PEER REVIEW HISTORY

BMJ 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)

Protocol for an Analytical Cross-Sectional Study on the Association of Iron Deficiency Anemia with Dental Caries in the Permanent First Molars of Children Aged 7 to 12 Years in Karachi, Pakistan.

Authors

Syed, Sadaf; Azam, Iqbal; Jamalvi , Waseem Ahmed; Khan, Farhan Raza; Saleem, Ali Faisal; Tikmani, Shiyam S

VERSION 1 - REVIEW

Reviewer **1**

Name **Zameer, Mohammed**

Affiliation **Dental Department, Armed Forces Hospital, Abu Arish, Jazan, Saudi Arabia**

Date **04-Sep-2024**

COI **No**

The writer has considered all the aspects well.

Reviewer **2**

Name **Rahman, Syed**

Affiliation **Uppsala Universitet, Global Health and Migration Unit, Deaprtment of Women's and Children's Health**

Date **22-Sep-2024**

COI review **I declare that I have no conflict of interest related to this review**

This paper addresses an important topic, particularly relevant to current global public health challenges. However, I have identified a few observations and comments that I believe should be considered to further strengthen the study, as outlined below:

1. **Title:** I would like to suggest a revision of the title to better reflect the nature of the manuscript as a protocol paper. The current title, *"Association of iron deficiency anemia with dental caries in the permanent first molars of children aged 7 to 12 years: An analytical cross-sectional study in Karachi, Sindh, Pakistan,"* presents itself as a completed study rather than a protocol for a future study. To make it more clear that this is a protocol paper, consider revising the title to something like:

"Protocol for an analytical cross-sectional study on the association of iron deficiency anemia with dental caries in the permanent first molars of children aged 7 to 12 years in Karachi, Pakistan."

2. **Overall language clarity:** The manuscript can be improved in terms of word choice, tense consistency, flow, and grammatical accuracy. The sentence structures throughout the manuscript can be enhanced to improve clarity.

- Repetitions: There are repetitive phrasing, particularly in the case of informed consent and assent, which are mentioned across multiple sections. The topic of informed consent is mentioned several times across different sections and pages, like on pages 6, 7, 10, and again on page 22 under 'Ethics'. It should be enough to include it only in the Ethics section.

- Abbreviations: For conciseness and ease of reading, consider introducing an abbreviation after the first mention of 'Dr. Ruth K.M. Pfau, Civil Hospital Karachi, Sindh, Pakistan.' For instance, you could shorten it to 'Civil Hospital Karachi (CHK)' in subsequent references.

3. Methods and analysis:

Authors mention that data cleaning and detection of missing variables will be handled by the 'complete-subject analysis approach'. It would be better to add a reference for the complete-subject analysis approach - Page 18

Although there is a long paragraph on the selected hospital, it would be more beneficial to explain the patient flow in the paediatric outpatient department. I have identified some concerns about the sampling strategy, patient recruitment, and the practicality of conducting all procedures within a tight timeframe.

- What was the rationale/reasoning behind including the **paediatric emergency ward** in the recruitment process? The paediatric emergency ward typically deals with acute

cases, which may not reflect the general population of children with iron deficiency anaemia or dental caries.

- Can all procedures, including patient selection, lab tests, and oral examinations, be realistically completed in a single visit? Will **follow-up visits** be required for patients to collect ferritin test results or for further assessments? Ferritin test results typically take a few days to be processed, while CRP results can be available more quickly. The study does not clarify any information regarding the timeline and follow-up issue regarding this procedure.

- Authors mentioned 141 children will be selected from aged 7 to 12 years visiting physicians in the pediatric outpatient department and emergency room of Dr. Ruth K.M Pfau. It is important to know the **patient flow** to understand how many patients the outpatient and the emergency department might get in one single day,

- Are the researchers going to take all children from the paediatric outpatient and emergency departments in one day or over multiple days? How will they ensure that the sample represents the broader paediatric population?

Confusion Regarding Inclusion, Exclusion Criteria, and Sampling:

- There are confusions regarding the inclusion criteria. Is the key requirement is that children reside with an adult family member in the same household, or is it that they are prescribed C-reactive protein (CRP), ferritin, and complete blood count (CBC) tests by a physician? The authors should make it clear how both criteria work together, and if both are equally important for participant inclusion?

- There is an overlap between purposive sampling and the inclusion criteria in this protocol. How does purposive sampling differ from the stated inclusion criteria?

- It is also better to mention why is it necessary for both the child and the accompanying adult to reside in the same household. How do these criteria contribute to the study's objectives?

- In page seven, it was mentioned that “If deemed eligible, these children and the adult family members will be invited for participation in the study”. It is not clear why did authors include family members? – page 7

- The **inclusion criteria** stated that both children and their accompanying adults must be able to communicate **effectively** in **Urdu.** Although Urdu is the national language of Pakistan, languages like Sindhi, Pashto, or Balochi are also used in Karachi district of Sindh province. Thus, in my point of view, the reasoning for restricting participation to those who can only communicate effectively in Urdu should have been explicitly stated. Also, wouldn’t the study be more **inclusive** by accommodating speakers of other regional languages? or by employing translators? Page 6 and 7. Thus, the inclusion and exclusion criteria should be revised and edited to avoid the above-mentioned confusions.

