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)	Anemia Prevalence and its Associated Factors in Children under 5 Years in Western China: A Systematic Review
AUTHORS	Zhou, Huan Du, Yefan Liao, Ying Leng, Fangqun Li, Linhua Ye, Ruixue Mao, Yuping Raat, Hein

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Dr. Mohammad Yawar Yakoob Institution and Country: Indus Hosp Res Ctr, United Kingdom of Great Britain and Northern Ireland Competing interests: None
REVIEW RETURNED	12-Oct-2021

GENERAL COMMENTS	A good study. Please find some comments below. Define cross-sectional studies whether these were descriptive or analytical i.e. how many reported only prevalences and how many reported associated factors analyses as well. If both then mention analytical. How are surveillance studies different from cross-sectional? I would suggest to use the word 'pooled' or systematic 'meta-analysis' review in the manuscript freely with overall reported prevalences to indicate meta-analyzed effect estimates. Associated factors are mentioned but what are the pooled ORs or RRs with 95% CI for example, if Tibetan ethnic group has a higher risk of anemia then how much fold or times increased risk and whether it is statistically significant or not? It is not clear what statistical software was used for analysis i.e. to pool the prevalences. Also, the authors mention about heterogeneity in the Abstract but I2 statistics and Chi-square P-values are not mentioned to distinguish between-study from within-study variation. I would also suggest to include meta-regression to adjust for confounding effects of different associated factors rather than univariate analyses and to justify study of effect modification only for significant factors like that done for age, gender and ethnicity with P-values for interaction. I would suggest to explore if data is available on use of iron supplements or prevalence of childhood helminthiasis in each of the regions to do some rough correlation for Discussion section even if that available from different other studies in published literature. In the PRISMA flow-diagram I would suggest to add labels in big

	boxes to the left of the figure to add headings of steps of review like screening, eligibility assessment, data extraction, full-text review, final inclusion and meta-analysis, etc. I would like to know what was the range of prevalences in studies which were excluded due to low quality or sample size less than 50 in the Results section to see if excluded or missing studies had different sizes compared to included studies.

REVIEWER	Reviewer name: Dr. Luis Rajmil Institution and Country: Homer 22 1rst 1, Barcelona, 08023, Spain Competing interests: None
REVIEW RETURNED	19-Oct-2021

GENERAL COMMENTS	The article presents the objectives, methods and results in a clera and acceptable way.

REVIEWER	Reviewer name: Dr. Sarah Nevitt Institution and Country: University of Liverpool, Biostatistics Competing interests: None
REVIEW RETURNED	27-Oct-2021

GENERAL COMMENTS	I have conducted a statistical review of the manuscript "Anemia prevalence and its Associated Factors in Children under 5 Years Old in Western China: A Scoping Review" The authors aim to conduct a scoping review of anemia prevalence and factors associated with anemia in Western China. The authors have clearly comprehensively searched the literature and extracted a lot of information from the included studies related to the objective. But I think that what the authors have actually done here is closer to a systematic review than a scoping review. Scoping reviews are generally conducted where the evidence base for a particular question are unclear and further information regarding types of available evidence, research methods used, definitions and concepts etc. related to a question are of interest. Scoping reviews may be conducted as a learning exercise prior to a systematic review, to inform the inclusion criteria, outcomes etc. of a systematic review. Scoping reviews are generally not suitable for synthesising numerical data and/or providing effect sizes. The authors seem to justify the scoping review approach because previous systematic reviews have shown heterogeneity between studies. This is not a necessarily a reason to perform a scoping review rather than a systematic review, but it would be a reason to perform a systematic review without meta-analysis. If the authors wished to gain further insight about the design, methods used, factors examined etc. in the prevalence studies a scoping review would be suitable. However, as the aim of the authors is to provide estimates (i.e. medians and ranges) of prevalence rates across different regions, to compare subgroups and to identify factors significantly associated with prevalence rates, a systematic review would be more suitable. Please see the following reference for further details on the differences between systematic reviews and scoping reviews: Munn, Z., Peters, M.D.J., Stern, C. et al. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 18,

	143 (2018). https://doi.org/10.1186/s12874-018-0611-x I suggest that the authors start again and conduct a systematic review. I have a few other specific comments on the approach and presentations of results: 1) Throughout the manuscript the term 'child under the age of five' is used. Please replace with 'children under the age of five.' 2) Methods: Please note that PRISMA (for systematic reviews or for scoping reviews) are reporting guidelines rather than conducting guidelines or methodological framework. 3) Methods: The search strategy, screening process for the studies and quality assessment are described in detail but more information is needed in the methods section of a review (whether a scoping review or a systematic review). Further details are needed of the data which will be extracted, how this data will be summarised (including any effect size used – e.g. median and range of prevalence) and how data will be synthesised and presented (in other words, whether a quantitative synthesis is planned (i.e. meta-analysis) or a narrative synthesis in figures and tables). 4) Inclusion of studies with sample size of over 50: While this may be a reasonable inclusion criterion, I'm not sure about this reference (i.e. a systematic review of the influence of grandmothers on breastfeeding rates). Does this reference demonstrate that studies with sample size <50 may not have adequate statistical power or is this simply also an inclusion criterion of this review? If the latter, I suggest that a methodological reference should be cited instead. 5) Quality assessment: "Each of the item was identified by a score from 0 (unclear) to 2 (yes)," I assume that this means unclear = 0, no = 1 and yes = 2? Given that the questions of the JBI checklist consider whether methods used are appropriate or adequate where the wording of the questions means that yes = appropriate methods, no=inappropriate methods and unclear = insufficient information, to assign a higher score to 'no' (where methods are clearly inappropriate) than to 'unclear' (where there is uncertainty whether methods are appropriate) doesn't make sense. The JBI checklist doesn't assign scores, rather the reviewers should decide whether to include or exclude the studies based on the responses to the 9 questions, and where studies are excluded, reasons should be provided for this. I suggest that the tools should be used as intended rather than assigning scores. 6) Table 1: 58 studies are included in the review but only 51 studies are summarised in Table 1. What sort of information did the 7 studies not included in Table 1 provide which were of relevance to the review? 7) Figure 3: I think I understand what the authors are trying to show here but there are a few issues with this plot. Firstly, the plot does not show 'time of each study' but shows estimates across different regions over time. By showing summary estimates across regions (presumably the median values), this does not capture the uncertainty and variability across studies. I assume that studies report only prevalence estimates for specific years or time intervals, rather than specific information about the rate of change in prevalence over time. By joining up the point estimates at different time points with straight lines, this implies a linear change (increase or decrease) in prevalence over time. Commenting on differences in prevalence estimates in different regions in different years is fine (ideally with uncertainty in the estimates captured), but the authors should avoid making

	assumptions about how prevalence changes over time without details from the studies to inform this.

REVIEWER	Reviewer name: Dr. Peter Rohloff Institution and Country: 2 Calle 5-43 Zona 1, Santiago Sacatepéquez, 3006, Guatemala Competing interests: None
REVIEW RETURNED	03-Nov-2021

GENERAL COMMENTS	This is a remarkably interesting study, which does a great service by comprehensively reviewing Chinese language research literature and medical thesis and getting this data into an accessible format for the global pediatric community. I do have a few comments. Overall English could be improved - a close edit is needed to improve readability Methods -Overall I am mostly wondering why the authors haven't conducted a prevalence metaanalysis. Most of the studies are cross sectional, quality assessment has been done, and the search strategy is systematic. At least for these cross sectional studies, it seems like it would be logical to do the metaanalysis This seems to me the biggest limitation, and the publication would be higher impact if this was done. -Exclusion criteria: excluding studies with samples less than 50: again this seems to set up the metaanalysis as the authors talk about only including studies with adequate statistical power -Exclusion criteria need to be explicitly detailed (it seems like only examples are given, e.g prematurity and aplastic anemia) -I'm not totally clear on exclusion for quality score - "only the best quality and relevant articles included" for a given "database/study population" - in the next section it sounds like articles with low quality (<13) were excluded, but were other articles excluded for different reasons based on geography/redundancy? Results -Table 1 is great, but I wonder if we could get a map showing prevalence by province/autonomous region. This would be very helpful perhaps as a separate panel to the existing map! -In Tables 1/2 if possible to do without impacting readabily too much, it would be good to have the pooled n of subjects for each category -The wide ranges in reported prevalences would respond well to a metanalytic presentation I think -The figure of prevalence over time should have confidence intervals for each point Figure 4: I really like this visual presentation of the qualitative findings. However, some clarification is needed. Some things here don't seem to match the text - for example the text finds no sex difference in anemia, but here in the figure sex is mentioned as associated. Similarly, I think all the items listed here need to be clarified in terms of their directionality - some are associated with reduced and others with increased anemia. I suggest editing this table so the directionalility is immediately clear in all instances. Discussion -discussion is adequate, although again I'm curious why the authors don't think they can do pooled estimates - the included studies aren't actually that heterogenous it seems (>70% are cross sectional). If authors decided to not pursue a pooled analysis, then I think this needs stronger justification.

