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)	Age, sex and ethnic differentials in the prevalence and control of epilepsy among Sri Lankan children: a population-based study
AUTHORS	Wanigasinghe, Jithangi; Arambepola, Carukshi; Murugupillai, Roshini; Chang, Thashi

VERSION 1 – REVIEW

REVIEWER	Reviewer name: Tan CT Institution and Country: Division of Neurology, University of Malaya, Kuala Lumpur, Malaysia Competing interests: None
REVIEW RETURNED	14-Jan-2019

GENERAL COMMENTS	This is a worthwhile effort to determine the prevalence of children with epilepsy in an area of mixed population in Sri Lanka. However, a number of issues need clarification. 1. As the district chosen for the study, Ampara was badly affected by the civil war, was the population stable during the period of study? Could the civil unrest affect the result of the study? 2. To enable the readers appreciate the population under study, can you briefly describe the socio-economic background of the Ampara population? Is it largely rural and agricultural? If agriculture, what is the main crops planted, rice, tea, rubber, vegetable, or animal husbandry? What is the income level and employment rate of the study area, as it may affect the attendance of schools and treatment. 3. As mentioned by the authors, the prevalence of epilepsy in the developing countries is affected by perinatal injury, head trauma, and neurocysticercosis. Can the authors address these factors in the discussion? How about consanguineous marriage, is this practiced in any of the three groups studied, which may affect the prevalence? 4. Is there any theoretical reason why the three groups should have different prevalence of epilepsy? Is there racial (genetic) difference between the Tamils (Hindus), Sinhalese (Buddhist) and Muslim, or the difference is only in the religious belief (cultural). 5. Is your questionnaire efficient in capturing patient whose seizures manifestation is predominantly that of impaired awareness? 6. As mentioned in the limitation of the study, for children 6-16 years, there could be drop out from school due to stigma and parental/teachers' misconception. As Sri Lanka has free health care, did you double check your result with the data from the health clinic or hospital records, whether there are some children followed up in these clinics who are not attending schools, thus not captured in your study. 7. Is there previous study on the epilepsy KAP of the rural population of Sri Lanka? Is the use of complementary medicine for epilepsy common in Sri Lanka?

	How about Ayuvedic medicine, use of herbs, or forms of alternative medicine or treatment including spiritual healers. Can these affect your results? Please comment in the discussion. 8. The issue of treatment gap is not mentioned in the paper. 9. How about the incidence of "hot water epilepsy", said to be common in neighboring south Indian population. Is it seen in your study population?

REVIEWER	Reviewer name: Rachel Hilliam Institution and Country: The Open University, UK Competing interests: None
REVIEW RETURNED	15-Jan-2019

GENERAL COMMENTS	This paper is a useful addition to the literature and makes a valid contribution to an area which is not only difficult to quantify, but important in terms of distribution of resources. The researchers make a good attempt at using cluster randomisation and have used appropriate statistical methods of adjustment here. However if would be helpful to the reader if the limitation of covering children only up to 16 is mentioned in the section on sample size, since this is based on the assumption that a prevalence 0.6% in 0-18 (rather than the 0-16 covered in the study). The objectives in the abstract state that the ages are split into three groups 0-5, 6-10 and 10-16, however the cluster randomisation only splits 0-5 and 6-16. I understand why this has been done due to the way that the clusters have been chosen, however since some of the results split the older group, it would be helpful to see how the distribution on age is similar in the complete sample and the population (or not). Particularly if there is a difference between the sample and population in terms of the split of ages 6-10 and 10-16. Whilst there is some explanation to ensure that there is the correct mix of ethnicity, I'm not clear this is the case for age distribution since 10 classes per school were randomly chosen. Given the numbers involved there should be enough variation for this not to bias the results so I am not concerned about this, but a clearer explanation would help the reader. Apart from this point the paper is very well written.