4. Data collection method and tools -

Authors define iron deficiency anaemia (IDA) as the presence of anaemia along with ferritin levels less than 15 µg/L. Also considered, factors that might affect ferritin levels, such as inflammation or infection, as ferritin is an acute-phase reactant and may be elevated during these conditions. This is an appropriate criterion. Since study focuses on children aged 7-12 years, it may be helpful to unify the haemoglobin thresholds for this specific age group. For example, 11.5 g/dL threshold for children 7-12 years. Clarifying this will avoid confusion during data collection and interpretation. -Page 9

For accuracy and consistency, it is recommended to use standardized equipment for height and weight measurements. Instead of an inch tape and a bathroom weighing machine, consider using a stadiometer for height and a calibrated digital scale for weight. These tools will provide more precise measurements, which is important for calculating BMI accurately. Also, better to mention manufacturer information. - Page 14

Unfortunately, I am not confident enough to comment on oral examinations because I am not a dental professional. Moreover, the phrasing in these sections can be streamlined for clarity and flow. (page 14 – 15).

Authors should clarify the rationale behind drawing different blood volumes (2 ml vs. 6 ml) and the use of EDTA vs. gel tubes. It may also be worth specifying the expected timeframe for sample transportation and analysis, as well as the conditions under which the samples will be maintained during transit, especially in relation to the use of ice packs. In which laboratory will they analyze, and what is the quality standard? -Page 16, 17

The paragraph on page 18 that starting with "To protect confidentiality… using erasure software" could be more appropriately placed under the **Ethics** section for better flow and organization.

Although it is mentioned on page 6 that data collection will take place over a one-month period from mid-September to mid-October 2024, a detailed timeline chart outlining the study period including recruitment and data collection phases would help readers better understand the study's feasibility. The protocol mentions on page 18 that "pre-testing of the study tools will be conducted on 5% of the study sample." Is the pilot study included in the same time plan?

Authors plan to collect information on **smokeless tobacco use** among children, which is a very sensitive issue in South Asia. Although tobacco use might have relevance to the study due to its potential impact on oral health, the necessity of including this question in the study needs further justification. It is important to clarify whether it is mentioned in the consent form and how the ethical considerations are addressed. Consent from parents is essential for this question, and it should be clear how the investigators will manage this process ethically. Page 13, 18, 19

5. Result: As it is a protocol manuscript there will be no result. However, it would better to add a section on expected results.

6. Questionnaires: There are several sets of questionnaires that need to be filled out, covering topics such as socio-demographic information, oral hygiene-related behaviour, medical history, and dietary information.

- Who is responsible for completing these questionnaires—will it be the physician, outpatient nurses, or someone else? How long will it take to complete these forms?
- In the paediatric outpatient department, if 100 children are seen per day, is it feasible to complete these questionnaires for all the children in a single day? Is it realistic to expect busy hospital staff or patients to complete them, especially in a high-volume setting like the paediatric outpatient department?
- Is it practical to administer and filled these questionnaires in the emergency ward?

Reviewer	**3**
Name	**Aguirre-Ipenza, Rubén**
Affiliation	**Universidad Continental**
Date	**14-Oct-2024**
COI	**I have no competing interests to declare.**

The protocol proposes using Cox proportional hazards regression to analyze the data from a cross-sectional study. This method is suitable for survival analysis where the time until the occurrence of an event is evaluated, which is not appropriate in cross-sectional studies. It is recommended to use statistical methods appropriate for cross-sectional data, such as Poisson regression with robust variance or log-binomial regression, which allow for the direct estimation of prevalence ratios. The manuscript presents grammatical and writing errors that could make it difficult for readers to understand. It is suggested to conduct a review of the English language.

VERSION 1 - AUTHOR RESPONSE

S. #	Reviewer's comments	Author's response
	Reviewer 1 comment ***Dr. Mohammed Zameer, Dental Department, Armed Forces Hospital, Abu Arish, Jazan, Saudi Arabia***	
1.	The writer has considered all the aspects well.	Thank you for your positive feedback. I appreciate your recognition of the efforts made to address all aspects of the study.
	Reviewer 2 comments ***Dr. Syed Rahman, Uppsala Universitet***	
1.	Title: I would like to suggest a revision of the title to better reflect the nature of the manuscript as a protocol paper. To make it more clear that this is a protocol paper, consider revising the title to something like: "Protocol for an analytical cross-sectional study on the association of iron deficiency anemia with dental caries in the permanent first molars of children aged 7 to 12 years in Karachi, Pakistan."	Thank you for your feedback. The protocol title is updated as suggested.
2.	The manuscript can be improved in terms of word choice, tense consistency, flow, and grammatical accuracy. The sentence structures throughout the manuscript can be enhanced to improve clarity.	Thank you for your valuable feedback regarding the language and structure of the manuscript. We have revised the grammar, flow, ar sentence structure to the best of our ability, aiming to improve clar and overall quality.