VERSION 1 – AUTHOR RESPONSE

Dear Prof. Imti Choonara,

My coauthors and I thank you for your comments and suggestions concerning our manuscript "Anemia Prevalence and its Associated Factors in Children under 5 Years in Western China: A Systematic Review" (ID: bmjpo-2021-001185). We also appreciate the thoughtful comments from the reviewers, which have much improved the paper.

We have studied the comments carefully and have revised our paper accordingly. This letter provides point-by-point responses to each comment and summarizes relevant changes in the manuscript. Primary changes to the paper include the following:

- We have revised the manuscript from a scoping review to a systematic review without meta-analysis.
- We have made modifications to the grammar/syntax, references, abbreviations, numerals, and personal pronouns based on the copy-editing comments to improve the English of the manuscript.

All changes are highlighted using a track change function in the revised manuscript (marked copy). Please see the attachment for the details of the response letter.

The material in the manuscript has not and will not be submitted elsewhere for possible publication as long as it is under consideration by *BMJ Paediatrics Open*.

Once again, we are very grateful for your consideration of our work. We look forward to receiving your feedback on the revised manuscript.

Sincerely.

Zhou Huan

Response to Editor in Chief Comments

Comment 1: Please rewrite as a systematic review - see the comments of the statistical reviewer (Reviewer: 3; Dr. Sarah Nevitt, University of Liverpool).

Response: Thanks for your comment. We have rewritten the manuscript as a systematic review based on the comments of the Reviewer 3. Accordingly, we have also supplemented the protocols required for the

system review.

Comment 2: Please replace Key Messages with What is already known and What this study adds sections. Your review is an original article.

Response: Thank you for your comment. We have replaced the Key Messages with the "What is Already Known" and "What this Study Adds" sections, and we have added them on page 2 of the manuscript (added text in italics).

What is known about the subject?

- *Iron deficiency anemia disproportionately affects infants and children in low- and middle-income areas.*
- *Western China, which covers 72% of China's total area and is home to 27% of the total population, is one of the least economically developed regions in the country.*
- *In China, 4 of the 5 provinces with the highest rates of childhood anemia are located in Western China.*

What this study adds?

- *In Western China, the median prevalence of anemia in children under 5 years is 40%, which is much higher than the national average.*
- *The highest prevalence rates (59.1% to 75.74%) were located in Qinghai province, and the highest levels were reported among children aged 6-24 months.*
- *Regional contexts, individual sociodemographic characteristics and feeding behaviors, and nutritional program interventions play important roles in the prevalence of childhood anemia in Western China.*

Response to Reviewer 1

General Comment: A good study. Please find some comments below.

Response to General Comment: Thank you for recognizing the value of our research and for pointing out some of the limitations of this paper. In the following pages, we provide a point-by-point response to your comments, including specific changes made to the revised manuscript. All changes are highlighted using a track change function in the revised manuscript.

Comment 1: Define cross-sectional studies whether these were descriptive or analytical i.e., how many reported only prevalence and how many reported associated factors analyses as well. If both then mention analytical. How are surveillance studies different from cross-sectional?

Response: In this comment, the reviewer makes two points. The first point asks us to clearly show the number of studies of descriptive or analytical. The second point asks how the surveillance studies differ from cross-sectional studies. For clarity, we will respond to each point separately.

In response to the first point:
We agree with your suggestion to define studies as descriptive or analytical. A total of 29 articles are descriptive, and the remaining 29 are analytical, reporting associated factors analyses as well. We have added this information to the RESULTS section of the paper, page 7, in the revised manuscript (revised text in italics).

"Most studies were cross sectional studies (n = 41, 70.69%), followed by prospective cohort studies (n =6, 10.34%), surveillance data (n=5, 8.62%), RCTs (n=3, 5.17%), and quasi-experiments (n=3, 5.17%). 29 studies were descriptive and the remaining 29 were analytical (reporting associated factors analyses). The manner of reporting data varied across the studies, and we report data in their original format (see in Appendix Table 1)."

In response to the second point:
According to MIT's [1] definition of surveillance study, at the most general level, surveillance study of humans can be defined as regard or attendance to others (whether a person, a group, or an aggregate as with a national census) or to factors presumed to be associated with these subjects.

We consider surveillance data not as a specific type of study, but as a data set from a particular source. Such data sets are formed by continuous, dynamic, and quantitative observation of indicators reflecting the internal and external states of the study subjects and their influencing factors using relevant measurement instruments [2]. The data mainly come from disease and environmental surveillance systems, including information on disease outcomes and exposure to pathogenic factors, etc., without additional input. In contrast, cross-sectional studies are conducted at a specific time, i.e., at a point in time or over a short period of time, to investigate the relationship between factors of interest and disease or health status in a specific population by means of a census or sample survey, thus describing the distribution of disease and observing the relationship between certain factors and disease over this time period.

Therefore, we currently prefer to distinguish surveillance data from cross-sectional studies and look forward to your additional comments and guidance.

REFERENCES:
[1] Gary T. Marx. Surveillance Studies. International Encyclopedia of the Social & Behavioral Sciences, Second Edition, 2015, 733–741. http://dx.doi.org/10.1016/B978-0-08-097086-8.64025-4

[2] Zhao Zhe, Wang Haitao, Jiang Baofa. Applications of statistical models on surveillance data in ecological study. Chinese Journal of Epidemiology, 2019, 40(8): 1010-1017. http://dx.doi.org/10.3760/cma.j.issn.0254-6450.2019.08.026

Comment 2: I would suggest to use the word 'pooled' or systematic 'meta-analysis' review in the manuscript freely with overall reported prevalence to indicate meta-analyzed effect estimates.

Response: Thank you for your comment. After our review of the literature and thoughtful consideration, we generally agree with you to revise the type of this review. However, we believe that a systematic review (without meta-analysis) may be more appropriate for this study. There are two reasons for this:

First, there were substantial variations in the sample sizes, study settings, survey methodologies, and populations among the included studies. Given the heterogeneity among the included studies, it felt inappropriate to combine all studies and perform a meta-analysis to provide pooling statistics [1-3]. According to *Cochrane Handbook* [4], meta-analysis should only be considered when a group of studies is sufficiently homogeneous in terms of participants, interventions, and outcomes to provide a meaningful summary. Furthermore, meta-analysis techniques are not suitable because nearly all the studies were representative of the whole population and not restricted to samples [5]. Although the data were not suitable for meta-analysis, the systematic approach is a useful and clear method for providing a data summary and clearly demonstrating where gaps exist. Therefore, based on your kind suggestions, we have re-examined the design of our review and revised the entire manuscript to a systematic review (without meta-analysis).

