nature portfolio

# Peer Review File

# Racial/ethnic differences in post-acute sequelae of SARS-CoV-2 in children and adolescents in the United States

Corresponding Author: Professor Yong Chen

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
Thank you very much for the opportunity to review this manuscript. Overall, that's an interesting and well written article about Racial/Ethnic differences in Long-COVID-associated symptoms among a pediatrics population. Strength includes the multi-center approach, the large data set and the thorough statistical analysis.

Concerns/questions
Data sources: Only described very rudimentary.
→ What data were received from each institution?
→ Was the data set standardized, e.g., by using the same data abstraction form?
→ Were only EHR used? Are some information patient- or clinician-reported?
→ How were the data from different institutions harmonized?
Please describe this much more in detail. An overview of variables collected as an appendix would be helpful.

Classification of ethnic/racial groups: Description of the group construction and the underlying rational is described too superficially.
→ Some basic information for all ethnic/racial groups (raw data) should be presented
→ Based on this, process of categorization of ethnic/racial groups used in this article should be described.
→ In the flowchart should be described what ethnic/racial groups were excluded, including number of patients for each group
→ What is considered a small sample size? What was the cutoff for exclusion of ethnic/racial groups based on too small numbers?

Calculation of outcomes: Overall, it's unclear where the outcomes are coming from and how they were calculated.
→ Are the outcomes based on the EHR or another data collection tool?
→ How were they calculated? Using ICD-Codes?
→ Did all hospitals deliver information to outcomes?
→ For what outcome variables data were missing?

Results: The racial/ethnic differences very small – even with the high number of patients included.

Other comments
→ Discussion underdeveloped
→ Figure numbers are not in order

→ Reference “Khullar D, Zhang Y, Zang C, et al. Racial/ethnic disparities in post-acute sequelae of SARS-CoV-2 infection in New York: an EHR-based cohort study from the RECOVER program. J Gen Intern Med. 2023;38(5):1127-1136.” is listed twice (Reference 16 and 30)

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
This manuscript explores ethnic/racial differences in the association between SARS-CoV-2 infection and a list of 24 long-COVID symptoms among young individuals up to the age of 21 in a US setting. The focus of the manuscript is well introduced. However, some key details of the methods and results appear missing and should be included to improve the readability and understanding of what has been done. The discussion section should be developed further. Since the format of a manuscript for the Nature journals means the methods section is presented at the end of the manuscript, generally, a few pointers in the introduction and at the start of the results section will be helpful to the readers.
INTRODUCTION
- Page 4 - Lines 119-120 - The last sentence of the introduction is confusing. The authors refer to the “relationship between COVID-19, racial/ethnic differences and potential PASC symptoms and conditions” when it might be more appropriate to rephrase as “racial/ethnic differences in the relationship between COVID-19 and potential PASC symptoms and conditions”?
Is the word “potential” necessary? Is your end point “potential symptoms” or “symptoms”? If these symptoms/conditions are recorded in hospital, there are likely to be real symptoms/conditions? By potential, do you mean they might not be related to PASC? If not related to PASC then why would they be included in the list of PASC symptoms/conditions for this study?
RESULTS
- Page 4 – lines 125-129 – The population analysed is made of both those with COVID-19 (COVID-19 positive cohort) and those without (COVID-19 negative cohort). First, to improve readability, you can introduce both “COVID-19 positive cohort” and “COVID-19 negative cohort” in the first sentence of your results section as follows:
“The study involved 225,723 children and adolescents with COVID-19 (COVID-19 positive cohort) and 677,448 without COVID-19 (COVID-19 negative cohort).”
Second, a description of the characteristics of the “COVID-19 negative cohort” should be included in this paragraph following that or along the description of the characteristics of the “COVID-19 positive cohort”. Descriptive statistics of the “COVID-19 negative cohort” should be included in table 1 as well. You should have each cohort as column and maybe decide to keep the breakdown of either ethnicity or COVID-19 severity as column if you think this is useful and the other could go as in rows like age.
- Page 4 – line 128 – “March 2020 to October 2022” is introduced in this paragraph. This is later on referred as “entry month”. This could be simply mentioned here or further explained if preferred.
- Page 4 – lines 134-143 – the first sentence of this paragraph introduces a first key result of the manuscript i.e. incidence. Table 2 presents incidence for both cohorts and stratified by severity status.
First, the “severity status” is mentioned but has not yet been introduced to the reader. Perhaps a small description into parenthesis of what is this “severity status” would be useful.
Second, some statistical tests seem to have been performed between the 2 cohorts, but this is not clear between what and what. A p-value is provided in the text but not in table 2.
Third, the colour code might be wrongly used since the titles are presented in blue in table 2 when this should represent lower incidence in the COVID-19 positive cohort compared to the COVID-19 negative cohort.
Table 2 should provide more information on the tests and comparison performed, possibly in footnotes. If it is red for a condition, is this red colour code applicable to “all”, “severe” and “non-severe”?
Legend is unclear: “COVID-19 positive higher incidence” and might need to state what is it compared to.
Finally, to improve the structure of table 2, the way the conditions are listed could be reorganised to have “systemic” and all the conditions that are included within “systemic” and then “syndromic” and the listing of all the conditions that are included in “syndromic”.
- Page 4, line 146 – What does SMD stands for? Write in full. Readability will be improved if what has been done is clarified in a first sentence of this paragraph, even though this might have been detailed in methods/supplementary material.
- Figure 3. It is unusual to have the legend of the figure in the middle of the figure. Maybe “at least one groups ratio of relative risk <1” and “at least one groups ratio of relative risk > 1” should be moved at the bottom of the figure. The legend is misleading because the interpretation and colour system are based on whether the confidence interval (CI) includes 1. So, it is not just about the value of the risk ratio (RR) here. Maybe the legend needs to be rephrased to be more accurate and mention the role of the CI in judging on whether the RR is below or above 1.