3.	Repetitions: There are repetitive phrasing, particularly in the case of informed consent and assent, which are mentioned across multiple sections. The topic of informed consent is mentioned several times across different sections and pages, like on pages 6, 7, 10, and again on page 22 under 'Ethics'. It should be enough to include it only in the Ethics section.	The informed consent and assent process plays a critical role in bot the inclusion criteria and the screening procedures, which is why it was referenced in multiple sections to highlight its importance in participant eligibility and data collection. However, I understand th concern regarding repetition and have made adjustments to reduce redundancy. The relevant changes have been made within the **'Ethics and Dissemination' section on page 22, 'Inclusion Criteria'** **on page 7 and 'Data Collection Methods and Tools' on page 11** for clarity.
4.	Abbreviations: For conciseness and ease of reading, consider introducing an abbreviation after the first mention of 'Dr. Ruth K.M. Pfau, Civil Hospital Karachi, Sindh, Pakistan.' For instance, you could shorten it to 'Civil Hospital Karachi (CHK)' in subsequent references.	Thank you for the suggestion. I have introduced abbreviations whe appropriate to improve readability and conciseness, including the u of 'Civil Hospital Karachi (CHK)' and 'CHK' for subsequent references. I hope these revisions contribute to enhancing the overa clarity and flow of the manuscript.
5.	Although there is a long paragraph on the selected hospital, it would be more beneficial to explain the patient flow in the paediatric outpatient department.	Thank you for your thoughtful recommendation. I have made the necessary amendments to the protocol accordingly under **'Study si and setting' on page 5 & 6**.
6.	What was the rationale/reasoning behind including the paediatric emergency ward in the recruitment process? The paediatric emergency ward typically deals with acute cases, which may not reflect the general population of children with iron deficiency anaemia or dental caries.	Initially, the pediatric emergency room (ER) was included in the recruitment process to help expedite data collection, as it operates 24/7, unlike the pediatric outpatient department (OPD), which only functions during daytime hours. Given the high patient volume in t OPD, many cases are referred to the ER after OPD hours, making i potentially valuable recruitment site. To align with the study's focu we only considered P2 and P3 cases in the ER, which are less seve and more representative of the general population, rather than acute emergency cases. However, we have since revised the study to limit data collection t the pediatric OPD. This decision was made because ER physicians rely on the Central Laboratory for blood tests, which provides faste results needed to expedite patient management. To ensure consistency and validity by using a single laboratory (DOW lab), th ER was excluded from the recruitment process.
7.	Can all procedures, including patient selection, lab tests, and oral examinations, be realistically completed in a single visit? Will follow-up visits be required for patients to collect ferritin test results or for further assessments? Ferritin test results typically take a few days to be processed, while CRP results can be available more quickly. The study does not clarify any information regarding the timeline and follow-up issue regarding this procedure.	Thank you for your insightful comments. In response to your concerns: No follow-up visits are required from patients in this study. The fir author (Dr. Sadaf Syed) is covering the cost of the lab tests, which allows both the patient and the researcher to access the test results directly. Typically, the results for CBC, CRP, and ferritin tests fror DOW lab are available within 12- 48 hours. All procedures, including patient selection, lab tests, and oral examinations, can be realistically completed during a single visit. While these procedures take some time, they are feasible to comple within one day. To ensure patient comfort and to facilitate data collection, informed consent is obtained in advance, addressing bot ethical considerations and patient convenience. These details have been added to the protocol under '**Recruitment participants' on page 9**.

8.	Authors mentioned 141 children will be selected from aged 7 to 12 years visiting physicians in the pediatric outpatient department and emergency room of Dr. Ruth K.M Pfau. It is important to know the patient flow to understand how many patients the outpatient and the emergency department might get in one single day	Thank you for your valuable feedback. In response to your comme I would like to clarify that the pediatric emergency room has been excluded from the study, and recruitment is now focused solely on the pediatric outpatient department (OPD). The OPD operates from 9:30 am to 2:00 pm, Monday through Saturday, and sees approximately 130 children aged 7 to 12 years each day on average. While the limited operating hours, the fact that only one data collector (the first author) is handling recruitment and the extensive selection criteria may restrict the number of participants enrolled daily, it remains feasible to recruit between 1 to 4 patients per day. This is particularly facilitated by the high volume of patients, supporting a steady and manageable recruitment process within the available timeframe. This has been added to the protocol under **'Study site and setting' on page 5 & 6**.
9.	Are the researchers going to take all children from the paediatric outpatient and emergency departments in one day or over multiple days? How will they ensure that the sample represents the broader paediatric population?	Thank you for your thoughtful comments. Data collection will occu over multiple days and months to ensure the required sample size i reached. The study has prioritized internal validity by applying stri selection criteria, ensuring that the results accurately reflect the tar population. While broader sample inclusion might increase externa validity, it could compromise internal validity, which is crucial for the study's aims. Although increasing the sample size or expanding to alternative stu sites, such as schools, could enhance representativeness, logistical and financial constraints make this impractical. Furthermore, paren are often reluctant to consent to blood tests for their children in non-clinical settings, which could further delay the study. The pediatric outpatient department (OPD) was selected as the stuc site as it offers a more diverse patient population, including healthi children (compared to hospitalized patients) and those who may no attend school, thus providing a balanced and feasible setting for recruitment. While external validity may be limited, conducting the study in this setting ensures scientific rigor and practical feasibility particularly given the timeline and ethical considerations. These details have been added to the protocol under **'Study site an setting' on page 5**.
10.	There are confusions regarding the inclusion criteria. Is the key requirement that children reside with an adult family member in the same household, or is it that they are prescribed C-reactive protein (CRP), ferritin, and complete blood count (CBC) tests by a physician? The authors should make it clear how both criteria work together, and if both are equally important for participant inclusion?	Thank you for your valuable feedback. Both criteria—prescription blood tests (CRP, ferritin, and CBC) by a physician and the requirement for the child to reside with an adult family member in the same household—are important for inclusion, though for different reasons. The blood tests must be prescribed by the physician at the study sit for ethical and practical reasons. This ensures that the tests are clinically justified, making parents more comfortable with their chi undergoing them, as they align with the child's medical needs rath than being solely research-driven. The requirement for the child to be accompanied by an adult family member who resides in the same household is necessary for data accuracy. Children visiting the pediatric OPD are not always accompanied by their parents; they may be with other relatives or