REFERENCES:

[1] Systematic Review VS Meta-Analysis. https://scientific-publishing.webshop.elsevier.com/manuscript-review/systematic-review-vs-meta-analysis/
[2] Rastin M, Mahmoudi M, Sahebari M, Tabasi N. Clinical & immunological characteristics in systemic lupus erythematosus patients. Indian J Med Res. 2017 Aug;146(2):224-229. doi: 10.4103/ijmr.IJMR_1356_15.
[3] Hoofwijk DM, van Reij RR, Rutten BP, Kenis G, Buhre WF, Joosten EA. Genetic polymorphisms and their association with the prevalence and severity of chronic postsurgical pain: a systematic review. Br J Anaesth. 2016 Dec;117(6):708-719. doi: 10.1093/bja/aew378.
[4] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd Edition. Chichester (UK): John Wiley & Sons, 2019.
[5] Stöckl H, Devries K, Rotstein A, Abrahams N, Campbell J, Watts C, Moreno CG. The global prevalence of intimate partner homicide: a

systematic review. Lancet. 2013 Sep 7;382(9895):859-65. doi: 10.1016/S0140-6736(13)61030-2. Epub 2013 Jun 20. PMID: 23791474.

Comment 3: Associated factors are mentioned but what are the pooled ORs or RRs with 95% CI for example, if Tibetan ethnic group has a higher risk of anemia, then how much fold or times increased risk and whether it is statistically significant or not?

Response: We have revised the manuscript as a systematic review based on your comment. However, due to limitations, such as the heterogeneity of included studies mentioned above, our review was only a descriptive synthesis of the characteristics and findings of the included studies. It was not possible to use statistical methods to summarize the results of independent studies and to provide pooled *ORs* or *RRs* with *95% CIs*.

In addition, only two studies on anemia differences between Tibetan and Han children (*a24* and *a50*) were included in this review. Both studies performed multivariate logistic analysis of anemia prevalence (controlling for confounding factors such as child's age, sex, mother's age, and education level), where *OR*=2.301 (*95% CI*: 1.863 to 2.843) in *a24* and *OR*=3.123 (*95% CI*: 1.473 to 6.623) in *a50*. Specifically, both articles supported children of Tibetan nationality were 2.301 (*a24*) and 3.123 (*a5*) times more likely to be anemic ($p<0.05$).

Comment 4: It is not clear what statistical software was used for analysis i.e., to pool the prevalence. Also, the authors mention about heterogeneity in the Abstract but I^2 statistics and Chi-square P-values are not mentioned to distinguish between-study from within-study variation.

Response: Thanks again for your comment. In this comment, the reviewer makes two points. The first point asks us to clearly show the statistical software used for analysis. The second point asks about the reason why we not mentioned I^2 statistics and Chi-square P-values to distinguish between-study from within-study variation. For clarity, we will respond to each point separately.

<u>In response to the first point:</u>
We apologize that the data analysis software and analysis methods used were not clearly described in the original manuscript. After we revised the entire manuscript to a systematic review, we have added this **Data Synthesis** section on pages 6-7 (revised text in italics):

"The analysis consisted of four steps: (1) calculation of anemia prevalence estimate in children under 5 years of age per province (distinguishing between urban and rural areas if reported), using the median percentage with IQR; (2) stratification of prevalence estimates by sex, age, and ethnic group, separately; (3) collation of factors associated with childhood anemia; and (4) selection of one estimate per study-year,

scatter chart plotting, and linear regression predictions. Data were analyzed with Stata version 16.0."

In response to the second point:
The *Cochrane Handbook* [1] defines *"heterogeneity"* as: any kind of variability among studies in a systematic review. It can be distinguished between different types of heterogeneity:

1. Clinical heterogeneity: variability in the participants, interventions, and outcomes studied;
2. Methodological heterogeneity: variability in study design and risk of bias;
3. Statistical heterogeneity: variability in the intervention effects being evaluated in the different studies, and is a consequence of clinical or methodological diversity, or both, among the studies.

Based on the substantial variations in the sample size, study settings, survey methodology, and populations, there was a large clinical and methodological heterogeneity among the included studies. As we know, the chi-squared (χ^2, or Chi^2) test and I^2 are commonly used in statistical heterogeneity of meta-analysis. It is legitimate for a systematic review to focus on examining the relationship between some clinical characteristic(s) of the studies and the size of intervention effect, rather than on obtaining a summary effect estimate across a series of homogeneous studies. As we mentioned in Comment 2, we think that a systematic review (without meta-analysis) may be more appropriate for this study. Thus, because we do not conduct a meta-analysis, we cannot use the I^2 statistics and Chi-square *P*-values to distinguish between-study from within-study variation. Additionally, we conducted an assessment of the risk of bias (or "quality") of studies in the METHODS section.

REFERENCES:
[1] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. 2nd Edition. Chichester (UK): John Wiley & Sons, 2019.

Comment 5: I would also suggest to include meta-regression to adjust for confounding effects of different associated factors rather than univariate analyses and to justify study of effect modification only for significant factors like that done for age, gender, and ethnicity with P-values for interaction.

Response: Thank you for your comment. After our thoughtful consideration, we believe that a systematic review (without meta-analysis) may be more appropriate for this study. We use the systematic review to focus on descriptive synthesis of the characteristics and findings of the included studies. Due to limitations of data from the included studies, we were unable, in this review, to include meta-regression to

adjust for confounding effects of different associated factors and to justify study of effect modification only for significant factors.

We hope that we have answered and clarified all comments on the selection of review types and are looking forward to your reply.

Comment 6: I would suggest to explore if data is available on use of iron supplements or prevalence of childhood helminthiasis in each of the regions to do some rough correlation for Discussion section even if that available from different other studies in published literature.

Response: Thanks for your suggestion and you make two points. The first point asks us whether data on the use of iron supplements in each of the regions are available. The second point asks about whether the data on prevalence of childhood helminthiasis in each of the regions are available. For clarity, we will respond to each point separately.

In response to the first point:
At present, there is no way to supplement data on the use of iron supplements in each region. Although anemia rates were reported in each of the included studies, the vast majority of studies were limited to the status and influencing factors of anemia, and only a few studies conducted interventions with iron supplements. In addition, there are many kinds of iron supplements, including "Ying Yang Bao" (YYB), multiple-micronutrient sprinkles, iron-fortified foods, micronutrient supplements, and so on, which are difficult to be classified and summarized.

However, according to a review of the literature, the YYB (a free government-distributed nutritional supplement, with iron as the main ingredient) is popular in Western China, which we have mentioned in the DISCUSSION section (page 13).

"In 2013, 187 counties in Western China were covered by this project, and between 2012 and 2017, the national anemia rate decreased from 32.9% to 17.6%. These results are illustrated in Figure 3, which shows the prevalence of childhood anemia in the western region has been decreasing year by year."

In response to the second point:
There are no studies on the effect of helminthiasis on anemia in children in China, but the conclusions of foreign studies are different [1-3]. Chinese research shows that the detection rate of key parasites (including soil-borne nematodes, Enterobius vermicularis, Clonorchis sinensis, Taenia solium, and intestinal protozoa) among children aged 3-6 years is 3.30%, and the weighted infection rate of key parasites is 5.96% [4]. For children aged 3-9 years in 736 surveillance sites in 30 provinces (municipalities and autonomos regions) in China, the pinworm infection rates in 2016-2018 were 2.50%, 2.84% and 2.46%, respectively. Among them, 9

provinces (municipalities and autonomous regions) in western China in 2018 had a rag worm infection rate of less than 1% [5]. To summarize, because the prevalence of helminthiasis is low in China (including Western China), think there is little to no need to focus on the effect of helminthiasis on anemia in this particular review.

REFERENCES:

[1]. Gutierrez-Rodriguez C, Trujillo-Hernandez B, Martinez-Contreras A, et al. Frequency of intestinal helminthiasis and its association with iron deficiency and malnutrition in children from western Mexico. *Gaceta Medica De Mexico* 2007;143(4):297-300.
[2]. Atwa ZTH, Thabet MM. INTESTINAL PARASITIC INFECTION IN EGYPTIAN CHILDREN: COULD IT BE A RISK FACTOR FOR IRON DEFICIENCY ANEMIA? *Journal of the Egyptian Society of Parasitology* 2016;46(3):533-540.
[3]. Bartoloni A, Cancrini G, Roselli M, et al. Iron deficiency in an area of Bolivia and high prevalence of intestinal helminthiasis. *Parassitologia* 1990;32(3):335-338.
[4]. Chen Y, Zhou C, Zhu H, et al. National survey on the current status of important human parasitic diseases in China in 2015. *Chinese Journal of Parasitology and Parasitic Diseases* 2020;38(01):5-16.
[5]. Huang J, Zhang M, Zhu H, et al. National surveillance on Enterobius vermicularis infections among children at ages of 3 to 9 years in China from 2016 to 2018. *Chinese Journal of Schistosomiasis Control* 2020;32(1):54-59.

