Author’s response to reviews

Title: Human metapneumovirus prevalence and patterns of subgroup persistence identified through surveillance of pediatric pneumonia hospital admissions in coastal Kenya, 2007-2016

Authors:

John Oketch (oketchjohn9@gmail.com)

Everlyn Kamau (EKamau@kemri-wellcome.org)

Grieve Otieno (grieveotio@gmail.com)

James Otieno (JOTieno@kemri-wellcome.org)

Charles Agoti (CNyaigoti@kemri-wellcome.org)

James Nokes (JNokes@kemri-wellcome.org)

Version: 1 Date: 22 Jun 2019

Author’s response to reviews:

Dear Editor,

We wish to submit the revised version of the manuscript entitled ‘Human metapneumovirus prevalence and patterns of subgroup persistence identified through surveillance of pediatric pneumonia hospital admissions in coastal Kenya, 2007-2016’ for consideration for publication in BMC Infectious diseases.

Following the comments from the reviewers, we have undertaken extensive revision of the initial draft. We thank the reviewers for their valuable contributions, which we believe have resulted in considerable improvement of the manuscript. The revisions undertaken include:

• Clearly defined nomenclature of HMPV subgroups

• Phylogenetic analysis

• Specific revisions throughout the manuscript as indicated by the reviewers

• Improved the quality of the figures.
We have provided a point-by-point response to the comments from the reviewers together with this resubmission.

We apologize for the late submission of the revised draft, this was not intentional.

We look forward to your consideration.

Yours sincerely,

Corresponding author,

John W. Oketch,

KEMRI - Wellcome Trust Research Programme, Centre for Geographic Medicine Research – Coast, Hospital Road,
P.O. Box 230-80108, Kilifi, Kenya; Tel: +254 (0)417 522535; Fax: +254 (0)417 522390;
Email: oketchjohn9@gmail.com

OUR RESPONSES TO REVIEWER COMMENTS

REVIEWER #1

1. No consensus has been reached on the subgrouping of HMPV strains among the researches. Each researcher uses different methods and different gene regions or length to propose new sub-clusters, such as A2c, A2b1, A2b2, or novel sub-lineage of A2b. The 'A2c' has been provisionally proposed only using the short F gene region (321-nucleotide region) and limited numbers of HMPV strains. The provisional new sub-lineage of A2b strains, including A2c strains, have been still classified into the subgroup A2b in certain studies. The variations of genotype nomenclature among researchers may confuse the situations of HMPV molecular epidemiology. First, please define clearly the nomenclature of HMPV subgroups in this study, comparing with those in other studies. This reviewer expects that subgrouping analysis is conducted systematically using as many available sequence data from database as possible (maybe in other study).
Our Response: We agree indeed that there is no definitive HMPV nomenclature in the literature. We have clarified in the introduction and methods in the revised manuscript, the classifications as reported in the literature. Also in the discussion section, we acknowledge the possible ambiguity of HMPV nomenclature. In our analysis, we adopted the term ‘A2c’ from two studies by (Jagušić M et al. 2017 and Rahman et al. 2019), and think this is fitting as the clade evidently clusters differently from A2b (refer to Figure 1 and Figure 2). Subgrouping has been repeated with more sequences as described in the methods section. We have defined nomenclature in the manuscript (line 60-67) and provided the trees in the results section.

2. The authors analyzed the full-length F and G coding regions of 114 HMPV strains (line 114). If the full-length sequence data is used, is the A2c sub-clustering still evident?

Our Response: Yes, based on F and G gene phylogenies the sub-clustering is still evident, as illustrated by Figure 1 and supplementary Figure 1.

3. Recently, unique A2b HMPV strains, which possess 180- or 111-nucleotide duplication in G gene, have been detected in several countries. Could the similar strains be detected in Kenya? When these strains are included in the phylogenic tree, where are they located in the tree?

Our Response: The strains cluster within A2c, we have highlighted the strains in the trees (Figure 1) and mention in the results (lines 187-189) and discussion (lines 295-298) sections. To our the best of our knowledge, the novel strains with the duplication have not been observed or reported in Kenya.