		even non-family members. However, family members who live wi the child are better positioned to provide accurate information abou the child's health and behaviors, as well as important household variables, such as monthly expenditure, which are relevant to the study. This combination of criteria helps ensure the study's feasibility whi improving the quality of the data collected, as both medical and socio-economic information is necessary for addressing the study's objectives. These details have been added to the protocol under **'Inclusion Criteria' on page 7.**
11.	It is also better to mention why is it necessary for both the child and the accompanying adult to reside in the same household. How do these criteria contribute to the study's objectives?	The rationale for requiring the accompanying adult family member reside with the child has been addressed in my response to the previous comment. I hope this provides the necessary clarification, but I would be happy to elaborate further if needed.
12.	In page seven, it was mentioned that "If deemed eligible, these children and the adult family members will be invited for participation in the study". It is not clear why did authors include family members? – page 7	I understand that the rationale for including the accompanying adul family member who resides with the child has been addressed in m earlier response. This criterion is vital for ensuring accurate information regarding the child's home environment and related factors impacting the study's objectives. These details have been added to the protocol under **'Inclusion Criteria' on page 7.**
13.	Authors mention that data cleaning and detection of missing variables will be handled by the 'complete-subject analysis approach'. It would be better to add a reference for the complete subject analysis approach - Page 18	Thank you for your valuable suggestion. To minimize missing data the first author will personally collect the data and ensure thorough completion by actively probing participants for responses. Additionally, the use of REDCap will enhance data integrity by marking all questions as 'required,' further reducing the likelihood missed responses. As a result, missing data is not anticipated, and we have therefore removed the management method for missing data from the protocol. The updated information can be found under the **'Data Management' section on page 18.**
14.	There is an overlap between purposive sampling and the inclusion criteria in this protocol. How does purposive sampling differ from the stated inclusion criteria?	Thank you for your insightful comment regarding the overlap between purposive sampling and the inclusion criteria. I would like to clarify that this study employs a consecutive sampling technique, where children are recruited in the order they visit the pediatric outpatient department (OPD). During the initial screening process, the data collector will quickly assess potential participants based on key criteria, which include: th child must be accompanied by a family member residing in the sam household; both the child and accompanying family member must understand Urdu; the child must have been prescribed relevant bloo tests; and the child should have at least one fully erupted permanen central incisor in both arches, along with fully erupted permanent first molars. Following this initial screening, a more thorough assessment will b conducted to evaluate the participants against the full set of selecti criteria. These details have been added to the protocol **under 'Methods an analysis' in 'Abstract' on page 1 and under 'Sampling techniqu and study population' on page 6.**