- page 5 – line 158 and line 169 – the text says “Hispanic patients showed no increase in any of the conditions” but later says there is an increase in “hair loss” in that same group. There is a problem of logic and terminology. This “any of the conditions” terminology also referred to “At least one condition” is an outcome that is not clear and might need to be replaced by another wording to ensure clear interpretation of the findings.
- Page 6, lines 224-225 - the supplementary file is large but some of the extra analysis performed are vaguely mentioned with no interpretation of the key findings. The supplementary file might need to be reviewed to include what is really necessary.
- Section S3 and figure S13 are not referred to in the text of the manuscript. The title of figure S13 might be misleading. Is it about using regression analysis only (to compare to DiD), or using both DiD and regression? What is the purpose of Figure S13?

DISCUSSION
The discussion is short and a few limitations/strengths could be added. Other points that could be added include:
- Explain how the findings might be generalisable to the rest of the population.
- Is the sample size robust enough to generate strong evidence?
- For example, it is interesting that there is a lack of racial/ethnic differences found in figure 3 for most racial/ethnic groups apart from AAPI compared to the NHW group for the combined outcomes. In such cases, we tend to think of an issue of small sample size leading to not being able to distinguish ethnic differences. However, AAPI is the smallest racial/ethnic group and for this group, the results show evidence of differences compared to NHWhite. Therefore, this strengthens the finding for the other groups. This point could be reflected on in the discussion section.
- Discuss bias due to the way the cohorts have been selected.
- Asymptomatic cases are less likely to be captured because less likely to be tested and therefore this will have an impact in a specific direction to the findings for the non-severe COVID-19. Discuss how this affect your results.
- How could you capture more information on socio-economic differences to strengthen your analysis. As the discussion explains, socio-economic differences are a key element in understanding ethnic/racial differences in health. If you have individuals’ addresses, could these be linked to geographical level data that can inform on an area-based measure of deprivation for example?

- Page 6, line 229 – In the first summary sentence of the discussion, it would be useful to specify this is about children and younger individuals, so it flows better with the reference to evidence in adults appearing in a later sentence.
- Page 6, lines 233-234, there is a bit of redundancy between these two sentences. The statement that findings in younger individuals align with previous findings in adults should be mentioned once only.
- Page 7, lines 255-259 – You can state the type of bias as “ascertainment bias” in this paragraph.