4. Letters in all figures are very small, nearly unrecognizable. Certain figures do not meet the criteria or requirements of BMC journal. Additional information is necessary for Figure 5.

Our Response: We have corrected this.

5. This reviewer partially disagrees with the authors' conclusion that the genotype prevalence pattern in Kilifi was similar to that of global one (lines 193-196). They show a certain similarity, but differ significantly in many points. In many years, A2a has been detected globally, but not in Kilifi. A2c has been detected every year globally. The dominance of B1 is evident in 2011
globally. Unlike in Kilifi, B1, A2c, and A2b have persisted over the 9 years globally (lines 195-196). The dominance of A2b between 2007 and 2011 and the dominance of B1 between 2012 and 2014 are less evident globally (lines 268-269).

Our Response: There is a general but not exact agreement of genotype prevalence patterns between Kilifi and other locations. We suggest that at finer (subgroup) levels there are differences due to variation in sample size or sampling methods. Our study is based on hospital surveillance, representing only a small proportion of cases in the community, and it is likely we missed some genotypes, which may have only occurred in mild, asymptomatic cases not presenting at the hospital.

We have clarified this in the revised manuscript (line 210-213).

6. If only partial short regions (345-nucleotide regions) were used for the analysis of evolutionary rate, genetic distance estimation, or selection pressures, significance of the resulting value or data in this study is quite limited.

Our Response: The authors appreciate the reviewer’s concerns that the short region limits inferences compared to longer genomic regions. In order to compare our data with what is currently available from other studies, we had to trim the F gene sequences to 354 nucleotides. For the G gene data we analysed nearly complete gene fragment (640 nucleotides), which gives better inferences on evolutionary rate, genetic distance estimation or selection pressures for the different subgroups, as this genomic region is more variable.

7. The difference in methods may significantly affect the results of genotype rate data in this study (line 307-309).

Our Response: We thank the reviewer for his comment. Subgroup assignment was done based on both F and G gene sequences. All sequences sequenced for G gene were sequenced for F gene.
Minor comments

1. Many molecular epidemiology studies have been reported to data, although the data in Africa is still limited. Line 25-26.

Our Response: We have modified the sentence to refer to the low-resource setting.

2. There are many unnecessary citations. For example, as many as 23 papers are cited to describe the general information of HMPV.

Our Response: We have corrected this, we have kept the reference to minimal.

3. The paper (reference 31) does not use the term 'A2c' (line 66 and 269). In the same paper, HMPV strains detected between 2012 and 2014 are used. Was the A2c detected in 2007 (line 66)?

Our Response: We have changed the reference, and further clarified in the discussion. Line (66)

4. Dose 'few' mean almost none or several? If it is not 'none', authors may use 'a few'. (Lines 68, 72, and 74)

Our Response: We have corrected this to mean a few.

5. Please define 714F and 500G (line 192). –revise in the manuscript document

6. Was phylogenetic tree constructed using 345- or 714-nucleotide region?

Our Response: 345bp was used for F gene and 640bp for G gene. We clarified this in the manuscript (line 150 and line 153).

7. What is 'inadequate temporal signal'?

Our Response: We have rephrased the statement.

‘The effective sample size (ESS) for all parameters failed to reach the cut-off (200) required for confirmation of convergence in BEAST analysis’

We have corrected this (line 253).

8. The reference 56 may be changed. The reference 56 shows only data in a single city, Japan, but not the global data.

Our Response: We have revised the sentence to refer to a single country. Line 289

9. Are there any supportive evidence or reports to show genotype-specific herd immunity for HMPV? (Lines 272-274)

Our Response: We are speculating on this matter in using the term ‘herd immunity’. Of course, the occurrence of the various genotypes as observed in this study may be no more than chance events. We do not have specific supportive of genotype specific herd immunity.
10. Why does reduction in bacterial pneumonia result in a reduction in viral pneumonia? (Lines 297-298)

Our Response: Studies have reported the role of infecting viruses in facilitating bacterial spread in the pathogenesis through viral destruction of respiratory epithelium. In contrast, Pneumococcal conjugate vaccine reduces pneumonia associated with respiratory viral infections, presumably by preventing superimposed bacterial coinfection (Fathima et al. 2018, Madhi et al. 2004).