Comment 7: In the PRISMA flow-diagram I would suggest to add labels in big boxes to the left of the figure to add headings of steps of review like screening, eligibility assessment, data extraction, full-text review, final inclusion and meta-analysis, etc.

Response: We agree with the reviewer that adding labels in big boxes to the left of the figure to illustrate the headings of steps of review, would make the figure simpler and more intuitive. We have added this point to the Figure 2.

Comment 8: I would like to know what was the range of prevalence in studies which were excluded due to low quality or sample size less than 50 in the Results section to see if excluded or missing studies had different sizes compared to included studies.

Response: Thank you for your comment. Given that most of the studies with sample sizes less than 50 were clinical studies or case reports of children in a particular hospital or community, studies under this size were considered not having adequate statistical power for generalization, and we believe that the results of these studies are hardly representative of the regional prevalence of anemia. Therefore, their prevalence is not reported in our study.

Response to Reviewer 2

General Comment: The article presents the objectives, methods and results in a clear and acceptable way.

> **Response to General Comment:** Thank you for recognizing the value of our research.

Response to Reviewer 3

General Comment:
I have conducted a statistical review of the manuscript "Anemia prevalence and its Associated Factors in Children under 5 Years Old in Western China: A Scoping Review". The authors aim to conduct a scoping review of anemia prevalence and factors associated with anemia in Western China.

The authors have clearly comprehensively searched the literature and extracted a lot of information from the included studies related to the objective. But I think that what the authors have actually done here is closer to a systematic review than a scoping review.

Scoping reviews are generally conducted where the evidence base for a particular question are unclear and further information regarding types of available evidence, research methods used, definitions and concepts etc. related to a question are of interest.

Scoping reviews may be conducted as a learning exercise prior to a systematic review, to inform the inclusion criteria, outcomes etc. of a systematic review.

Scoping reviews are generally not suitable for synthesizing numerical data and/or providing effect sizes. The authors seem to justify the scoping review approach because previous systematic reviews have shown heterogeneity between studies. This is not a necessarily a reason to perform a scoping review rather than a systematic review, but it would be a reason to perform a systematic review without meta-analysis.

If the authors wished to gain further insight about the design, methods used, factors examined etc. in the prevalence studies a scoping review would be suitable. However, as the aim of the authors is to provide estimates (i.e., medians and ranges) of prevalence rates across different regions, to compare subgroups and to identify factors significantly associated with prevalence rates, a systematic review would be more suitable.

Please see the following reference for further details on the differences between systematic reviews and scoping reviews:
Munn, Z., Peters, M.D.J., Stern, C. et al. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 18, 143 (2018). https://doi.org/10.1186/s12874-018-0611-x

I suggest that the authors start again and conduct a systematic review.

Response to General Comment: Thank you for your comments, which have provided us with a more in-depth and detailed understanding of the selection of review types. After our careful consideration and group discussion, we absolutely agreed with your opinion. Based on your comments and the interpretation of the recommended literature from BMC Med Res Methodol, we do accept that this study is more appropriate as a systematic review rather than a scoping review.

There are two main reasons to support this: (1) Firstly, heterogeneity between included studies is not a strong reason to perform a scoping review. Scoping reviews are best designed for [1]: when a body of literature has not yet been comprehensively reviewed, or exhibits a large, complex, or heterogeneous nature not amenable to a more precise systematic review. Instead, both narrative review and systematic review (without meta-analysis) can also be used as a solution of clinical/methodology heterogeneity between included studies [2]. (2) In addition, the "research questions" of our study are also inapplicable to the problem addressed by the scoping review. We realize that the logical structure and specific approach of this paper is closer to that of a systematic review.

Therefore, we agree that a systematic review (without meta-analysis) may be more appropriate for this study. And we have revised the whole manuscript as a systematic review following the guideline of PRISMA checklist. We hope that we have answered and modified the manuscript properly and are looking forward to your reply.
REFERENCES:
[1] Peters M, Godfrey C, Khalil H, et al. Guidance for Conducting Systematic Scoping Reviews. *Int J Evid Based Healthc*. 2015;13:141-146.
[2] Munn Z, Peters M, Stern C, et al. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*. 2018;18:143.

Specific comments: I have a few other specific comments on the approach and presentations of results:

Comment 1: Throughout the manuscript the term 'child under the age of five' is used. Please replace with 'children under the age of five.'

Response: Thanks for your comment. We apologize for the inappropriate wording. In the revised manuscript, we have thoroughly checked the full text and corrected all grammatical errors.

Comment 2: Methods: Please note that PRISMA (for systematic reviews or for scoping reviews) are reporting guidelines rather than conducting guidelines or methodological framework.

Response: Thanks for your comment. We have recognized that this was a writing mistake on our part, and actually, the PRISMA checklist was used as a framework for writing. We have removed this sentence in the METHODS section of the paper, page 5.

Comment 3: Methods: The search strategy, screening process for the studies and quality assessment are described in detail but more information is needed in the methods section of a review (whether a scoping review or a systematic review).

Further details are needed of the data which will be extracted, how this data will be summarized (including any effect size used – e.g., median and range of prevalence) and how data will be synthesized and presented (in other words, whether a quantitative synthesis is planned (i.e., meta-analysis) or a narrative synthesis in figures and tables).

Response: Thanks for your comment. We highly appreciate your opinion and have revised the METHODS section entirely according to the PRISMA reporting framework.

"METHODS
This systematic review was conducted according to PRISMA guidelines[13]. The project protocol was registered with PROSPERO.

Eligibility criteria
Studies were eligible for inclusion only if they stated the prevalence of children under 5 years with IDA in Western China. "Western China" is not a specific administrative division, but the region includes 12 provinces, autonomous regions, and municipalities: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang (Figure 1)[14]. As prealence data may be sourced from different study designs, we included all relevant cross-sectional studies, randomized controlled trials (RCTs), cohort studies, and published surveillance data. If results based on the same data were presented in more than one publication, results from only one publication were included.

We excluded clinical studies or case reports of children in specific hospitals or communities, and studies with sample sizes less than 50 participants, as the results of these studies are hardly representative of the regional prevalence of anemia. We excluded studies on children in selective samples (premature, low birthweight, birth defects) or with specific types of anemia (aplastic anemia, thalassemia, megaloblastic anemia). Studies lacking clear presentation of prevalence or diagnostic criteria were also excluded.

Search Strategy

Literature search strategies were developed using medical subject headings (MeSH) and text words related to childhood anemia. The search terms for studies published in English were (anemia OR anaemia OR iron deficiency anemia OR IDA OR nutritional anemia) AND (infants OR children OR preschool) AND (China OR Chinese). The search terms for studies published in Chinese were (贫血 OR 缺铁性贫血 OR 营养性贫血) AND (婴儿 OR 婴幼儿 OR 幼儿 OR 儿童). We searched Medline (Ovid interface, 1948 onwards), Embase, PubMed, Web of Science, CNKI, WanFang Data, and VIP. The literature search was limited to studies written in English and Chinese languages, published from 1 January 2011 to 30 June 2021. We carefully examined reference lists of published articles to find other related publications not identified in the database search.

Selection process

EndNote X9 was used to manage search results and delete duplicates. Two researchers (YF and FQ) independently screened the titles and abstracts identified through the search against the inclusion criteria. As the wide range of geographic locations in Western China cannot be defined by search terms, the two researchers screened the full text reports and determined whether studies met the inclusion criteria. Disagreements were resolved by including a third researcher (LY) to make the final decision.

Assessment of methodological quality

We used standardized forms from Joanna Briggs Institute Critical Appraisal Checklist for Studies Reporting Prevalence Data[17] to determine the methodological quality of included studies. Two researchers (LY and FQ) independently evaluated 9 areas of study design, conduct, and analysis for each included study. Each of the 9 areas were qualified as "poor quality," "moderate quality," or "high quality," receiving a score from 0 (poor quality) to 2 (high quality). Researchers then discussed and made final quality assessments. Total quality scores ranged from 0 to 18 and studies that scored less than 13 were excluded.