METHODS
- Page 7, lines 278-285. Even though a link is provided, a few sentences should be added to describe what the data is about. Right now, this is unclear what the data contains and what are the exact sources. Is this about hospitalisation data only (as the list of sites suggests)? Is it linked to other information from other sources (death, community testing, GP data, vaccination data, …)? What type of data sources are used? Electronic Health Records or self-reported survey information or a mix of both?
- Page 7, lines 292-293 – From which data source is the COVID-19 information collected/captured?
- Page 8, lines 296-298 – For the “COVID-19 negative patients”, from which data source are they selected from? Have they all had a visit in hospital? Are they meant to be representative of the general population or a specific subpopulation of the general population? Could you clarify? Reference to figure 1 is not enough to explain the selection of the cohorts. This should be explained in a couple of sentences in the text as well.
- Page 8, reference to figure 1 appears after reference to figures 2-3, figures will need to be renumbered.
- Page 8, line 320, entry month is “March 2020 to October 2022”. A clear range of information is collected at baseline (entry month). However, this is not clear whether patients are followed over time to collect information on PASC symptoms/conditions and over which period. More information on follow-up, censoring, median follow-up time for example should be described.
- Page 8, line 337. Incidence here is likely a measure of the number of individuals with new symptoms/conditions developed after baseline over all individuals, all followed over a specific period of time. Therefore, it might be more appropriate to use person-year as a denominator, especially is the follow-up period varies across individuals.
- Page 9, line 356 – the end of the statistical methodology does not develop on how pre-infection racial/ethnic differences are accounted for and using which methodology. This should be described at the end of this paragraph, possibly after mention of Poisson regression.
- Page 9, line 379, when mention of “31 negative control” is given, a pointer to where to find these conditions in supplementary material should also been given more precisely (i.e. table s1).

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)
The study utilized data from RECOVER and examined differences in PASC across racial and ethnic groups in a pediatric population. The research question is very important, and the authors have sufficient data resource to address it. However, the paper requires significant improvements in clarity and is missing several key pieces of information. In its current form, the importance of the study is diminished due to unclear writing and a lack of focus on clinical questions.

1. The study provided a comparison of outcomes across racial and ethnic groups, as well as the unadjusted rates in COVID-19 and control groups within the same racial and ethnic groups. However, it lacks the adjusted rates of outcomes (which represent PASC) by comparison between the COVID-19 and control groups. The incidence of PASC in this population itself holds major clinical importance. Moreover, it is impossible to estimate the difference in PASC without determining whether the outcome qualifies as PASC in this population.

2. The baseline characteristics of the non-COVID group, as well as all groups after matching, need to be presented. Additionally, please include the number of participants matched for each matching criterion.

3. The study calculated the incidence rate by dividing the number of patients experiencing an incident outcome by the total number of patients (lines 336-337). Such a comparison of incidence rates presents a clear bias when comparing groups with different follow-up times and when comparing groups with different baseline prevalences of outcomes. Unless I have overlooked it, I did not see any effort mentioned in the manuscript to balance the follow-up or to adjust for the baseline prevalence of outcomes.

4. The 'DID' analyses presented by the authors are very unclear.
According to the manuscript, the study only conducted pairwise matching and did not match between the COVID and non-COVID groups.
A valid comparison of PASC between two racial groups would require balanced baseline characteristics both between the COVID and non-COVID groups, and between racial groups within each COVID-19 status. This would require multiple arms matching, which does not appear to be conducted.
Given the differences in baseline characteristics between the COVID and non-COVID groups, the beta2 and beta3 (which, beta3 are the main results of the study) in the author's 'DID' formula could be biased by confounding effects.
Please further justify the methodology and provide a clear explanation of how confounding effects between COVID and non-COVID were adjusted for. Additionally, justify the decision to conduct matching separately instead of employing other statistical methods such as weighting or multiple treatment matching, which could provide comparisons within the same target population.

Figure 2 shows the advantage of the study as examining the difference in outcomes between two groups during follow-up after removing the differences that existed at baseline. Should this be considered as basic concept of causal inference? How is it more advanced compared to studies that balance the baseline prevalence of outcomes through the propensity score method or g-estimation? Or compared to studies that estimate purely incident outcomes in populations without the outcome at baseline?
More clarity is needed to understand how this study represents an advancement in methodology. Moreover, given the lack of clarity, it is also unclear how the application of negative controls used in the study is superior to existing methods around empirical calibration.

