PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Swissped-RECOVERY – Masked Independent Adjudication for the Interpretation of Non-randomised Treatment in a Two-arm Open-label Randomised Controlled Trial (Methylprednisolone vs Immunoglobulins) in Paediatric Inflammatory Multisystem Syndrome Temporally Associated with SARS-CoV-2 (PIMS-TS) Involving Ten Secondary and Tertiary Paediatric Hospitals in Switzerland
AUTHORS	Schöbi, Nina; Sanchez, Carlos; Welzel, Tatjana; Bamford, Alasdair; Webb, Kate; Rojo, Pablo; Tremoulet, Adriana; Atkinson, Andrew; Schlapbach, Luregn; Bielicki , Julia

VERSION 1 – REVIEW

REVIEWER	Nachtigall, Irit Helios Klinikum Bad Saarow, Infectious Diseases
REVIEW RETURNED	31-Aug-2023

GENERAL COMMENTS	My thanks to the authors for their important study and for bringing forward an equally important RCT. In my view, it is very important to address exactly the problem that was addressed here, namely the concomitant medication that was not randomised. I have only one point: I understand that the groups are already small, but I find it difficult that the data are not adjusted for gender, since it is known that the immune system already differs between the sexes in infants.

REVIEWER	Welty, Stephen University of Washington, Pediatrics
REVIEW RETURNED	06-Oct-2023

GENERAL COMMENTS	With the emergence of Covid19. a syndrome in pediatric patients has been described in which they present with evidence of an exaggerated systemic inflammatory responses. The syndrome has been labeled as Pediatric Inflammatory Multisystem Syndrome temporally associated with SARS-CoV-2 (PIMS-TS) in which the patients may present with the syndrome described and in which best method(s) of treatment have not been elucidated, the two primary approaches being systemic administration of steroids and/or intravenous immunoglobulin. The randomized open label trial has already been published (reference 6) and found no difference in the primary outcome (LOS) and did find a difference in the rate of respiratory support need after baseline (higher in the IVIG group). The authors point out that given the emergence of the syndrome and a lack of systematically acquired data that a

BMJ Open: first published as 10.1136/bmjopen-2023-078137 on 25 April 2024. Downloaded from http://bmjopen.bmj.com/ on August 3, 2024 by guest. Protected by copyright.

	pragmatic open label trial is indicated and the results can be hard to interpret when non-randomized therapies are undertaken termed intercurrent events (ICEs) as deemed appropriate by the clinical care team. The authors undertook this method by forming an independent adjudication committee (IAC) consisting of experts blinded to the study allocation arm that determined whether the ICE was indicated treatment. The findings were described and determined that most (2/3) of the ICEs were indicated and mostly consisted of additional steroid dosing. Furthermore, the category at baseline and the evolution of the category of disease (shock-like, KD like or undifferentiated) seemed to have an effect whether the ICE was indicated. While the problem the authors address are important and the authors indicate that masked adjudication if ICEs contributes to the interpretation of results, there are a few concerns that should be addressed. The primary issue is addressing the importance of the evaluation done and how will it improve the interpretation of treatments. The authors do not articulate how the data in this analysis adds to the initial report specifically and the reader/reviewer is left to his/her own thoughts about the conclusions drawn by the authors. In a study design such as this one with the incorporation of adjudication of ICEs, perhaps the clinical care team will intervene more uniformly and/or judiciously, if they know that non-randomized interventions are being closely scrutinized by an IAC for being indicated so that the study groups will be managed differently than if care beyond the study arm is not scrutinized? Perhaps if ICE rates are different between study groups, and are viewed as indicated more in one group than the other, it is an endpoint that indicates a difference in efficacy in the treatments different than the prospective outcomes of efficacy identified in the initial study design? The authors expressly say that the method described in this manuscript “contributes to the interpretation of results in open-label trials and should be consistently incorporated in the future”. I think that the authors may be right, but they should spend a little more energy explaining how their results support their conclusion. There are some other concerns that are not as concerning that would enhance the manuscript. 1. The manuscript was hard to read because of the non-standard abbreviations used and I did not find easily the abbreviations and what they meant. I was unfamiliar with the term KD-like and had to refer to their initial study to find that KD stood for kawasaki’s disease. 2. The term PIM-TS should have been described and one paragraph explaining the difference between the three different categories on presentation would have been helpful and made the rationale for the two treatment arms being studied more obvious. Furthermore, while not stated explicitly, I think that the three categories viewed are viewed as different severities (shock-like>KD-like>undifferentiated). However, I guess they could also be viewed as different phenotypes with different mechanisms for development. My guess these questions are not entirely answered, but the manuscript reads as the categories are based on severity of disease with similar mechanisms. Do not miss the opportunity to educate the reader. 3. Line 169 says consensus decisions were required. Does that mean majority or all members of the IAC needed to agree?

BMJ Open: first published as 10.1136/bmjopen-2023-078137 on 25 April 2024. Downloaded from http://bmjopen.bmj.com/ on August 3, 2024 by guest. Protected by copyright.