Data abstraction

A standardized reporting form was used to extract data from each publication (Appendix Table 1). The form included: study ID, first author's name, year of publication, language, study design, year of data collection, place where the study was conducted, sample size, age range of study subjects, prevalence estimates (stratified by sex, age, and ethnic group), and quality score.

Data Synthesis

The analysis consisted of four steps: (1) calculation of anemia prevalence estimate in children under 5 years of age per province (distinguishing between urban and rural areas if reported), using the median percentage with IQR; (2) stratification of prevalence estimates by sex, age, and ethnic

group, separately; (3) collation of factors associated with childhood anemia; and (4) selection of one estimate per study-year, scatter chart plotting, and linear regression predictions. Data were analyzed with Stata version 16.0."

Comment 4: Inclusion of studies with sample size of over 50: While this may be a reasonable inclusion criterion, I'm not sure about this reference (i.e., a systematic review of the influence of grandmothers on breastfeeding rates). Does this reference demonstrate that studies with sample size <50 may not have adequate statistical power or is this simply also an inclusion criterion of this review? If the latter, I suggest that a methodological reference should be cited instead.

Response: Thank you for your comment. Since the British statistician William Sealy developed the "small sample theory" in the early 20th century, it is generally accepted that a sample size of less than 50 (or less than 30) is not considered to have sufficient statistical power in statistics.

However, we sincerely apologize for not being able to find an explicit methodological reference. We did find that Dr. Joost de Winter et al. mentioned a similar statement in his methodological article on Exploratory Factor Analysis [1]: *"Exploratory factor analysis (EFA) is generally regarded as a technique for large sample sizes (N), with N = 50 as a reasonable absolute minimum. This study offers a comprehensive overview of the conditions in which EFA can yield good quality results for N below 50."*

REFERENCES:

[1] de Winter JC, Dodou D, Wieringa PA. Exploratory Factor Analysis With Small Sample Sizes. Multivariate Behav Res. 2009 Mar-Apr;44(2):147-81. doi: 10.1080/00273170902794206.

Meanwhile, we have also found other systematic reviews that include "sample size >50" as an inclusion criterion. But unfortunately, we still could not find an explicit methodological reference from the method sections of these articles. Here are these references:

1. LeBrun DG, Banskota B, Banskota AK, Rajbhandari T, Baldwin KD, Spiegel DA. Socioeconomic Status Influences Functional Severity of Untreated Cerebral Palsy in Nepal: A Prospective Analysis and Systematic Review. Clin Orthop Relat Res. 2019;477(1):10-21. doi:10.1097/CORR.0000000000000476

2. Bell JA, Galaznik A, Huelin R, et al. Systematic Literature Review of Treatment Options and Clinical Outcomes for Patients with Higher-Risk Myelodysplastic Syndromes and Chronic Myelomonocytic Leukemia. Clin Lymphoma, Myeloma Leuk. 2018;18(4):e157-e166. doi:10.1016/j.clml.2018.02.001

3. Dong S, Yang Y, Wang Y, et al. Prevalence of Cryptosporidium Infection in the Global Population: A Systematic Review and Meta-

analysis. Acta Parasitol. 2020;65(4):882-889. doi:10.2478/s11686-020-00230-1
4. Morkisch N, Upegui-Arango LD, Cardona MI, et al. Components of the transitional care model (TCM) to reduce readmission in geriatric patients: A systematic review. BMC Geriatr. 2020;20(1):1-18. doi:10.1186/s12877-020-01747-w
5. Panda A, Bhalla AS, Goyal A. Bronchial artery embolization in hemoptysis: A systematic review. Diagnostic Interv Radiol. 2017;23(4):307-317. doi:10.5152/dir.2017.16454
6. Roheger M, Kalbe E, Liepelt-Scarfone I. Progression of cognitive decline in Parkinson's disease. J Parkinsons Dis. 2018;8(2):183-193. doi:10.3233/JPD-181306

To solve this issue completely, we modified the presentation of the inclusion and exclusion criteria in the methods section and attempted to circumvent references involving. we have
added **Eligibility criteria** section on page 5 (revised text in italics):

"Eligibility criteria
Studies were eligible for inclusion only if they stated the prevalence of children under 5 years with IDA in Western China. "Western China" is not a specific administrative division, but the region includes 12 provinces, autonomous regions, and municipalities: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang (Figure 1)[14]. As prevalence data may be sourced from different study designs, we included all relevant cross-sectional studies, randomized controlled trials (RCTs), cohort studies, and published surveillance data. If results based on the same data were presented in more than one publication, results from only one publication were included.

We excluded clinical studies or case reports of children in specific hospitals or communities, and studies with sample sizes less than 50 participants, as the results of these studies are hardly representative of the regional prevalence of anemia. We excluded studies on children in selective samples (premature, low birthweight, birth defects) or with specific types of anemia (aplastic anemia, thalassemia, megaloblastic anemia). Studies lacking clear presentation of prevalence or diagnostic criteria were also excluded."

Comment 5: Quality assessment: "Each of the item was identified by a score from 0 (unclear) to 2 (yes)," I assume that this means unclear = 0, no = 1 and yes = 2?
Given that the questions of the JBI checklist consider whether methods used are appropriate or adequate where the wording of the questions means that yes = appropriate methods, no=inappropriate methods and unclear = insufficient information, to assign a higher score to 'no' (where methods are clearly

inappropriate) than to 'unclear' (where there is uncertainty whether methods are appropriate) doesn't make sense.

The JBI checklist doesn't assign scores, rather the reviewers should decide whether to include or exclude the studies based on the responses to the 9 questions, and where studies are excluded, reasons should be provided for this. I suggest that the tools should be used as intended rather than assigning scores.

Response: Thank you for bringing this to our attention. In order to better address the question, we will answer it in two parts

First of all, we acknowledged that the JBI checklist does not assign scores, rather, the reviewers should decide whether to include or exclude the studies based on the responses to the 9 questions. We apologize for our misstatement of assigning the items by a score from 0 (unclear) to 2 (yes). So, we clarified the ideas according to your opinions and made corresponding modifications in the manuscript.

Actually, our intention was to evaluate the quality of the studies according to the 9 questions of JBI and exclude poor quality study. For this reason, we have extensively searched relevant studies and found that there is no standard cut-off value in the JBI Checklist to exclude studies. The published studies have set cut-off values according to the purposes of their own. For example, Ofori-Asenso R et al. classified studies were ineligible if fewer than 5 of the criteria were achieved span style="font-family:'Times New Roman'; font-size:8pt; vertical-align:super">[1]; Torgbenu, E. et al. classified studies into low or high risk of bias using a cut-off of 70% [2].

We refer to a similar study published in the Lancet [3], two researchers independently evaluated 9 questions, and the evaluation results of each question were "poor quality", "moderate quality" and "high quality", received a score from 0 (poor quality) to 2 (high quality). Studies with a total score of less than 13 were excluded. For example, in question 3, "Was the sample size adequate?" If a sample size calculation is performed and the sample size is sufficient, it is "high quality," and receives a score of 2. If the study did not conduct sample size calculation and determined it was not a large national survey, we would conduct our own sample size analysis which was recommended by the JBI Checklist. In this case, sufficient sample size was rated as "moderate quality" (score of 1) and insufficient sample size was rated as "poor quality" (score of 0).

Details of full-text articles retrieved for critical appraisal have been added in the **Supplementary File**. Additionally, the **Assessment of methodological quality** section of the manuscript has been revised for clarity in the revised manuscript, page 6 (revised text in italics):

"We used standardized forms from Joanna Briggs Institute Critical Appraisal Checklist for Studies Reporting Prevalence Data[17] to determine

the methodological quality of included studies. Two researchers (LY and FQ) independently evaluated 9 areas of study design, conduct, and analysis for each included study. Each of the 9 areas were qualified as "poor quality," "moderate quality," or "high quality," receiving a score from 0 (poor quality) to 2 (high quality). Researchers then discussed and made final quality assessments. Total quality scores ranged from 0 to 18 and studies that scored less than 13 were excluded."