15.	The inclusion criteria stated that both children and their accompanying adults must be able to communicate effectively in Urdu. Although Urdu is the national language of Pakistan, languages like Sindhi, Pashto, or Balochi are also used in Karachi district of Sindh province. Thus, in my point of view, the reasoning for restricting participation to those who can only communicate effectively in Urdu should have been explicitly stated. Also, wouldn't the study be more inclusive by accommodating speakers of other regional languages? or by employing translators? Page 6 and 7. Thus, the inclusion and exclusion criteria should be revised and edited to avoid the above-mentioned confusions	Thank you for your thoughtful comment. While I acknowledge that the study could have been more inclusive by accommodating participants who speak other regional languages, the decision to limit participation to those who can effectively communicate in Urdu was made for several important reasons. Primarily, financial constraints have limited the study's budget, making it unfeasible to allocate additional resources for translators or to hire more data collectors. Furthermore, the first author, as the sole data collector, is proficient only in Urdu. This limitation means that effective communication in other local languages would be challenging, potentially affecting the quality and accuracy of the data collected. Ensuring clear communication is essential for obtaining reliable and valid information from participants, and these inclusion criteria were set to uphold the integrity of the data collection process. I hope this clarifies the reasoning behind the inclusion criteria, and I have made the necessary revisions to the protocol (under **'Inclusion criteria' on page 7**) to avoid any potential confusion.
16.	Authors define iron deficiency anaemia (IDA) as the presence of anaemia along with ferritin levels less than 15 µg/L. Also considered, factors that might affect ferritin levels, such as inflammation or infection, as ferritin is an acute-phase reactant and may be elevated during these conditions. This is an appropriate criterion. Since study focuses on children aged 7-12 years, it may be helpful to unify the haemoglobin thresholds for this specific age group. For example, 11.5 g/dL threshold for children 7-12 years. Clarifying this will avoid confusion during data collection and interpretation. -Page 9	Thank you for your helpful suggestion. In response, I have standardized the hemoglobin threshold for anemia diagnosis in children aged 7-12 years to 11.5 g/dL, as recommended. This clarification has been incorporated under the section **'Study Variables' on page 9**, which should help ensure consistency in data collection and interpretation.
17.	For accuracy and consistency, it is recommended to use standardized equipment for height and weight measurements. Instead of an inch tape and a bathroom weighing machine, consider using a stadiometer for height and a calibrated digital scale for weight. These tools will provide more precise measurements, which is important for calculating BMI accurately. Also, better to mention manufacturer information. - Page 14	Thank you for your valuable suggestions regarding the use of standardized equipment for height and weight measurements. While I fully acknowledge the importance of using a stadiometer and a calibrated digital scale for precision, certain logistical constraints at the pediatric OPD of Civil Hospital Karachi will make their implementation challenging. The OPD is often overcrowded and lacks sufficient infrastructure to accommodate these tools. The existing digital weighing scale, located at a considerable distance from the data collection area, is in constant use during operational hours, making it impractical for consistent use in the study. Additionally, there is no stadiometer available at the OPD. The only height-measuring tool available is a wall-mounted tape, which also presents logistical difficulties due to its location and the busy nature of the setting.

		For weight measurements, we plan on using a manual scale from the Tanita brand (HA621 manual weight scale), which, while not as precise as a calibrated digital scale, is the most feasible option under the current circumstances. I understand the significance of precision in height and weight measurements for BMI calculations, and we are making every effort to maintain consistency in the use of the available tools to ensure the accuracy of the data collected. As suggested, I have added the manufacturer information to the protocol under **'Physical exmaintion' on page 14.**
18.	Unfortunately, I am not confident enough to comment on oral examinations because I am not a dental professional. Moreover, the phrasing in these sections can be streamlined for clarity and flow. (page 14 – 15).	Thank you for your valuable feedback regarding the sections on oral examinations. I understand your concerns and have made the necessary revisions to streamline the phrasing for improved clarity and flow (under **'The simplified oral hygiene index' and 'DMFT/dmft index of the permanent first molar and its adjacent teeth' from page 15-17**). I appreciate your guidance and hope that these changes address your comments effectively.
19.	Authors should clarify the rationale behind drawing different blood volumes (2 ml vs. 6 ml) and the use of EDTA vs. gel tubes. It may also be worth specifying the expected timeframe for sample transportation and analysis, as well as the conditions under which the samples will be maintained during transit, especially in relation to the use of ice packs. In which laboratory will they analyze, and what is the quality standard? -Page 16, 17	Thank you for your insightful comment regarding the blood volume collection, tube types, and the sample transportation process. I appreciate the opportunity to clarify these points. All blood tests for the study participants will be conducted at the Dow Diagnostic Research and Reference Laboratory (DDRRL), located adjacent to Civil Hospital Karachi. DDRRL has been providing high-quality diagnostic services since 2007, adhering to international standards. The laboratory is known for its commitment to delivering reliable and accurate results, with internal quality assurance maintained through daily controls. Members of its medical directorate have also undergone specialized training in compliance with the latest ISO 15189:2022 standards for medical laboratories. In the study, 2 ml of venous blood will be drawn into an EDTA tube for hemoglobin analysis. EDTA tubes will be used because they prevent blood coagulation, preserving the cellular components. Additionally, 6 ml of blood (split into two 3 ml gel tubes) will be collected for CRP and ferritin analyses. Gel tubes are necessary for serum separation, which is required for these tests. After collection, the samples will be transported to DDRRL on the same day, utilizing ice packs to maintain the appropriate temperature throughout transit. The ice packs will be regularly monitored and replaced as needed to ensure the integrity of the samples is preserved. Upon arrival at the laboratory, the samples will be promptly analyzed, with results typically available within 12 to 48 hours of blood collection. I hope this explanation addresses your concerns, and I have made revisions to the protocol (under **'Laboratory-related information' on page 17**) to reflect these clarifications.
20.	The paragraph on page 18 that starting with "To protect confidentiality… using erasure software" could be more appropriately placed under the Ethics section for better flow and organization.	Thank you for your valuable feedback. The suggested revisions have been made to improve the clarity and organization of the protocol. The paragraph on confidentiality has been relocated to the **'Ethics and Dissemination' section on page 23**, to enhance the flow and ensure that all ethical considerations, including the