11. Why do the purifying selection pressures drive virus evolution?

Our Response: Study by Austin L. Hughes and Mary A. Hughes (2007), show purifying selection appear to be more effective due to effective population and higher recombination rate. It is predicted that genomes or genomic regions with low recombination rates will in general show elevated accumulation of nonsynonymous mutations (Berlin and Ellegren 2006; Wykoff et al. 2002).

We have referenced this article in the manuscript (Line 335).

12. Globally persistent circulation is observed also for B1, A2b, and A2c (line 303-304).

Our Response: Line 303-304 discusses subgroup occurrence in Kilifi, only B2 was persistent in the region. However, we also recognise the persistence of A2b and A2c in global data. Considering our data is from a single location, persistence of B2 and diversifying selection observed within these strains is of interest.

We have clarified this in the manuscript at the beginning of the paragraph (line 333).

13. No journal name is shown in the references.

Our Response: We have corrected this.
Abstract

1. Line 34: The part you described about the prevalence of HMPV among all samples tested, you can just use 1 decimal place like other parts in the paper.

Our Response: We have changed this to 4.1% and not 4.10% (line 34).

Materials and Methods

2. Line 92-93: for the inclusion criteria, it is mentioned that those children with either severe or very severe pneumonia were enrolled in the study. Did you also do HMPV screening test for those with ARI hospitalized children without pneumonia symptom? If so, it might be good to put those data in supplementary data as it might be an interest for some readers. It might be better if you could include the definition for hospitalization in this hospital setting as well.

Our Response: Cases were defined as patients presenting with syndromic pneumonia (severe or very severe), and only cases were tested for HMPV. Controls (children without syndromic pneumonia) were not tested. The reviewer is referred to a previous study by (Owor et al. 2016) who has reported HMPV occurrence in both cases and control, in this same hospital. For admission or hospitalization criteria, the clinician recommends hospitalisation upon review at the time of patient presentation, for instance, if the child has the above-mentioned pneumonia symptoms, they will be admitted. We have included this definition in the Materials and Methods section (line 90).

3. Line 90-93: You stated that you collected throat swabs, nasal aspirates and washes, or sputum specimen. Is it consistent in terms of type of samples being collected from each study participant? For example, it may be technically difficult to collect nasal aspirate / wash or
sputum from younger children compared to adult participants. Ideally, the type of samples collected should be nearly identical among all age groups.

Our Response: Sputum samples were not part of the study. It is true that over the time period of this study there have been changes to the collection procedures. However, we would argue there is some consistency in our collection method in that all sample involved collection of material from the deep nasopharynx. Our staff are well trained and experienced in collecting samples from all ages and we do not believe the patient age to be an issue in terms of sample quality. We have revised the sentence to clarify sample collection methods (line 97-99).

4. Line 95-96: For the sample collection, it was described that nasal wash, NP, or NP/OP were collected from enrolled cases. In terms of the samples used for the HMPV screening step, was it consistent throughout the current study period?

Our Response: We used all the sample types mentioned. Samples from different sites may favour collection of one pathogen over another (eg Hammitt et al, 2011). However, as stated above the consistent aspect is that all include a sample from the deep nasopharynx.

5. Line 100-111: For the HMPV screening as well as nucleic acids extraction methods used in different study period, did you verify that detection sensitivity etc. do not statistically vary among different kit?

Our Response: Provisional (unpublished) comparison data from our studies show that the different extraction methods used vary insignificantly in the rRT-PCR cycle threshold (Ct) values. For instance, Ct values from the Viral RNA Mini kit (QIAgen, Germany) method are 0.75 fold higher than QIAcube HT kit (QIAgen, Germany).