REFERENCES:

[1] Ofori-Asenso R, Chin KL, Mazidi M, Zomer E, Ilomaki J, Zullo AR, Gasevic D, Ademi Z, Korhonen MJ, LoGiudice D, Bell JS, Liew D. Global Incidence of Frailty and Prefrailty Among Community-Dwelling Older Adults: A Systematic Review and Meta-analysis. JAMA Netw Open. 2019 Aug 2;2(8):e198398. doi: 10.1001/jamanetworkopen.2019.8398.

[2] Torgbenu E, Luckett T, Buhagiar MA, Chang S, Phillips JL. Prevalence and incidence of cancer related lymphedema in low and middle-income countries: a systematic review and meta-analysis. BMC Cancer. 2020 Jun 29;20(1):604. doi: 10.1186/s12885-020-07079-7.

[3] Borschmann R, Janca E, Carter A, Willoughby M, Hughes N, Snow K, Stockings E, Hill NTM, Hocking J, Love A, Patton GC, Sawyer SM, Fazel S, Puljević C, Robinson J, Kinner SA. The health of adolescents in detention: a global scoping review. Lancet Public Health. 2020 Feb;5(2):e114-e126. doi: 10.1016/S2468-2667(19)30217-8.

Comment 6: Table 1: 58 studies are included in the review but only 51 studies are summarized in Table 1. What sort of information did the 7 studies not included in Table 1 provide which were of relevance to the review?

Response: We apologize. Due to our mistake, three studies were omitted in the final number of studies that reported the anemia rate by province. These three studies are *a52*, which reported the anemia prevalence of Sichuan Province; *a53*, which reported the anemia prevalence of Gansu Province; and *a10*, which reported the overall anemia prevalence of Sichuan and Gansu Province. Therefore, a total of 54 studies reported the anemia prevalence by province. In addition, 4 studies are not included in Table 1 because they did not report the overall anemia prevalence by province. Among them, three studies (*a7, a27 and a35*) reported the anemia prevalence by age group, and one study (*a42*) only reported the anemia prevalence of the experimental and control groups respectively after intervention.

We have revised this accordingly in Table 1 and attached at the end of this response letter.

Comment 7: Figure 3: I think I understand what the authors are trying to show here but there are a few issues with this plot.

Firstly, the plot does not show 'time of each study' but shows estimates across different regions over time. By showing summary estimates across regions (presumably the median values), this does not capture the uncertainty and variability across studies.

I assume that studies report only prevalence estimates for specific years or time intervals, rather than specific information about the rate of change in prevalence over time. By joining up the point estimates at different time points with straight lines, this implies a linear change (increase or decrease) in prevalence over time.

Commenting on differences in prevalence estimates in different regions in different years is fine (ideally with uncertainty in the estimates captured), but the authors should avoid making assumptions about how prevalence changes over time without details from the studies to inform this.

> **Response:** Thank you for your comment and we generally agree with your opinion. We ignored the fact that direct linkage was not allowed between individual studies in different regions at different times. After consulting experts and referring to similar studies, we selected one prevalence per study-year, plotting a new scatter chart and making linear regression predictions. If provinces have more than one estimate available for a given year, to avoid double-counting in the time trend chart, we gave preference to the study with a higher quality assessment score. We have revised the Figure 3 and attached at the end of response letter.

Response to Reviewer 4

General Comment: This is a remarkably interesting study, which does a great service by comprehensively reviewing Chinese language research literature and medical thesis and getting this data into an accessible format for the global pediatric community.

> **Response to General Comment**: Thank you for your comment and for recognizing the value of our study. We agree that our findings provide a synthesis of Chinese language research literature for the global pediatric community. We also appreciate your comments and have been corrected point-by-point.

Specific comments:

Comment 1: Overall - English could be improved - a close edit is needed to improve readability

Response: Thank you for pointing this out. A native-English speaker has reviewed the entire manuscript and supplementary files and has made language edits for readability.

Comment 2: Methods - Overall I am mostly wondering why the authors haven't conducted a prevalence meta-analysis. Most of the studies are cross sectional, quality assessment has been done, and the search strategy is systematic. At least for these cross-sectional studies, it seems like it would be logical to do the meta-analysis. This seems to me the biggest limitation, and the publication would be higher impact if this was done.

Response: Thank you for your comment. After our review of the literature and thoughtful consideration, we generally agree with you to revise the type of this review. However, we believe that a systematic review (without meta-analysis) may be more appropriate for this study. There are two reasons for this:

First, there were substantial variations in the sample sizes, study settings, survey methodologies, and populations among the included studies. Given the heterogeneity among the included studies, it was felt to be inappropriate to combine all the studies and perform a meta-analysis to provide pooling statistics [1-3]. According to *Cochrane Handbook* [4], meta-analysis should only be considered when a group of studies is sufficiently homogeneous in terms of participants, interventions and outcomes to provide a meaningful summary. Furthermore, meta-analysis techniques are not suitable because nearly all the studies were representative of the whole populationand not restricted to samples [5]. Although the data were not suitable for meta-analysis, the systematic approach is a useful and clear method for providing a data summary and clearly demonstrating where gaps exist. Therefore, based on your kind suggestions, we have re-examined the design of our review and revised the entire manuscript to a systematic review (without meta-analysis).

REFERENCES:

[1] Systematic Review VS Meta-Analysis. https://scientific-publishing.webshop.elsevier.com/manuscript-review/systematic-review-vs-meta-analysis/

[2] Rastin M, Mahmoudi M, Sahebari M, Tabasi N. Clinical & immunological characteristics in systemic lupus erythematosus patients. Indian J Med Res. 2017 Aug;146(2):224-229. doi: 10.4103/ijmr.IJMR_1356_15.

[3] Hoofwijk DM, van Reij RR, Rutten BP, Kenis G, Buhre WF, Joosten EA. Genetic polymorphisms and their association with the prevalence and severity of chronic postsurgical pain: a systematic review. Br J Anaesth. 2016 Dec;117(6):708-719. doi: 10.1093/bja/aew378.

[4] Stöckl H, Devries K, Rotstein A, Abrahams N, Campbell J, Watts C, Moreno CG. The global prevalence of intimate partner homicide: a systematic review. Lancet. 2013 Sep 7;382(9895):859-65. doi: 10.1016/S0140-6736(13)61030-2. Epub 2013 Jun 20. PMID: 23791474.

Comment 3: Methods - Exclusion criteria: excluding studies with samples less than 50: again, this seems to set up the meta-analysis as the authors talk about only including studies with adequate statistical power.

Response: Thanks again for your comment. We agree with your comments on "including studies with adequate statistical power also support a more systematic review" and have revised the entire manuscript to a systematic review.

Comment 4: Methods - Exclusion criteria need to be explicitly detailed (it seems like only examples are given, e.g., prematurity and aplastic anemia).

Response: Thank you very much for your comments. We have modified the Exclusion criteria more explicitly based on your suggestion in the revised manuscript, page 5 (revised text in italics):

"We excluded clinical studies or case reports of children in specific hospitals or communities, and studies with sample sizes less than 50 participants, as the results of these studies are hardly representative of the regional prevalence of anemia. We excluded studies on children in selective samples (premature, low birthweight, birth defects) or with specific types of anemia (aplastic anemia, thalassemia, megaloblastic anemia). Studies lacking clear presentation of prevalence or diagnostic criteria were also excluded."

Comment 5: Methods - I'm not totally clear on exclusion for quality score - "only the best quality and relevant articles included" for a given "database/study population" - in the next section it sounds like articles with low quality (<13) were excluded, but were other articles excluded for different reasons based on geography/redundancy?

Response: We apologize for the confusion. Regarding the selection process, if there were multiple articles based on the same data, the two researchers usually made independent judgments, such as merit inclusion based on the quality of the literature. Disagreements were resolved through discussion or, if necessary, by including a third researcher to make the final decision.