5. Several terms used in the paper appear inappropriate:
a. The term 'potential PASC' has been used in multiple places. It is unclear what this implies. Does it reflect the authors' lack of confidence in determining PASC, or does it suggest that the outcome measured may not be associated with COVID-19? I suggest removing the word 'potential.'
b. In lines 148 and 150, what do 'moderate' or 'strong evidence' mean? Are these based on effect size, P values, or another measure? Please clarify the quantification of 'moderate' and 'strong.'
c. The authors suggest they aim to account for time-varying unmeasured confounders. However, in this study, there are no time-varying confounders given the fixed exposures (racial/ethnic group and COVID status defined in this study). Please review the definition of a time-varying confounder.
d. In line 159, why is an RR of 0.93 (0.85, 1.24) described as a 'minor decrease'?
e. In line 164, what does 'almost no decrease' mean given an RR of 0.99 (0.89, 1.11)?
f. In lines 172-173, why is an RR of 0.74 described as 'increased risk'?
g. Line 134: There should be no PASC for participants without COVID-19."

6. The discussion fails to contextualize the study results in terms of their clinical interpretation and significance. The three strengths highlighted by the authors in the discussion (using matching, balancing pre-infection differences, and employing negative control outcomes) do not appear substantial enough to be emphasized as the central elements of the discussion of a methodological study nor for a clinical research study.

7. The study requires participants, both infected and uninfected, to have at least one healthcare visit within 28-179 days after the index date. I am concerned that this may result in selection bias, as the majority of the study population may not require such frequent healthcare visit, especially those who were included during 2020 and 2021 and uninfected, due to hospital policies.
Similarly, the study selected individuals who tested negative as the control group. Given that tests were not commonly conducted within the study population, does this suggest that the study is selecting a control group that is likely to have symptoms, or have an appointment for surgery or a procedure?

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
Thank you very much for the hard work you put into the revision of this manuscript. All of my comments have been addressed, and I have no further comments.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
This is a review of the first revised version of the manuscript titled "Racial/Ethnic Differences in Long-COVID-Associated Symptoms among Pediatrics Population: Findings from Difference-in-differences Analyses in RECOVER Program".

This revised version has hugely improved compared to the first submitted version of the manuscript and the efforts of the authors to respond to reviewer's comments are commendable.

The authors have addressed my comments thoroughly and appropriately. A few minor points:
- Line 157: This seems to be the first mention of “NCO”. Make sure all acronyms are spelt out at first mention.
- Line 306: replace “the other minor groups” by “the other minority groups”. Double-check if other mentions of “minor group” appear in the text and replace by “minority group” which is a better accepted terminology.
- Line 431: suggest using “who had at least one visit from 7 days before up to 18 months after the index date” rather than the other way around.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)
I thank the authors for their efforts in revising the manuscript and for providing a detailed response. The manuscript has shown improvement through the clarification and correction of the previously highlighted errors. However, I still have a few remaining concerns.

1. The authors have now clarified that only participants with positive COVID-19 were used in their analyses, without comparison to those without COVID-19. This raises concerns about whether the effects measured are truly reflective of PASC. It is important to note that the references offered by the authors in their response used both those with and without COVID-19 to quantify PASC.
The authors could claim they are measuring the difference in PASC or long COVID between races without a COVID-19 negative control group only if: 1. They can demonstrate that these conditions did not occur within the COVID-19 negative group, or 2. They can show there is no difference in these outcomes across races within the COVID-19 negative group. Otherwise, they are only measuring differences in diagnosis or healthcare utilization after COVID-19, which does not necessarily indicate PASC. While this remains an important research question, there is a key distinction compared to measuring PASC.

2. I thank the authors for their clarification regarding their consideration and utilization of the DID approach. In addition to this, I suggest the authors clearly specify the two "differences" involved in the DID for this study early in the manuscript. Furthermore, while the authors state that DID can remove confounding effects from unmeasured confounders, could the authors please comment on the following scenario: if COVID-19 modifies the effect of an unmeasured confounder on the outcome (similar to how it modifies the effect of race on the outcome), would DID introduce bias in the presence of such unmeasured confounders?