REVIEWER	Reviewer name: David Flood Institution and Country: Wuqu' Kawoq, Guatemala Competing interests: None
REVIEW RETURNED	15-Jan-2019

GENERAL COMMENTS	General comments: I enjoyed reviewing this well-written and succinct cross-sectional survey of epilepsy in Sri Lanka. I am a pediatrician and health services researcher who has significant clinical experience caring for children with seizure disorders in a low- and middle-income country in Latin America, so I am personally aware of the challenges and importance of the issue described in this paper. While I am not a survey methodologist, as I describe below, my main concern is that the survey methods are incompletely explained and that there could be some bias inserted in the prevalence estimates. A statistician or research better versed in survey sampling techniques may be able to better comment on this issue, but I summarize my questions below.

	Utilizing a clinic-based, neurologist-driven, peer-reviewed, and 1-year validation epilepsy diagnosis are significant strengths of this study. I would be interested in reading more about the health-services implications of some of the data presented here. For example, did most children with epilepsy receive appropriate care at an appropriate age? Were most children cared for in the public or private sector? Is it the authors impression that the generally good seizure control reported in these children is a reflection of the tendency of childhood epilepsy to remit or a result of a high-functioning health system? I ask these questions because, in my anecdotal experience in a rural setting in a Latin American LMIC, seizures and other pediatric neurologic disorders are not well recognized, understood, or manages; the experience reported here in Sri Lanka seems very different. The English-language scientific writing is quite good overall, but very light English editing from a scientific editor would be useful. Summary comments on methodologies: - I would like the authors to give more detail about the sampling methods and calculations. I am not a survey methodologist, but I was left with some questions when trying to duplicate the sample sizes estimates using Stata and an online survey sample size calculator. - How was the ICC of 0.1 assumed? - Was a statistical software or website utilized in the sample size calculations? - Were the clusters (i.e., the PHM and classrooms, respectively) assumed to be the same size in these calculations? - Why was the non-response rate of 5% assumed? This seems reasonable but perhaps could use justification. - In the first survey on pre-school children, if the PHM-based clusters are not of approximately equal size, then using simple random sampling will introduce bias, no? - The second survey on school-age children is complex and, from what I can discern, is essentially a two-stage stratified sample, with cluster = school and stratification = ethnicity. It would be important to know how many schools were in the total sample, if schools and class sizes were assumed to be approximately equal, and, if not, if probability-proportional-to-size sampling was carried out. - Are schools ethnically homogenous, and is this an assumption that is made? - What other assumptions were made in the sample size calculations? Please include enough that we reviewers could attempt to duplicate the work. - These and other survey details may need to be included in a technical appendix.

	Line-by-line comments: Page 3, line 2: The estimates from the Global Burden of Disease should be updated to the most recent available version of the GBD study. Page 4, line 19: I would recommend separating elements of the methods into descrete sections such as survey methodologies, data collection, analysis, ethics, etc. Page 4, line 21: What is the population of this district and precise ethnic proportions, if these data exist? Page 4, line 44: An unstated assumption in the methods here is that the rate of epilepsy would be the same in preschool vs. school age children. While this is a reasonable assumption, in my view, it should be explicitly justified. Page 4, line 51: Please define more clearly that, as terms used in the paper, "preschool" = 0-5 and "school-age" = 6-16. This is done somewhat in this sentence, but could be more clear. Page 5, line 57 How many neurologists were involved in this process? Were they pediatric neurologist specialists? Page 7, line 24: I do not understand why 6-16 age group is separated within the results. It seems the most intuitive groupins are 0-5 and 6-16, since these represent separate surveys. Page 8, line 3 onward: Comparisons between groups or seizure types should have p-values associated with the comparison to test for significance. No p-values are reported. Additionally, the statistical technique used to test for significance should be reported in the methods. Page 9, line 21: Some speculation about the ethnic disparities in epilepsy prevalence is merited. Are Sinhala generally poorer or more wealthy than other groups, for example? Do Tamil groups have greater risk for neonatal insults leading to epilepsy? Page 10, line 8. Please cite the statement that the burden of epilepsy is higher in LICs than HICs. This may include the same citations from the introduction, but would still be important to include. Page 11, line 44: This is a serious limitation in the school-aged sample and am glad it is addressed. Other limitations to to be mention might include the setting of a single district, more detail on lack of generalization to rural areas, and lack of data on seizure etiology (i.e., genetic, congenital, idiopathic, due to an episode of meningitis, due to hypoxic-ischemic birth injury) that would lend itself to interventions.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

This is a worthwhile effort to determine the prevalence of children with epilepsy in an area of mixed population in Sri Lanka.

