember that it can be important to include variables that are small and NOT significant if they a) Act as good covariates - that is, including them changes the other parameter estimates. b) Are generally thought to be important. Showing that a relationship is weak is important if everyone thinks it is strong. (As an example, suppose you were studying the relationship of height and sex in humans and found some group where the difference was very small --- that would be a HUGE discovery! or c) Are part of your</p>
-------------------------	--

	hypotheses or research questions.
	Otherwise - very good revisions!

VERSION 2 – AUTHOR RESPONSE

Dear editors and reviewers,

We appreciate the opportunity to revise the manuscript again, and thank you for the editor's and reviewers' comments concerning our manuscript entitled "Risk factors for misdiagnosis in children with developmental dysplasia of the hip- a retrospective single centre study". We have studied the comments carefully and have made corrections which we hope meet with approval. Revised portions are marked in different colors in the revised manuscript. The main corrections in the paper and the responses to the reviewers' comments are as follows.

Responds to the reviewer's comments:

To: Editor in Chief

1. Comment 1: Formatting Amendments (where applicable):

Response: Thank you for your reminder. We reviewed the format of the article and found no obvious errors.

To: Associate Editor

1. Comment 1: See the further comments made re. statistical approach used and respond to these.

Response: Thank you for your reminder. We have rebuilt the multivariate model as suggested by reviewer 1. Although some of the statistical results changed, the final study conclusions did not change.

To Reviewer 1:

Comment 1: It is certainly a good idea to look at other studies, and I will readily admit that bivariate screening is widely used. I see it used all the time in articles I review, for one thing. But it's still not a good method. As I noted in my first review, all the output will be incorrect (see that review for details). These issues apply to logistic regression, linear regression, and other forms of regression as well.

I recommend two possible courses of action: 1. (Preferable): Use substantive knowledge and interest to build the model. Include variables that others have found important, that you are particularly interested in, that make substantive sense, and so on. Try to include any covariates which readers would think it odd to omit. (I don't know what these are, but the authors probably do).

Remember that it can be important to include variables that are small and NOT significant if they a) Act as good covariates - that is, including them changes the other parameter estimates. b) Are generally thought to be important. Showing that a relationship is weak is important if everyone thinks it is strong. (As an example, suppose you were studying the relationship of height and sex in humans and found some group where the difference was very small --- that would be a HUGE discovery! or c) Are part of your hypotheses or research questions.

Response: Thank you very much for your detailed guidance, which has given us a further understanding of statistical methods. In the revised manuscript, we have rebuilt the multivariate model.

Under your guidance, we studied the statistical methods and consulted a statistical expert. It is unreasonable to exclude variables from the multivariate model just because the P value of > 0.1 in the univariate analysis of this study. When reconstructing the multivariate model, we not only included variables with a value of < 0.1 in univariate analysis, but also included factors that were considered to have a significant impact on the diagnosis of DDH in previous studies (main symptom, participation in the infant physical examination or not, family history and severity of DDH). To find meaningful variables, we adjusted the reference variables for some multi-categorical variables. Although some of the statistical results changed, the final study conclusions did not change.

Special thanks to you for your good comments and suggestions, which have improved the accuracy of the article. If there are still shortcomings, we will make further corrections in time.

VERSION 3 – REVIEW

REVIEWER	Reviewer Name: Dr. Peter Flom Institution and Country: Peter Flom Consulting, USA
REVIEW RETURNED	30-Apr-2023

GENERAL COMMENTS	The authors have addressed my concerns and I now recommend publication
-------------------------	--

VERSION 3 – AUTHOR RESPONSE

N/A