		protection of participant data, are presented in a cohesive and logical manner.
21.	Although it is mentioned on page 6 that data collection will take place over a one-month period from mid-September to mid-October 2024, a detailed timeline chart outlining the study period including recruitment and data collection phases would help readers better understand the study's feasibility. The protocol mentions on page 18 that "pre-testing of the study tools will be conducted on 5% of the study sample." Is the pilot study included in the same time plan?	I appreciate your observation, regarding the timeline. The pre-testing phase was not originally included in the timeline as the data collected from the pre-test was not considered part of the main study sample. However, to clarify, the pre-testing phase will be conducted over one week, from December 2nd to December 7th, 2024. Following the pre-test, participant recruitment and data collection will begin simultaneously in 8th December 2024 and are expected to be completed by mid-February 2025. After the data collection phase, data cleaning and analysis will take approximately one month. Manuscript preparation and submission are anticipated to be finalized by mid-April 2025. These details have also been added to the protocol under **'Study site and setting' on page 6**.
22.	Authors plan to collect information on smokeless tobacco use among children, which is a very sensitive issue in South Asia. Although tobacco use might have relevance to the study due to its potential impact on oral health, the necessity of including this question in the study needs further justification. It is important to clarify whether it is mentioned in the consent form and how the ethical considerations are addressed. Consent from parents is essential for this question, and it should be clear how the investigators will manage this process ethically. Page 13, 18, 19	Thank you for your valuable feedback regarding the inclusion of smokeless tobacco use among children in our study. I appreciate the opportunity to clarify the necessity and ethical considerations surrounding this sensitive topic. According to the Global Adult Tobacco Survey, approximately 7.7% of adults in Pakistan use smokeless tobacco, indicating that this practice may also be prevalent among children, particularly within lower socioeconomic groups. Given that our study primarily focuses on children from these backgrounds, it is essential to include a question on smokeless tobacco use to account for any potential confounding factors related to this variable and its impact on oral health. To ensure ethical transparency, the topic of smokeless tobacco use is explicitly addressed in both the consent and assent forms. In the consent form, it is included under the section titled "What is my child going to be asked to do?" and similarly in the assent form under "If I am in the study, what will happen to me?" This approach ensures that both parents and children are fully informed about the nature of the data being collected. The consent process also emphasizes their right to refuse participation or withdraw from the study at any time, thereby safeguarding their autonomy and ensuring ethical treatment. Furthermore, data collection will be conducted in a private space to protect participants' confidentiality. Sensitive questions, including those related to tobacco use, will be posed in a neutral and non-judgmental manner. This approach is designed to foster a comfortable environment for participants, thereby minimizing any potential discomfort associated with discussing sensitive topics and upholding the highest ethical standards in our research. I hope this clarification addresses your concerns, and I have made the necessary revisions to the protocol (under **'Study**

		variable' and 'Ethics and Dissemination' on page 10 and 22 respectively) to reflect these considerations.
23.	Who is responsible for completing these questionnaires—will it be the physician, outpatient nurses, or someone else? How long will it take to complete these forms?	Thank you for your thoughtful questions regarding the administration of the questionnaires in our study. I appreciate the opportunity to provide clarification. The first author is solely responsible for administering and completing the questionnaires as part of her thesis. Each questionnaire typically takes between 35 to 50 minutes to complete, depending on the participant's responses. Given the busy environment of the pediatric outpatient department (OPD), it would not be feasible for other staff members, such as physicians or outpatient nurses, to assist in the data collection process. The role of the physicians is limited to referring children aged 7 to 12 years who are prescribed CRP, Ferritin, and CBC tests to the first author. Their workload, coupled with the high patient volume, makes it impractical for them to engage further in data collection. Thank you again for your insights, and I hope this response adequately addresses your concerns. Moreover, I have also added made the necessary adjustments to the protocol (under **'Recruitment of participants' and 'Data collection methods and tools' on page 9 and 11 respectively**) to reflect this information.
24.	In the paediatric outpatient department, if 100 children are seen per day, is it feasible to complete these questionnaires for all the children in a single day? Is it realistic to expect busy hospital staff or patients to complete them, especially in a high-volume setting like the paediatric outpatient department?	It would not be practical to complete forms for 100 children in a single day. As mentioned in earlier response, we anticipate being able to cover 1 to 4 children per day. Additionally, considering that many patients may be uneducated and the hospital staff are often too busy, it is not realistic to expect them to fill out the forms. Therefore, the first author will personally administer and complete the questionnaires. This information can also be seen in the protocol under **'Study site and setting' and 'Recruitment of participants' on page 6 and 9 respectively**.
25.	Is it practical to administer and filled these questionnaires in the emergency ward	Thank you for your valuable feedback. I would like to clarify that the emergency ward is no longer included as a study site, so this particular concern does not apply to the current study design.
	Reviewer 3 comments ***Dr. Rubén Aguirre-Ipenza, Universidad Continental***	
1.	The protocol proposes using Cox proportional hazards regression to analyze the data from a cross-sectional study. This method is suitable for survival analysis where the time until the occurrence of an event is evaluated, which is not appropriate in cross-sectional studies. It is recommended to use statistical methods appropriate for cross-sectional data, such as Poisson regression with robust variance or log-binomial regression, which allow for the direct estimation of prevalence ratios.	Thank you for your insightful comment regarding the statistical methods proposed for data analysis in our study. I appreciate the opportunity to clarify our approach. Log-binomial regression is indeed a suitable method for directly estimating prevalence ratios in cross-sectional studies. However, it can encounter challenges such as non-convergence of parameter estimates, especially when the outcome is common or when covariates are continuous. In such instances, alternative methods like Poisson regression and Cox proportional hazards regression with robust variance estimation have been shown in the literature to be effective options, even within cross-sectional contexts. While Poisson regression typically estimates incidence rate ratios (IRR), in cross-sectional studies, we can assign a constant