However, a number of issues need clarification.

1. As the district chosen for the study, Ampara was badly affected by the civil war, was the population stable during the period of study? Could the civil unrest affect the result of the study?

This study was conducted in 2014, six years after the conclusion of this civil war. Ampara though located in the Eastern Province was not significantly involved in the war as much as its neighbouring district Batticaloa. Ampara district was never an LTTE controlled area and has been under the administrative control of the Sri Lanka Army with least civil unrest. All three communities have been living together in this district. Population has been stable with no major internal migration during or after the war and re-settlement period. Therefore, there is absolutely no impact of the past situation on the findings of the study.

2. To enable the readers appreciate the population under study, can you briefly describe the socio-economic background of the Ampara population? Is it largely rural and agricultural? If agriculture, what is the main crops planted, rice, tea, rubber, vegetable, or animal husbandry? What is the income level and employment rate of the study area, as it may affect the attendance of schools and treatment.

Included in the revised text.

3. As mentioned by the authors, the prevalence of epilepsy in the developing countries is affected by perinatal injury, head trauma, and neurocysticercosis. Can the authors address these factors in the discussion? How about consanguineous marriage, is this practiced in any of the three groups studied, which may affect the prevalence?

Revised the discussion to include in the text.

4. Is there any theoretical reason why the three groups should have different prevalence of epilepsy? Is there racial (genetic) difference between the Tamils (Hindus), Sinhalese (Buddhist) and Muslim, or the difference is only in the religious belief (cultural).

There is no plausible theoretical reason. There is no known role of religion either on the disease outcome. However, considering the higher rate of consanguineous marriages compared to the other two ethnic groups, we expected an increased incidence among the Muslim population in this region. Also, since Muslim communities advocate inter-marriages, having epilepsy could cause stigma and loss of marriage prospects, and therefore reluctant to reveal the disease status. It is also shown that their health seeking behavior is poor compared to other two ethnic groups. Though stringent criteria were used in the study to screen and confirm the diagnosis of epilepsy, it was still dependent on the reporting by parents. Owing to the above issues, they may not have revealed the illness status during the survey, and therefore a lower prevalence reported among them.

5. Is your questionnaire efficient in capturing patient whose seizures manifestation is predominantly that of impaired awareness?

The developed a screening questionnaire that had 23 items that help to screen for those with epilepsy. This item number is higher than many of the validated screening tools used for diagnosis of epilepsy in the community. We pretested this questionnaire among 50 parents prior to its administration. The tool was completed by the parent and the questionnaire ensured that the respondent is familiar with the child's illness.

6. As mentioned in the limitation of the study, for children 6-16 years, there could be drop out from school due to stigma and parental/teachers' misconception. As Sri Lanka has free health care, did you double check your result with the data from the health clinic or hospital records, whether there are some children followed up in these clinics who are not attending schools, thus not captured in your study.

We did not cross check this. We accept that this study has the limitation of drop outs from school. The percentage increases mainly in the secondary schools and increased from 2.75% to 9.37% (grade 6 to 10) as per Education Ministry estimates. We have discussed this under limitations.

7. Is there previous study on the epilepsy KAP of the rural population of Sri Lanka? Is the use of complementary medicine for epilepsy common in Sri Lanka? How about Ayuvedic medicine, use of herbs, or forms of alternative medicine or treatment including spiritual healers. Can these affect your results? Please comment in the discussion.

There is no published study looking into this. An abstract in local Epilepsy Association describe use of these different treatment strategies. They often occur concurrent to allopathic medications.

8. The issue of treatment gap is not mentioned in the paper.

We did not consider treatment gap for few reasons. The control of epilepsy was good in the majority. The reason for poor control was related to the epilepsy rather than the medication. Secondly, there is no significant regional difference on the availability of AEDs. Being a small country, people have access to all medications.