	4. The authors used median, interquartile ranges and Wilcoxon rank sum tests for continuous variables. I guess that means that the data were not appropriate for parametric testing. If that is so, the authors should state that, or if they did not consider parametric testing perhaps, they should have. 5. Results: a. Figure one, excellent. b. Figure 2? (page 20) would benefit from a legend. c. Table 1? (IVMP versus IVIG) would benefit from another column to the far right with the individual p values. There are many variables being assessed (ICE (none, indicated, not indicated), and category of disease and the treatment group allocation and the statistics are thus difficult. This table is probably important to explain better as the results may be important information supporting (or not) their conclusion). 6. Discussion: Overall the discussion was excellent a. Lines 257 and 258, “This information can be incorporated in to (sic) pre-specified analyses of the main trial results”. The authors should expand how the information can be useful in understanding better the main trial results. b. Lines 262 and 263, Those ICEs predominantly comprised of added oral glucocorticoids and were most likely based on local guidelines rather than clinical necessity.” That statement is generous and I wonder how the authors came to this speculation?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

My thanks to the authors for their important study and for bringing forward an equally important RCT. In my view, it is very important to address exactly the problem that was addressed here, namely the concomitant medication that was not randomised.I have only one point: I understand that the groups are already small, but I find it difficult that the data are not adjusted for gender, since it is known that the immune system already differs between the sexes in infants.

*Response: Thank you very much for bringing up this important aspect. Because of the small sample and group sizes, we prefer to present crude data. We note, however, a *slight* preponderance of male participants among the trial participants.*

Reviewer 2

The primary issue is addressing the importance of the evaluation done and how will it improve the interpretation of treatments. The authors do not articulate how the data in this analysis adds to the initial report specifically and the reader/reviewer is left to his/her own thoughts about the conclusions drawn by the authors. In a study design such as this one with the incorporation of adjudication of ICEs, perhaps the clinical care team will intervene more uniformly and/or judiciously, if they know that non-randomized interventions are being closely scrutinized by an IAC for being indicated so that the study groups will be managed differently than if care beyond the study arm is not scrutinized? Perhaps if ICE rates are different between study groups, and are viewed as indicated more in one group than the other, it is an endpoint that indicates a difference in efficacy in the treatments different than the prospective outcomes of efficacy identified in the initial study design? The authors expressly say that the method described in this manuscript “contributes to the interpretation of results in open-label trials and should be consistently incorporated in the future”. I think that the authors may be right, but they should spend a little more energy explaining how their results support their conclusion.

BMJ Open: first published as 10.1136/bmjopen-2023-078137 on 25 April 2024. Downloaded from http://bmjopen.bmj.com/ on August 3, 2024 by guest. Protected by copyright.

Response: We thank the reviewer for his high quality review. The clinical team was not aware that non-randomised interventions will undergo adjudication. Therefore, we consider this study to have been embedded completely into clinical practice. However, the review process was in a somewhat artificial setting. We emphasised your comment and added line 330-335 to the discussion.

1. The manuscript was hard to read because of the non-standard abbreviations used and I did not find easily the abbreviations and what they meant. I was unfamiliar with the term KD-like and had to refer to their initial study to find that KD stood for kawasaki's disease.

Response: We acknowledge that the manuscript does include a substantial amount of abbreviations. We reduced the abbreviations and furthermore added an abbreviation list (161-170) to the manuscript.

2. The term PIM-TS should have been described and one paragraph explaining the difference between the three different categories on presentation would have been helpful and made the rationale for the two treatment arms being studied more obvious. Furthermore, while not stated explicitly, I think that the three categories viewed are viewed as different severities (shock-like>KD-like>undifferentiated). However, I guess they could also be viewed as different phenotypes with different mechanisms for development. My guess these questions are not entirely answered, but the manuscript reads as the categories are based on severity of disease with similar mechanisms. Do not miss the opportunity to educate the reader.

Response: Thank you for highlighting that the reader might not be familiar with the disease classification of PIMS-TS as it is not something that is internationally used. We elaborated a further on PIMS-TS and added a paragraph in the introduction (187-196).

3. Line 169 says consensus decisions were required. Does that mean majority or all members of the IAC needed to agree?

Response: All members had to agree. Changed accordingly (239).

4. The authors used median, interquartile ranges and Wilcoxon rank sum tests for continuous variables. I guess that means that the data were not appropriate for parametric testing. If that is so, the authors should state that, or if they did not consider parametric testing perhaps, they should have.

Response: Due to the small sample size and the skewed data parametric testing seemed not appropriate. Added this statement (271-272)

5. Results:
a. Figure one, excellent.

Response:Thank you very much.

b. Figure 2? (page 20) would benefit from a legend.

Response: Thank you for your comment. Indeed, a legend would improve the understanding of the figure. We added the legends to the main manuscript and included the tables in to the main manuscript, line 511-529, 485-509.

c. Table 1? (IVMP versus IVIG) would benefit from another column to the far right with the individual p values. There are many variables being assessed (ICE (none, indicated, not indicated), and category of disease and the treatment group allocation and the statistics are thus difficult. This table is probably important to explain better as the results may be important information supporting (or not) their conclusion).

Response: We submitted a new table 1 with p-values added (488).

6. Discussion: Overall the discussion was excellent

BMJ Open: first published as 10.1136/bmjopen-2023-078137 on 25 April 2024. Downloaded from http://bmjopen.bmj.com/ on August 3, 2024 by guest. Protected by copyright.

Response: Your feedback is highly appreciated.

a. Lines 257 and 258, “This information can be incorporated in to (sic) pre-specified analyses of the main trial results”. The authors should expand how the information can be useful in understanding better the main trial results.

Response: Thank you very much again for bringing this up. Added a paragraph. (330-335)

b. Lines 262 and 263, Those ICEs predominantly comprised of added oral glucocorticoids and were most likely based on local guidelines rather than clinical necessity.” That statement is generous and I wonder how the authors came to this speculation?
Response: Yes, agree. Written like this, it is speculative. We have deleted this half sentence and added the fact that many guidelines do recommend tapering steroids (342, 334)

BMJ Open: first published as 10.1136/bmjopen-2023-078137 on 25 April 2024. Downloaded from http://bmjopen.bmj.com/ on August 3, 2024 by guest. Protected by copyright.