We have modified both **Eligibility criteria** section for better understand of our selection preference of articles based on the same data in the revised manuscript page 5 (revised text in italics):

"Studies were eligible for inclusion only if they stated the prevalence of children under 5 years with IDA in Western China. "Western China" is not a specific administrative division, but the region includes 12 provinces, autonomous regions, and municipalities: Inner Mongolia, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang (Figure 1)[14]. As prevalence data may be sourced from different study designs, we included all relevant cross-sectional studies, randomized controlled trials (RCTs), cohort studies, and published surveillance data. If results based on the same data were presented in more than one publication, results from only one publication were included."

Comment 6: Results - Table 1 is great, but I wonder if we could get a map showing prevalence by province/autonomous region. This would be very helpful perhaps as a separate panel to the existing map!

Response: We thank the reviewer for your comment. We have carefully considered your suggestion, however, as we can see from Table 1, there are fewer than 4 studies for Tibet (*a43iv*), Ningxia (*a14*, *a54v*), Guangxi (*a44*, *a45*, *a54vi*), and Chongqing (*a22*, *a54v*). Due to the limited number of studies in each of these four regions, we cannot give overall median prevalence of anemia for any of the aforementioned areas. Furthermore, we think it is not appropriate to use individual studies to represent the overall prevalence of anemia in these autonomous regions and municipality. Unfortunately, given the limited data available, we decided not to draw a map (although we fully agree that a prevalence map would have been intuitive and helpful).

We sincerely appreciate your suggestions and are looking forward to your reply.

Comment 7: Results - In Tables 1/2, if possible, to do without impacting readably too much, it would be good to have the pooled n of subjects for each category.

Response: Thank you for your comments. We have added the pooled "n" of subjects for each category in Tables 1/2 according to your suggestion. Specifically, in Table 1 we added the number of studies by province and in Table 2 have added the number of studies by age, gender, and ethnic group. And we have attached the Tables 1/2 at the end of this response letter.

Comment 8: Results - The wide ranges in reported prevalence would respond well to a metanalytic presentation I think.

Response: As JBI's Systematic Reviews reported [1-2], *"The data synthesized within a systematic review are the results extracted from individual research studies relevant to the review question. As much as meta-analysis is preferred, it is not always possible in a systematic review*

if the included studies vary greatly from each other, either in terms of how they are conducted (different interventions), who they are performed on (different populations) or in their final result. When meta-analysis isn't possible, common alternatives for the synthesis of quantitative data in a systematic review include narrative summary of results, vote counting, and presenting data via tables."

Based on our response on Comment 2, we think that a systematic review (without meta-analysis) may be more appropriate for this study.

REFERENCES:

[1] Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C. Chapter 5: Systematic reviews of prevalence and incidence. In: Aromataris E, Munn Z (Editors). *JBI Manual for Evidence Synthesis.* JBI, 2020. Available from https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-20-06

[2] Munn, Zachary PhD; Tufanaru, Catalin MD, MPH; Aromataris, Edoardo PhD JBI's Systematic Reviews, AJN, American Journal of Nursing: July 2014 - Volume 114 - Issue 7 - p 49-54
doi: 10.1097/01.NAJ.0000451683.66447.89

Comment 9: Results - The figure of prevalence over time should have confidence intervals for each point.

Figure 4: I really like this visual presentation of the qualitative findings. However, some clarification is needed. Some things here don't seem to match the text - for example the text finds no sex difference in anemia, but here in the figure sex is mentioned as associated. Similarly, I think all the items listed here need to be clarified in terms of their directionality - some are associated with reduced and others with increased anemia. I suggest editing this table so the directionality is immediately clear in all instances.

Response: Thank you for your comment. In this comment, the reviewer makes two points. The first point asks us to add confidence intervals (CI) of prevalence over time on Figure 3.<="" span="" style="font-family: "Times New Roman";">about more clarification on Figure 4. For clarity, we will respond to each point separately.

<u>In response to the first point:</u>
In Figure 3, we selected of one estimate per province-year and plotted the trend of anemia prevalence in different western areas. Since we selected the time point prevalence from each study, it was not possible to calculate CI.

However, based on your comments, we realized that the original time trend chart is not the best representation of the included data. By only showing point-in-time prevalence across regions, we do not capture the uncertainty and variability across studies. Thus, after consulting experts and referring to similar studies, we selected one prevalence per study-year,

plotting the scatter chart, and making linear regression predictions. If provinces have more than one estimate available for a particular year, to avoid double-counting in the time trend chart, we gave preference to the study with a higher quality assessment score. We have revised the Figure 3 and attached it at the end of response letter.

In response to the second point:
We are not entirely sure what the reviewer means that the text finds no sex difference in anemia. We think the reviewer may be referring to the fact that, the prevalence of anemia in boys and girls under 5 years old was similar. However, we did not perform a univariate analysis to compare the differences between the two sex groups, but only an overall descriptive statement that the medians of the two sex groups were similar.

In Figure 4, we summarized the frequency of associated factors of childhood anemia in Western China to identify research hotspots in the field. We found that a total of 6 studies reported sex has a significant influence on the prevalence of childhood anemia (without attention to directionality): *a21, a43, a45, a54, a57, a58.*

However, it is difficult to clarify the directionality of each associated factors in Figure 4 because the direction of each influencing factor varies in different studies. Taking sex as an example, of the six studies that concluded that sex was an associated factor of anemia, four studies concluded that boys were more likely to have anemia than girls (*a21, a45, a54, a58*), while the other two studies concluded the opposite (*a43, a57*). Using age group as another example, several studies supported children in 6-12 months of age are at higher risk of anemia than children younger (0-6 months) or older (>12 months) than that age group *(a11, a18, a20, a49, a54*). Therefore, it is difficult to simply assume that older (or younger) age is a risk (or protective) factor for anemia, i.e., it is impractical to give a clear directionality to age.

Comment 10: Discussion - discussion is adequate, although again I'm curious why the authors don't think they can do pooled estimates - the included studies aren't actually that heterogenous it seems (>70% are cross sectional). If authors decided to not pursue a pooled analysis, then I think this needs stronger justification.

Response: Thank you for pointing this out. We have revised the manuscript to a systematic review (without meta-analysis) according to your Comment 2. Accordingly, we have modified some of the statements in the discussion.

VERSION 2 – REVIEW

REVIEWER	Reviewer name: Dr. Sarah Nevitt Institution and Country: University of Liverpool, Biostatistics Competing interests: None
REVIEW RETURNED	14-Dec-2021

GENERAL COMMENTS	Thank you to the authors for their responses and for their efforts to revise their work as a systematic review without meta-analysis. I have a few follow-up comments 1) Editorial comment: Perhaps I am missing it but I cannot find the supplementary materials within the Scholar One system so I have not been able to look at any of the Appendix documents as part of this review. 2) Related to my original comment 5 about Quality assessment. Although I commented on a problem with the values the authors have assigned to each score within the original submission (i.e. yes = 2, unclear = 1 and no = 0), the main point of this comment was that the JBI checklist doesn't assign scores, rather than that the authors should continue to use scores with different names. I still recommend that the tool should be used as intended rather than assigning scores. In other words, the reviewers record a response of Yes, No or Unclear for 9 questions and based on these responses decide whether to include or exclude studies, and where studies are excluded, reasons should be provided for this. Using the tool as intended may still result in some studies identified in the search being excluded, but this may be a different number and set of studies compared to those currently excluded. 3) Data synthesis: Please clarify in this section that prevalence estimates per province are based on unweighted pooling (in other words, adding together the number of children with anaemia and the total number of children within each study) rather than based on any weighted meta-analysis methods. 4) Related to my original comment 7 on Figure 3. I would like to clarify that I considered that the 'aim' of this Figure was fine, it was the interpretation of the figure in terms of time trends which I considered inappropriate. I didn't intend to imply that the authors should try to demonstrate a linear trend. Selecting one estimate per study year and fitting a linear line of best fit to a scatter plot includes only a subset of the data, does not capture uncertainty and any trends found cannot be assumed to be present for the evidence base as a whole as they are based on a selected subset. Rather, for the original Figure 3, I suggest that rather joining up the time points with straight lines, the point estimates which relate to all studies should be kept, but error bars (e.g the IQR) should be added to the point estimates to demonstrate uncertainty. Authors can certainly comment on any visual changes by region and over time, but to formally start testing for trends with any accuracy (whether linear or non-linear) generally requires information from patient level datasets (e.g. large cohort or registry studies) about how prevalence change over time within the same population. Such information usually is not available from summary data in published articles.