9. How about the incidence of "hot water epilepsy", said to be common in neighboring south Indian population. Is it seen in your study population?

This not a common experience in our usual practice.

Reviewer: 2

Comments to the Author

This paper is a useful addition to the literature and makes a valid contribution to an area which is not only difficult to quantify, but important in terms of distribution of resources.

The researchers make a good attempt at using cluster randomisation and have used appropriate statistical methods of adjustment here. However it would be helpful to the reader if the limitation of covering children only up to 16 is mentioned in the section on sample size, since this is based on the assumption that a prevalence 0.6% in 0-18 (rather than the 0-16 covered in the study).

Response: There is no previous study conducted among 0-16 age group in Sri Lanka or in the region, or studies conducted separately in 0-5, 6-10 and 11-16 age groups. Therefore, in order to calculate the sample size, we selected the study closest to our requirement, which is a study conducted in Asia among 0-18 age group.

We have made this clarification in the section on sample size calculation.

Comment: The objectives in the abstract state that the ages are split into three groups 0-5, 6-10 and 10-16, however the cluster randomisation only splits 0-5 and 6-16. I understand why this has been done due to the way that the clusters have been chosen, however since some of the results split the older group, it would be helpful to see how the distribution on age is similar in the complete sample and the population (or not). \Particularly if there is a difference between the sample and population in

terms of the split of ages 6-10 and 10-16. Whilst there is some explanation to ensure that there is the correct mix of ethnicity, I'm not clear this is the case for age distribution since 10 classes per school were randomly chosen. Given the numbers involved, there should be enough variation for this not to bias the results so I am not concerned about this, but a clearer explanation would help the reader.

Response: We thank the reviewer for highlighting this point.

As quite correctly stated by the reviewer, our aim was to provide prevalence rates separately for the three age groups. However, since both 6-10 and 11-16 age groups were recruited from school settings, sampling is described as for two groups (0-5 age group and 6-16 age group). The number to be included within each age group in the sample was calculated proportionate to the age distribution of the population data of Sri Lanka (Census 2012), as given here: (1) 24.7% of the sample to represent the 0-5 age group and 75.3% of the sample to represent the 6-16 age group, and subsequently (2). Within the 6-16 age group, 60 grade 1-5 classes to represent the 6-10 age group and 90 grade 6-11 classes to represent the 11-16 age group (i.e. within each selected school, 4 classes from grades 1-5 and 6 classes from grades 6-11). Therefore, the distribution on age in the complete sample is similar to that in the population for all three age groups, and thereby there is no difference between the sample and population in terms of the split of ages 6-10 and 10-16.

To make this point clear, we have provided additional information on the split of 6-16 age group into 6-10 and 11-16 age groups proportionate to population and how the sampling was done accordingly within the selected schools (methods).

Apart from this point the paper is very well written.

Reviewer: 3

Comments to the Author

General comments:

I enjoyed reviewing this well-written and succinct cross-sectional survey of epilepsy in Sri Lanka. I am a pediatrician and health services researcher who has significant clinical experience caring for children with seizure disorders in a low- and middle-income country in Latin America, so I am personally aware of the challenges and importance of the issue described in this paper.

While I am not a survey methodologist, as I describe below, my main concern is that the survey methods are incompletely explained and that there could be some bias inserted in the prevalence estimates. A statistician or research better versed in survey sampling techniques may be able to better comment on this issue, but I summarize my questions below.

Utilizing a clinic-based, neurologist-driven, peer-reviewed, and 1-year validation epilepsy diagnosis are significant strengths of this study.

I would be interested in reading more about the health-services implications of some of the data presented here. For example, did most children with epilepsy receive appropriate care at an appropriate age? Were most children cared for in the public or private sector? Is it the authors impression that the generally good seizure control reported in these children is a reflection of the tendency of childhood epilepsy to remit or a result of a high-functioning health system? I ask these questions because, in my anecdotal experience in a rural setting in a Latin American LMIC, seizures and other pediatric neurologic disorders are not well recognized, understood, or manages; the experience reported here in Sri Lanka seems very different.