		value of 1 to the follow-up time for each participant. This approach allows the IRR to approximate the prevalence ratio (PR). Similarly, although the Cox proportional hazards model is traditionally applied to longitudinal data for estimating hazard ratios, it can also be adapted for cross-sectional data. When time is assumed to be constant, the hazard ratio can approximate the prevalence ratio. Notably, previous studies have indicated that Cox proportional hazards regression with robust variance provides prevalence ratio estimates comparable to those derived from Poisson regression and the Mantel-Haenszel (MH) stratification method, with similar confidence intervals. This information can be found in this article [1] and these research articles that have used COX proportional algorithm in their respective cross-sectional studies have also been cited for reference [2, 3]. I hope this clarification addresses your concerns, and I appreciate your valuable feedback.
2.	The manuscript presents grammatical and writing errors that could make it difficult for readers to understand. It is suggested to conduct a review of the English language.	Thank you for your valuable feedback regarding the clarity and quality of the manuscript's language. I appreciate your suggestion to conduct a review of the English language, as it is crucial for ensuring that the content is easily understandable for all readers. In response to your comment, I have carefully revised the manuscript to address any grammatical and writing errors. Thank you again for your insightful comments, and I hope that the revised manuscript will meet your expectations.

Additional amendments made by the authors

1.	Changes in the categories of Parental and Child Education Status	A separate category for 'middle school education' (Grades 6 to 8) has been added for both mother and father. This will allow us to capture more granular and meaningful information regarding parental education levels. Moreover, literate but not more formal schooling and illiterate has been removed and ‘no formal education has been added instead’. Similarly, ‘illiterate’ and ‘literate but no formal schooling’ has been removed for child education level and ‘no formal schooling’ has been added instead. This change has been updated in the **study questionnaire on page 1 & 2**.
2.	Addition of 'Once a Week' in Tooth Brushing Frequency	The category 'once a week' has been included in the tooth brushing frequency section, which aligns with the WHO oral health survey. It bridges the gap between '2 to 3 times a month' and '2 to 6 times a week', providing a more accurate reflection of participants' habits. This change has been updated in the **study questionnaire on page 3.**
3.	Inclusion of 'Dandasa' in Other Oral Hygiene Practices	‘Dandasa' and its related categories have been added to capture regular oral hygiene habits. Many children in the lower SES strata use dandasa instead of miswak or a toothbrush, making it an important variable to include. This change has been updated in the **study questionnaire on page 4.** It has also been added to the protocol under **‘Study variables’ on page 10.**

4.	The keywords have also been edited.	The keywords have been edited and updated to reflect MeSH headings. These changes are reflected in the protocol under the **'Abstract' section on page 2**.
5.	The introduction paragraph on previous research on this topic and the proposed biological mechanism explaining the association have also been edited.	This was done to increase clarity of the paragraph. The updated content can be found under the **'Introduction' section on pages 3**.
6.	The method used for the conversion of the FFQ to a composite frequency of cariogenic food items score has been changed to Principal Component Analysis (PCA) from Factor Analysis.	The method used to convert the food frequency questionnaire (FFQ) data into a composite frequency score for cariogenic food items has been changed from Factor Analysis to Principal Component Analysis (PCA). This adjustment was made because Factor Analysis is typically used to identify underlying constructs or latent variables within the FFQ, which did not align with the objective of my study. Since the goal was to generate a single composite score representing the frequency of cariogenic food consumption in children (to account for in the study as a potential confounder), PCA was deemed a more appropriate method for this purpose. This score will also be further categorized into three levels (high, medium and low). This change has been updated in the protocol under **'Statistical analysis' on page 21**.

Reference:
1. Coutinho L, Scazufca M, Menezes PR. Methods for estimating prevalence ratios in cross-sectional studies. Revista de saude publica. 2008;42:992-8.
2. Peerwani G, Aijaz S, Sheikh S, et al. Predictors of non-obstructive coronary artery disease in patients undergoing elective coronary angiography. Global Heart. 2023;18(1).
3. Lakhdir MPA, Rozi S, Peerwani G, et al. Effect of parent-child relationship on physical aggression among adolescents: Global school-based student health survey. Health psychology open. 2020;7(2):2055102920954715.