REVIEWER	Reviewer name: Dr. Peter Rohloff Institution and Country: 2 Calle 5-43 Zona 1, Santiago Sacatepéquez, 3006, Guatemala

	Competing interests: None
REVIEW RETURNED	08-Dec-2021

GENERAL COMMENTS	Authors have comprehensively responded to the reviews.

VERSION 2 – AUTHOR RESPONSE

Dear Prof. Imti Choonara,

My coauthors and I thank you for your comments and suggestions concerning our manuscript "Anemia Prevalence and its Associated Factors in Children under 5 Years in Western China: A Systematic Review" (ID: bmjpo-2021-001185.R1). We also appreciate the thoughtful comments from the reviewers, which have much improved the paper.

We have studied the comments carefully and have revised our paper accordingly. This letter provides point-by-point responses to each comment and summarizes relevant changes in the manuscript. Primary changes to the paper include the following:

• We have reduced the number of included studies from 58 to 55 and revised the whole result section, after using the "JBI checklist" tool as intended.
• We have made modifications to Figure 3 based on reviewer's comment, removing the straight lines, and adding the confidential interval of each point estimates of prevalence.

All changes are highlighted using a track change function in the revised manuscript (marked copy). Please see the attachment for the details of the response letter.

The material in the manuscript has not and will not be submitted elsewhere for possible publication as long as it is under consideration by BMJ Paediatrics Open.

Once again, we are very grateful for your consideration of our work. We look forward to receiving your feedback on the revised manuscript.

Sincerely.

Zhou Huan

Response to Reviewer 1
General Comment: Authors have comprehensively responded to the reviews.

Response to General Comment: Thank you for recognizing our efforts!

Response to Reviewer 2
General Comment:
Thank you to the authors for their responses and for their efforts to revise their work as a systematic review without meta-analysis.

I have a few follow-up comments.

Response to General Comment: Thank you for recognizing our efforts. And we've provided point-by-point responses to each comment and revised relevant changes in the manuscript.

Comment 1: Editorial comment: Perhaps I am missing it but I cannot find the supplementary materials within the Scholar One system so I have not been able to look at any of the Appendix documents as part of this review.

Response: I think there maybe something wrong with the Scholar One system. To avoid the possibility of a similar situation later, we've attached all supplementary materials at the end of this response letter.

Comment 2: Related to my original comment 5 about Quality assessment.
Although I commented on a problem with the values the authors have assigned to each score within the original submission (i.e. yes = 2, unclear = 1 and no = 0), the main point of this comment was that the JBI checklist doesn't assign scores, rather than that the authors should continue to use scores with different names.
I still recommend that the tool should be used as intended rather than assigning scores. In other words, the reviewers record a response of Yes, No or Unclear for 9 questions and based on these responses decide whether to include or exclude studies, and where studies are excluded, reasons should be provided for this.

Using the tool as intended may still result in some studies identified in the search being excluded, but this may be a different number and set of studies compared to those currently excluded.

Response: Thanks for your comment. We were convinced and agreed with your comment. In the revised manuscript, the JBI checklist was used as intended instead of assigning scores. In other words, we decide whether to include or exclude the studies based on the responses to the 9 questions, and we excluded studies if fewer than 6 of the criteria were achieved [1, 2].

Of course, after using the tool as intended, the number of included studies was reduced from 58 to 55 (4 studies that were originally included were removed and 1 new study was included). Therefore, we rectified all the results, using a track change function in the revised manuscript (marked copy). The details quality scores of studies were provided in Appendix Table 2, attaching at the end of this response letter. The revision of "Quality assessment" was in the METHODS section of the paper, page 6.

"Assessment of methodological quality
We used standardized forms from Joanna Briggs Institute Critical Appraisal Checklist for Studies Reporting Prevalence Data17 to determine the methodological quality of included studies. Two researchers (YF and FQ) independently evaluated 9 methodological items of study design, conduct, and analysis for each included study. Each item has four choices: yes, no, unclear or not applicable. One point is assigned to a 'yes' response, and the quality score is the sum of the 9 items, ranging from 0 to 9, with a higher score indicating a lower risk of bias. Researchers then discussed and made a final decision, excluding studies whose scores were less than 6."

REFERENCES:
[1] Ofori-Asenso R, Chin KL, Mazidi M, Zomer E, Ilomaki J, Zullo AR, Gasevic D, Ademi Z, Korhonen MJ, LoGiudice D, Bell JS, Liew D. Global Incidence of Frailty and Prefrailty Among Community-Dwelling Older Adults: A Systematic Review and Meta-analysis. JAMA Netw Open. 2019 Aug 2;2(8):e198398. doi: 10.1001/jamanetworkopen.2019.8398.
[2] Luo W, Zhong BL, Chiu HF. Prevalence of depressive symptoms among Chinese university students amid the COVID-19 pandemic: a systematic review and meta-analysis. Epidemiol Psychiatr Sci. 2021 Mar 26;30:e31. doi: 10.1017/S2045796021000202.

Comment 3: Data synthesis: Please clarify in this section that prevalence estimates per province are based on unweighted pooling (in other words, adding together the number of children with anaemia and the total number of children within each study) rather than based on any weighted meta-analysis methods.

Response: Thanks for your comment. We highly agreed with your opinion and have revised the METHODS section on pages 6-7 (revised text in italics):

"Data synthesis
The analysis consisted of four steps: (1) calculation of anemia prevalence estimate in children under 5 years of age per province (distinguishing between urban and rural areas if reported), using the median percentage with IQR; (2) stratification of prevalence estimates by sex, age, and ethnic group, separately; (3) collation of factors associated with childhood anemia; and (4) extraction the point

estimates of prevalence (with confidential interval) and plotting by year in different provinces of the studies. The prevalence estimates are calculated based on unweighted pooling rather than based on weighted meta-analysis methods. Data were analyzed with Stata version 16.0."

Comment 4: Related to my original comment 7 on Figure 3.
I would like to clarify that I considered that the 'aim' of this Figure was fine, it was the interpretation of the figure in terms of time trends which I considered inappropriate. I didn't intend to imply that the authors should try to demonstrate a linear trend.

Selecting one estimate per study year and fitting a linear line of best fit to a scatter plot includes only a subset of the data, does not capture uncertainty and any trends found cannot be assumed to be present for the evidence base as a whole as they are based on a selected subset.

Rather, for the original Figure 3, I suggest that rather joining up the time points with straight lines, the point estimates which relate to all studies should be kept, but error bars (e.g the IQR) should be added to the point estimates to demonstrate uncertainty.

Authors can certainly comment on any visual changes by region and over time, but to formally start testing for trends with any accuracy (whether linear or non-linear) generally requires information from patient level datasets (e.g. large cohort or registry studies) about how prevalence change over time within the same population. Such information usually is not available from summary data in published articles.

Response: Thank you for your comment. We have removed the straight lines and added the CI (confidential interval) of each point estimates of prevalence on the original Figure 3. And according to your comment, we changed the "figure legends" from "Time Trend of Anemia Prevalence" to "Point Estimates of Prevalence by Year". We have revised the Figure 3 and attached at the end of response letter. The revision of "Data synthesis" in the METHODS section on pages 6-7 (revised text in italics):

"Data synthesis
The analysis consisted of four steps: (1) calculation of anemia prevalence estimate in children under 5 years of age per province (distinguishing between urban and rural areas if reported), using the median percentage with IQR; (2) stratification of prevalence estimates by sex, age, and ethnic group, separately; (3) collation of factors associated with childhood anemia; and (4) extraction the point estimates of prevalence (with confidential interval) and plotting by year in different provinces of the studies. The prevalence estimates are calculated based on unweighted pooling rather than based on weighted meta-analysis methods. Data were analyzed with Stata version 16.0."