Response: Review of patient details revealed that majority presented early in the disease, all had access to EEG. When indicated majority had access to CT and in the symptomatic focal group, some have had access to MRIs. In those well controlled, the management is mainly in the public health services by a paediatrician. Those poorly controlled had been seen by a neurologist in most instances. Interestingly, many had obtained treatment/ review from neurologist in the private sector, often from the capital, Colombo. The reason for this was their reluctance to attend the local services, due to the stigma.

The authors impression is that the good seizure control described in this paper is a reflection of the natural profile of most childhood epilepsies.

Comment: The English-language scientific writing is quite good overall, but very light English editing from a scientific editor would be useful.

Summary comments on methodologies:

- I would like the authors to give more detail about the sampling methods and calculations. I am not a survey methodologist, but I was left with some questions when trying to duplicate the sample sizes estimates using Stata and an online survey sample size calculator.

We thank the reviewer for this comment. Wherever relevant, we have added the information required for better understanding of the sampling used in the study.

- How was the ICC of 0.1 assumed?

ICC is usually estimated based on the results of previous studies of similar design and subject on the degree of homogeneity. Due to the unavailability of such data, we took 0.1, as recommended in the guidelines provided by - Bennet, S., Woods, T., & Liyanage, W. M. (1991). A Simplified general method for cluster sample surveys of health in developing countries. World Health Statistics Quarterly, 44(3), 98-106.

- Was a statistical software or website utilized in the sample size calculations?

Sample size was calculated using the following formula for estimating the prevalence of an attribute (Lwanga & Lemeshow, 1991).

$n = z^2 p (100-p) / d^2$

p = Expected prevalence

d = Required level of precision

z = Required level of confidence

As cluster sampling method was used in this study, correction to the homogeneity within the cluster was added using the following formula (Abramson and Abramson, 1999).

N= Design effect x n

Design effect = $1+ \delta (b-1)$

b= cluster size

δ = intra class correlation

- Were the clusters (i.e., the PHM and classrooms, respectively) assumed to be the same size in these calculations?

PHM areas and class rooms were approximately of equal size. From these, a pre-determined number of participants were recruited as clusters (19 selected from each PHM area; 25 selected from each classroom). These details have been added to the manuscript.

- Why was the non-response rate of 5% assumed? This seems reasonable but perhaps could use justification.

This is the standard non-response rate considered in epidemiological studies.

- In the first survey on pre-school children, if the PHM-based clusters are not of approximately equal size, then using simple random sampling will introduce bias, no?

We thank the reviewer for highlighting this point.

From each PHM area, a cluster of 19 children was selected systematically using the Birth & Immunization registers maintained and routinely updated by the respective area PHM. From each selected class, a cluster of25 children was selected systematically using the student attendance register.

We have included the above sentences in the manuscript.

- The second survey on school-age children is complex and, from what I can discern, is essentially a two-stage stratified sample, with cluster = school and stratification = ethnicity. It would be important to know how many schools were in the total sample, if schools and class sizes were assumed to be approximately equal, and, if not, if probability-proportional-to-size sampling was carried out.

In fact, this was multi-stage stratified cluster sampling. PPS was not applied as classes and PHM areas were reasonably of equal size.

Initially, a total of 15 schools were selected (5 schools each from the three ethnic group strata) randomly. Within each selected school, 10 classes were selected randomly (stratified as 5 classes from grades 1-5; and 6 classes from grades 6-11). Finally, within each selected class, 25 children were selected systematically as a cluster.

Section on sampling has been revised to incorporate the above information and for better understanding.

- Are schools ethnically homogenous, and is this an assumption that is made?

Schools are classified as ‘Sinhala’, ‘Tamil’ and ‘Muslim’ schools by the ethnic group that the majority children belong to. It does not however exclude a child of a different ethnic group. Therefore, when selecting the 25 children for each cluster, children belonging to the main ethnic group of that school were only selected.