VERSION 2 - REVIEW

Reviewer **2**

Name **Rahman, Syed**

Affiliation **Uppsala Universitet, Global Health and Migration Unit, Deaprtment of Women's and Children's Health**

Date **27-Nov-2024**

COI

Thanks to the authors for addressing the feedback in their revised protocol. The overall quality and clarity of the protocol have significantly improved, making it more impactful. The explanations are now well-articulated, and the structure is more cohesive, allowing for a smoother reading experience. Especially, changing the title to the suggested one and

updating the protocol to include smokeless tobacco use on the 'Ethics and Dissemination' page proved to be highly effective. However, avoiding run-on sentences by breaking longer ones into shorter, simpler sentences would further enhance the protocol's readability. For example, the sentence on page 7: "Both must understand Urdu, as the first author, the sole data collector, is proficient only in this language, and hiring translators or additional data collectors is not feasible due to financial constraints," can be rewritten as: "Both must understand Urdu because the first author, who is the sole data collector, is proficient only in this language. Hiring translators or additional data collectors is not feasible due to financial constraints."

That being said, the progress made is commendable. The authors have responded to every piece of feedback humbly, and the revisions reflect their dedication to refining the work.

Reviewer	**3**
Name	**Aguirre-Ipenza, Rubén**
Affiliation	**Universidad Continental**
Date	**03-Dec-2024**
COI	

I appreciate your detailed response and the references supporting the applicability of the Cox proportional hazards model in cross-sectional studies. While this approach can be adapted by assuming a constant time to estimate prevalence ratios, it introduces conceptual limitations, as the model was originally designed to analyze time-dependent events. This could lead to confusion among readers who are less familiar with this adaptation.

I understand the challenges associated with the log-binomial regression, such as convergence issues in certain cases. However, Poisson regression with robust variance is a widely validated alternative for cross-sectional studies, offering direct prevalence ratio estimates with greater interpretive clarity and methodological alignment.

Given the cross-sectional design of the study, I suggest prioritizing Poisson regression with robust variance for the analysis. If you choose to retain the Cox model, it would be essential to include in the protocol an explicit justification of how this model is adapted to estimate prevalence ratios, supported by relevant methodological references, to ensure methodological clarity and rigor.

VERSION 2 - AUTHOR RESPONSE

S. #	Reviewer's comments	Author's response
	Reviewer 2 comment ***Dr. Syed Rahman, Uppsala Universitet***	
1.	Thanks to the authors for addressing the feedback in their revised protocol. The overall quality and clarity of the protocol have significantly improved, making it more impactful. The explanations are now well-articulated, and the structure is more cohesive, allowing for a smoother reading experience. Especially, changing the title to the suggested one and updating the protocol to include smokeless tobacco use on the 'Ethics and Dissemination' page proved to be highly effective. However, avoiding run-on sentences by breaking longer ones into shorter, simpler sentences would further enhance the protocol's readability. For example, the sentence on page 7: "Both must understand Urdu, as the first author, the sole data collector, is proficient only in this language, and hiring translators or additional data collectors is not feasible due to financial constraints," can be rewritten as: "Both must understand Urdu because the first author, who is the sole data collector, is proficient only in this language. Hiring translators or additional data collectors is not feasible due to financial constraints." That being said, the progress made is commendable. The authors have responded to every piece of feedback humbly, and the revisions reflect their dedication to refining the work.	Thank you for your thoughtful feedback and for acknowledging the improvements in the revised protocol. In response to your valuable suggestion regarding run-on sentences, we carefully reviewed the protocol and made further refinements to simplify and break down longer sentences for improved readability. We are grateful for your constructive comments, which have greatly contributed to enhancing the quality of our protocol. Thank you for your kind words and for recognizing our efforts in addressing the feedback.
	Reviewer 3 comment ***Dr. Rubén Aguirre-Ipenza, Universidad Continental***	
1.	I appreciate your detailed response and the references supporting the applicability of the Cox proportional hazards model in cross-sectional studies. While this approach can be adapted by assuming a constant time to	Thank you for your thoughtful and constructive feedback regarding our statistical analysis methods. I appreciate your acknowledgment of our previous response and your insights into the applicability of the Cox proportional hazards model in cross-sectional studies.

	estimate prevalence ratios, it introduces conceptual limitations, as the model was originally designed to analyze time-dependent events. This could lead to confusion among readers who are less familiar with this adaptation. I understand the challenges associated with the log-binomial regression, such as convergence issues in certain cases. However, Poisson regression with robust variance is a widely validated alternative for cross-sectional studies, offering direct prevalence ratio estimates with greater interpretive clarity and methodological alignment. Given the cross-sectional design of the study, I suggest prioritizing Poisson regression with robust variance for the analysis. If you choose to retain the Cox model, it would be essential to include in the protocol an explicit justification of how this model is adapted to estimate prevalence ratios, supported by relevant methodological references, to ensure methodological clarity and rigor.	We understand your concerns about the conceptual limitations of adapting the Cox model for prevalence ratio estimation, particularly given its original design for time-dependent events. Your observation about potential confusion among readers unfamiliar with this adaptation is well noted. In line with your recommendations, we have decided to prioritize Poisson regression with robust variance for our analysis. This approach provides direct prevalence ratio estimates that align methodologically with our cross-sectional study design while enhancing interpretability and clarity. Thank you once again for your valuable feedback, which has been instrumental in strengthening our study methodology. The relevant changes have been incorporated under **'Methods and analysis' and 'Statistical Analysis' on page 1 and page 21 respectively** of the protocol manuscript.