Peer Review Report

Review Report on PHENOTYPIC VARIATION IN DISEASE SEVERITY AMONG HOSPITALIZED PEDIATRIC PATIENTS WITH COVID-19: ASSESSING THE IMPACT OF COVID-19 IN THE EPICO STUDY

Original Article, Int. J. Public Health

EVALUATION

Q 1   Please summarize the main findings of the study.

This paper aimed to defined phenotypes of COVID-19 representation in a paediatric population from Colombia and Spain using data from more than 2000 children. They identified 5 clusters that differed in terms of age, sex, comorbidities, symptoms, and outcomes.

Q 2   Please highlight the limitations and strengths.

Strengths of the study include the large sample size representing children seen in many different hospitals in Colombia and Spain. THe manuscript is limited by a lack of detail and lack of reflection of clinical relevance of the presented findings.

Q 3   Please provide your detailed review report to the authors. The editors prefer to receive your review structured in major and minor comments. Please consider in your review the methods (statistical methods valid and correctly applied (e.g. sample size, choice of test), is the study replicable based on the method description?), results, data interpretation and references. If there are any objective errors, or if the conclusions are not supported, you should detail your concerns.

Major
Introduction
Overall, I find the introduction lacks clarity and structure, and too few studies are cited with many statements left without citations.

The third paragraph is written in present tense beginning with "As time progresses and new virus variants emerge…", which reads a bit weird, now that quite some time has past since emergency visits for COVID-19 were common. The last sentence in the paragraph states that relying on symptoms to diagnose a SARS-CoV-2 infection may not be specific enough, but the focus of this manuscript does not seem to be on diagnostics. I suggest to delete this paragraph or rewrite it to fit the purpose of the paper better.

Page 3-4, lines 55-57. The authors write that studies exist on MIS-C in children but conclude that other disease presentations have been overlooked. This is quite an unspecific statement, and based on a quick literature search, several papers exist describing clinical phenotypes of COVID-19 in children (e.g. https://pubmed.ncbi.nlm.nih.gov/36919119/, https://pubmed.ncbi.nlm.nih.gov/36891409/, https://pubmed.ncbi.nlm.nih.gov/35133437/). I suggest to do a literature research and precisely describe where the knowledge gaps are and how this study contributes to closing that gap.

Page 4, lines 61-62; I find the aim described at the end of the introduction unspecific. I suggest to elaborate on the population (e.g. age range, hospitalized with COVID-19, countries etc.) and to elaborate that the aim not only is to describe but also defined phenotypes based on cluster analysis.

Methods

Page 4, subheading "participants": Recruitment procedures are unclear. Were parents of eligible participants approached and asked to participate and then they signed a consent form? And were all eligible children included in the study between April 2020 and November 2021, or did any parents refuse to participate? Please clarify

Page 5, subheading "Data sources": It is unclear if routine data from the medical records were collected or if a standard set of variables were collected specifically for the study. Kindly clarify.

Page 5, subheading data analysis: Many readers will not have heard about the agglomerative hierarchical ward analysis, please shortly describe this method and what kind of data are entered (e.g. categorical/continuous variables), ideal number of variables, samples size restrictions, advantages and disadvantages over other cluster methods. Add references to key publications describing or using this method in a similar way.

Results

Page 6, first paragraph under results. Please describe main characteristics from table 1 – e.g. frequency of comorbidities, most common clinical manifestations. Also, I suggest to describe shortly data on median days of hospitalisation, overall mortality rate, PICU admission for the total population, not just within clusters.

Page 6, results: I miss data on cluster selection. In the methods, several methods are mentioned for calculating similarity of patients and dendrograms and elbow methods for determining number of clusters. Please include data from these analysis in the results. It would also be relevant to know if the cluster analysis yielded clearly one model or if the investigators chose between several models, argue for why the selected model was chosen.

Figure 2 nicely displays characteristics of children in the different clusters, but the resolution is quite bad and the text and numbers really small making it hard to see. I suggest to increase font, so that all information in the figure can be easily understood.

Discussion

Page 8, first paragraph under discussion: The authors write that five distinct clinical phenotypes of COVID-19 disease were identified, some of which have been previously described in other studies. Please include references for these other studies in this sentence or take out the sentence – it is anyway described in detail below.

I miss a discussion on specific clinical impact of the findings from this study, which is only mentioned shortly in the conclusion. How exactly do the authors propose to utilize the findings from the cluster analysis in the future. Do the authors plan to do specific risk classification based on different characteristics, and how could physicians in the hospitals implement this in daily care. I suggest to elaborate on this and if space is a problem, the authors could shorten the discussion part describing and discussing the findings from each cluster (page 9 and a bit of page 10).

Minor
Introduction, page 3, first paragraph: The last sentence "Thus, making challenging to effectively differentiate and classify children based on their symptoms." This sentence seems incomplete, please revise.

Page 4, first sentence under "Participants": The described age range seems a bit weird, do you mean "aged 29 days to 17 years"?

Page 5, second sentence under subheading "Data analysis": please specify what is meant by qualitative variables.

PLEASE COMMENT

Q 4 Is the title appropriate, concise, attractive?

Yes

Q 5 Are the keywords appropriate?

yes

Q 6 Is the English language of sufficient quality?

The English language is fine, but the the structure of the text and logical arguments are not always of sufficient quality.

Q 7 Is the quality of the figures and tables satisfactory?

No.

Q 8 Does the reference list cover the relevant literature adequately and in an unbiased manner?)

No. I suggest the authors do a literature review of existing studies. The introduction include in total 7 references and many statements without references.

QUALITY ASSESSMENT

Q 9 Originality — ■ ■ ■ ■ ■

Q 10 Rigor — ■ ■ ■ ☐ ☐

Q 11 Significance to the field — ■ ■ ■ ■ ☐

Q 12 Interest to a general audience — ■ ■ ■ ■ ☐

Q 13 Quality of the writing — ■ ■ ☐ ☐ ☐

Q 14 Overall scientific quality of the study — ■ ■ ☐ ☐ ☐

REVISION LEVEL

Q 15 Please make a recommendation based on your comments:

Major revisions.