- What other assumptions were made in the sample size calculations? Please include enough that we reviewers could attempt to duplicate the work. - These and other survey details may need to be included in a technical appendix.

We have revised this section to include all relevant information necessary for duplication.

Line-by-line comments:

Page 3, line 2: The estimates from the Global Burden of Disease should be updated to the most recent available version of the GBD study.

Done. Reference changed to the GDB 2017 study.

Page 4, line 19: I would recommend separating elements of the methods into descrete sections such as survey methodologies, data collection, analysis, ethics, etc.

Is this the journal in house style?If so we are happy to change accordingly

Page 4, line 21: What is the population of this district and precise ethnic proportions, if these data exist?

Total 649,402

Sinhalese 252,458

Sri Lankan Tamils 112,457

Indian Tamils 846

Sri Lanka Moor 281,702

Burgher 1,036

Malay 187

Sri Lanka Chetty 5

Other 711

Page 4, line 44: An unstated assumption in the methods here is that the rate of epilepsy would be the same in preschool vs. school age children. While this is a reasonable assumption, in my view, it should be explicitly justified.

In the absence of prevalence data from studies conducted in 0-5, 6-10 and 11-16 age groups in Sri Lanka or in the region, the sample size was calculated for 0-16 age group, without considering the variation in the prevalence of epilepsy within it. However, given the large sample recruited within each age group, we assume that the study was adequately powered to estimate the prevalence of epilepsy stratified by the three age groups.

We have included this fact under limitations in the manuscript.

Page 4, line 51: Please define more clearly that, as terms used in the paper, "preschool" = 0-5 and "school-age" = 6-16. This is done somewhat in this sentence, but could be more clear.

We have revised this terminology to be more precise as:

Infant (0-1), young children (1-2) and pre schoolers (2-5), wherever relevant.

Page 5, line 57 How many neurologists were involved in this process? Were they pediatric neurologist specialists?

One Paedaitric Neurologist (JW)

One adult neurologist (TC)

Page 7, line 24: I do not understand why 6-16 age group is separated within the results. It seems the most intuitive groupins are 0-5 and 6-16, since these represent separate surveys.

We agree that they refer to two separate surveys, however, we feel the 6-16 age group is too broad and would be more informative if further grouped as 6-10 and 11-16 in the analysis. Further, it is easier to compare with other studies that may refer only to one age group of our sample.

Page 8, line 3 onward: Comparisons between groups or seizure types should have p-values associated with the comparison to test for significance. No p-values are reported. Additionally, the statistical technique used to test for significance should be reported in the methods.

Here, the data are described using descriptive statistics only (percentage, highest, lowest category, etc.) and not used for comparisons between groups or seizure types, hence no statistical tests and p values have been applied.

Page 9, line 21: Some speculation about the ethnic disparities in epilepsy prevalence is merited. Are Sinhala generally poorer or more wealthy than other groups, for example? Do Tamil groups have greater risk for neonatal insults leading to epilepsy?

Addressed in the discussion section and we accept that there may be a potential under-reporting in the minority communities.

Page 10, line 8. Please cite the statement that the burden of epilepsy is higher in LICs than HICs. This may include the same citations from the introduction, but would still be important to include.

Included

Page 11, line 44: This is a serious limitation in the school-aged sample and am glad it is addressed. Other limitations to to be mention might include the setting of a single district, more detail on lack of generalization to rural areas, and lack of data on seizure etiology (i.e., genetic, congenital, idiopathic, due to an episode of meningitis, due to hypoxic-ischemic birth injury) that would lend itself to interventions.

Thank you for the suggestion. We have included these aspects in the limitation section.