6. Line 162 and 246: I feel rather comfortable to use the term "genotype shift" over "genotype change" since "genotype change" may probably mean that a certain genotype (that already exists) change to another genotype, which is wrong in this context.
Our Response: The authors appreciate the reviewer concerns. In the manuscript, we implied change of observable genotype prevalence or predominance and have made this clarification in the revised manuscript (lines 168 and 183). We prefer not to use “genotype shift” as it has other associated connotations such as antigenic shift.

7. Line 163-164: could you justify the reason why use chi-squared test for the clinical severity comparison among different HMPV genotypes? As shown in the table 3, the value of some cells are less than 10; for example, (n=2) in A2c (very severe) and (n=10) in B2. In this case, the result of chi-square test may not reliable as the one from fisher's exact test.

Results

Our Response: Thank you. This was an error; the reported results were for Fisher’s exact test (line 42 and 352). We have corrected this (lines 42,170).

8. Line 172-173: it was described that clinical samples were successfully collected only from 74% of those study enrolled patients. It is considerably high and needs to be mentioned in discussion section. Also, it may be necessary to state that those samples-collected group and group without clinical sample are not significantly different in terms of characteristics if that is the case.

Our Response: The reviewer raises a good point. The pneumonia surveillance study aims to recruit all eligibles. We failed to collect from 24% of the eligibles in this study. We have previously reported similar results, which results from refusals and difficulty in collecting from the most severe cases (Nokes et al, 2009). Hence, we cannot argue that the clinical component of the cases for whom samples were collected is no different from those whose samples were collected. Hence, if some HMPV variants are associated with disease severity there may be a bias in the resultant composition in our study. However, the proportion not collected shows no systematic change over time and so it seems unlikely any bias in the prevalence of HMPV groups, genotypes and clusters have changed over time.

We have mentioned this in the manuscript (line 344-352).
9. Figure 1: in figure legend, it might be informative if you could put which bootstrap value was used as cut-off. For example, those bootstrap value higher than 70 was considered to be significant etc. Same for other ML-based phylogenetic analysis.

Our Response: We have included the the bootstrap values in all trees and mentioned the cut-off in the figure legends.

10. Figure 4: I understand that cluster unique amino acid substitutions were presented in figure 5, it might be also informative to put some of those information at the branches of Figure 4 if this will not make the trees too complicated. In that way, it might be easier for readers to visually understand the uniqueness of each genetic cluster.

Our Response: Done.
We have used same variant Id.s used in figure5 (now, supplementary figure2 and 3) to label the different variants along the branches in the revised figure 6.

11. Line 217-220 & Line 232-235: have you tried to statistically compare the difference in evolutionary rates among different genotypes? If this is your interest, you can easily compare using such as Welch's test, in which it assume two groups in comparison have normal distribution, yet the standard deviation may differ between two. In case of comparison among more than two groups, it might be worth using Kruskal Wallis test for testing null-hypothesis.

Our Response: Kruskal test may not be suitable for the comparison of the differences in Evolutionary rates between the groups. Instead we used highest posterior density intervals (HPD), if they overlap the difference is not significant.
Discussion

12. Line 279-284: You mentioned about the difference in evolutionary rates among genotypes, particularly in A2b and B1. Please elaborate the evolutionary implication etc. based on this results.

Furthermore, the faster evolutionary rate in hMPV in comparison with RSV reported elsewhere? You may want to add some discussion of this result in discussion section.

Our Response: The reviewer raises very interesting question. Had to speculate from our study. However, higher polymorphism has been reported in HMPV G gene and further within the resulting from events such changes in alternative transcription termination codons and insertions and hence differences in the evolution rates (Peret et al, 2004). Other reports show A2b to have evolved fastest towards heterogeneity and hence its dominant epidemic (Li et al, 2012).

In comparison to RSV; Mutations in a gene are likely to be either neutral or positively selected. As previously reported, for the G protein of RSV, there is a strong association between neutralizing epitopes and positively selected sites (Zlateva et al, 2004). In contrast to the case for other paramyxoviruses, such as RSV, the HMPV G protein is not a major neutralizing or protective antigen (Skiadopoulos et al, 2004)

We discussed this in the discussion section (line 307 and 313).

Reference