VERSION 2 – REVIEW

REVIEWER	Reviewer name: David Flood Institution and Country: Wuqu' Kawoq, Guatemala Competing interests: None
REVIEW RETURNED	20-Mar-2019

GENERAL COMMENTS	Thank you for this excellent resubmission. The manuscript is looking very good. A few brief comments: - Thank you for including additional information on the sampling design; this is helpful. - I would suggest even more clarity about the overall survey design. My understanding is this: the authors conducted 2 methodologically separate surveys: (1) in the 0-5 age group, a two-stage cluster sample without stratification in which all clusters (PMH) were of approximately the same size and were all were sampled; (2) in the 6-16 age group, a multi-stage stratified cluster sample. In survey (2),

	clusters = school and class (each of approximately the same size), and stratification = age and ethnicity. - With more undestanding of the sampling, I now have a concern about some of the calculations. Again, I am not an epidemiologist, but in my experience when calculating estimates from surveys I use, I have needed to account for design effects. Even if there is no weighting, design would include the stratication and clustering. Here, the authors report simple prevalence. Can the authors comment on why they chose not to incorporate design effects in each survey and corresponding confidence intervals for prevalence? Or would they prefer to use adjusted estimates, which would buttress their argument that their results reflect a representative sample? - If generated, confidence intervals could also be depicted graphically in the figures. - Regarding subheadings, especially in the methods section, I am not sure if this is a journal requirement, but using subheadings would be my recommendation (though not a requirement from this reviewer). I do think subheadings would be useful, but will defer to the authors and editor. - The added text in the results line 51-56 would be better placed in the discussion. I understand that this was added based on reviewer request. - Light typographic editing would be useful.

VERSION 2 – AUTHOR RESPONSE

Thank you for this excellent resubmission. The manuscript is looking very good. A few brief comments:

- Thank you for including additional information on the sampling design; this is helpful.

- I would suggest even more clarity about the overall survey design. My understanding is this: the authors conducted 2 methodologically separate surveys: (1) in the 0-5 age group, a two-stage cluster sample without stratification in which all clusters (PMH) were of approximately the same size and were all were sampled; (2) in the 6-16 age group, a multi-stage stratified cluster sample. In survey (2), clusters = school and class (each of approximately the same size), and stratification = age and ethnicity.

- With more undestanding of the sampling, I now have a concern about some of the calculations. Again, I am not an epidemiologist, but in my experience when calculating estimates from surveys I use, I have needed to account for design effects. Even if there is no weighting, design would include the stratication and clustering. Here, the authors report simple prevalence. Can the authors comment on why they chose not to incorporate design effects in each survey and corresponding confidence intervals for prevalence? Or would they prefer to use adjusted estimates, which would buttress their argument that their results reflect a representative sample?

As pointed out, since we adopted cluster sampling in both 0-5 and 6-16 age samples, we have added a design effect to the already calculated sample size, as a correction made for homogeneity within the clusters (Abramson & Abramson, 1999).

N = Design effect x n

Design effect = 1+ δ (b-1)

b = Cluster size; taken as 10

δ = Rho

We have already mentioned in the manuscript the design effect we have considered in both samples.

For example, design effect added to the 0-5 age group sample was 2.8, for a cluster size of 19 and rho value of 0.1 (in the absence of previous studies estimating epilepsy using cluster sampling method in Sri Lanka, as recommended by Bennett et al (1991)).

Since the cluster effect was already incorporated to the final sample, it is not preferable to adjust the estimates obtained.

If generated, confidence intervals could also be depicted graphically in the figures.

95% confidence intervals are already mentioned in the manuscript for 0-5 group and 6-16 groups.

0-5 age group= Overall prevalence rate of 73 per 10,000 children (95% CI: 37-144).

,6-16 group=52.5 per 10,000 children (95% CI: 33, 84)

(when further divided, 55.1 per 10,000 children aged 6-10 years (95% CI: 28, 108) and 50.4 per 10,000 children aged 11-16 years (95% CI: 26, 95)).

Figures drawn have also given the 95% Cis.

Regarding subheadings, especially in the methods section, I am not sure if this is a journal requirement, but using subheadings would be my recommendation (though not a requirement from this reviewer). I do think subheadings would be useful, but will defer to the authors and editor.

We will leave it for the editors to decide

- The added text in the results line 51-56 would be better placed in the discussion. I understand that this was added based on reviewer request.

Shifted to limiations.

- Light typographic editing would be useful. A few changes done.

-What this study adds

Has been changed accordingly