3. I also suggest the authors proofread and revise the manuscript throughout. For example, the latter portion of the revised sentence, "Hispanic patients showed almost no increase in at least one condition (RR 1.01, 95% CI 0.98 to 1.04, P=0.498), and the relative risk of 0.99 (95% CI: 0.89, 1.11) indicates no statistically significant change in risk," is unclear and needs revision.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #5

(Remarks to the Author)
I thank the authors for carefully considered my comments and their efforts on revising the manuscript. The results from analyses on COVID-19 negative cohort is convincing and I have no further comments.

(Remarks on code availability)

## Response to Reviewer 1:

**Thank you for giving us the opportunity to review this manuscript entitled: "Racial/Ethnic Differences in Long-COVID-Associated Symptoms among Pediatrics Population: Findings from Difference-in-differences Analyses in RECOVER Program".**

**The paper addresses the racial/ethnic differences in symptoms and conditions of post-acute sequelae SARS-CoV-2 infection (PASC) among children and adolescents stratified for COVID-19 severity and takes into account previous disparities with a difference-in-differences approach.**

**In general, it seems that the Methods section has been placed after the Discussion, according to Natures format, after the completion of the manuscript, as there are several mix ups in numbering of figures and tables, as well as lack of definitions of -abbreviations. We will further address our comments by section.**

**Our response**: Thank you for reviewing our manuscript. We appreciate your feedback. We have reorganized the manuscript to ensure the Methods section is correctly placed, renumbered all figures and tables, and included definitions for all abbreviations. We have also addressed specific comments in the Introduction, Methods, Results, Discussion, and Conclusion sections to enhance clarity and quality. Thank you again for your valuable input.

**Introduction:**

1. **Line 75: Could you use a sentence explaining the difference between syndromic and systematic PASC symptoms?**

   **Our response**: Thank you for the clarification question. Syndromic features are symptoms, signs, and non-specific laboratory abnormalities; Systemic features are diagnosed health conditions. We added the following words in the main manuscript and more details are provided in **Table S3** in the Supplementary Appendix, which reads:

   *"To investigate racial and ethnic differences attributable to COVID-19, we examined twenty-four PASC symptoms and conditions. We grouped the PASC symptoms and conditions into two categories: systemic and syndromic features (shown in **Table S3**). Syndromic features encompass symptoms, signs, and non-specific laboratory abnormalities, while systemic features are diagnosed health conditions."*

   **Table S3**. Classification of Symptoms and Conditions into Systematic and Syndromic Groups.

| Syndromic (symptoms, signs, non-specific laboratory abnormalities) | Systemic (diagnosed health conditions) |
|---|---|

| | |
|---|---|
| • abdominal pain<br>• abnormal liver enzyme<br>• cardiovascular signs and symptoms<br>• changes in taste and smell<br>• chest pain<br>• cognitive functions<br>• fatigue and malaise<br>• fever and chills<br>• generalized pain<br>• hair loss<br>• headache<br>• mental health disorders<br>• musculoskeletal pain<br>• respiratory signs and symptoms<br>• skin symptoms | • acute kidney injury<br>• acute respiratory distress syndrome<br>• arrhythmias<br>• fluid and electrolyte<br>• heart disease<br>• myocarditis<br>• myositis<br>• Postural Orthostatic Tachycardia Syndrome (POTS) or dysautonomia<br>• thrombophlebitis and thromboembolism |

2. **Line 115-116: What pre-infection disparities are you accounting for? Disparities in PASC symptoms?**

   **Our response**: Yes, the pre-infection disparities we are accounting for are disparities in PASC symptoms and conditions among different racial/ethnic groups. We added the following words right after the 115-116, which reads:

   *"Importantly, we quantified the racial/ethnic differences linked with COVID-19 by carefully accounting for pre-infection disparities in PASC symptoms and conditions through the application of difference-in-differences analyses."*

**Results:**

3. **Table 1:**
   - **Severe/non-severe COVID-19: How is that defined?**

     **Our response**: The severity of COVID-19 at cohort entry was stratified into the following categories: asymptomatic, mild (symptomatic), moderate (involving moderately severe COVID-19-related conditions like gastroenteritis, dehydration, and pneumonia), and severe (comprising unstable COVID-19-related conditions, ICU admission, or mechanical ventilation) [1]. In this paper, we categorized patients exhibiting either asymptomatic or mild symptoms as belonging to the "non-severe" group, while all other patients were classified as part of the "severe" group. We included a footnote to this table in the updated manuscript, which reads:

     *"Abbreviation: NHW, Non-Hispanic White; NHB, Non-Hispanic Black; AAPI, Asian American or Pacific Islander; PMCA, Pediatric Medical Complexity Algorithm [1].*

*Footnote:*
*1: According to the definition of the severity of COVID-19 [2] at cohort entry, the positive patients with moderate or severe COVID-19 were grouped into the severe group.*
*2: Similarly, the positive patients with asymptomatic or mild COVID-19 were grouped into the non-severe group."*

*Reference:*
*[1] Simon TD, Haaland W, Hawley K, Lambka K, Mangione-Smith R. Development and validation of the Pediatric Medical Complexity Algorithm (PMCA) version 3.0. Acad Pediatr. 2018;18(5):577-580.*
*[2] Forrest CB, Burrows EK, Mejias A, et al. Severity of acute COVID-19 in children< 18 years old March 2020 to December 2021. Pediatrics. 2022;149(4):e2021055765.*

- **PMCA: Abbreviation meaning? (It is explained later in the Methods section, but a definition earlier is needed).**

**Our response**: Thank you for pointing this out. PMCA means Pediatric Medical Complexity Algorithm, which categories no chronic condition, non-complex chronic condition, and complex chronic condition. More details are in **Table S2** in the Supplementary Appendix, and a footnote to Table 1 has been included in the updated manuscript:

*"PMCA denotes the Pediatric Medical Complexity Algorithm[1]."*

And the details about PMCA in **Table S2**:

**Table S2**. Confounding Variables used in the propensity model and other variables for eligible criteria.

| Variable | Functional form | Values | Detail | Codes/references |
|---|---|---|---|---|
| **Confounding variables** | | | | |
| **PMCA (Pediatric Medical Complexity Algorithm)** | 3 categories | No chronic condition (PMCA = 0)<br>Non-complex chronic condition (PMCA = 1)<br>Complex chronic condition comorbidities (PMCA = 2) | Based on the condition occurrence and visit occurrence domains. | Simon, T. D., Haaland, W., Hawley, K., Lambka, K., & Mangione-Smith, R. (2018). Development and validation of the Pediatric Medical Complexity Algorithm (PMCA) version 3.0. *Academic pediatrics*, *18*(5), 577-580. |

*Reference:*
*[1] Simon TD, Haaland W, Hawley K, Lambka K, Mangione-Smith R. Development and validation of the Pediatric Medical Complexity Algorithm (PMCA) version 3.0. Acad Pediatr. 2018;18(5):577-580.*

4. **Table 2:**
   - **It is not clear how or why you are dividing COVID-19-negative patients into being 'severe' and 'non-severe'? Is it the severity of COVID or PASC? Can you clarify. (According to the subheadings in the table)**

   **Our response**: Thank you for the clarification question. The severity classification pertains to COVID-19-related symptoms and conditions. Specifically, the severity status is based on the criteria outlined by Forrest et al. (2022) [2], which include COVID-19-related symptoms and conditions such as gastroenteritis, dehydration, pneumonia, ICU admission, and mechanical ventilation. For COVID-19-negative patients, we determined severity status based on the *symptoms and conditions observed during the acute phase (seven days before to thirteen days after the index date)*. The index date refers to a random COVID-19 negative test. We have clarified this in the Supplementary Materials, which now reads:

   *"For COVID-19-negative patients, we included patients who had neither a documented SARS-CoV-2 infection nor a diagnosis of multisystem inflammatory syndrome in children (MIS-C) within the same study period, and who had at least one negative COVID-19 test result. A random negative test was chosen as the index date for COVID-19-negative patients. The selection of participants for COVID-19-negative in real-world data is summarized in* ***Figure S13****.*

   *For COVID-19-negative patients, we determined severity status based on the symptoms and conditions observed during the acute phase (seven days before to thirteen days after the index date)."*

   *Reference:*
   *[2] Forrest CB, Burrows EK, Mejias A, et al. Severity of acute COVID-19 in children< 18 years old March 2020 to December 2021. Pediatrics. 2022;149(4):e2021055765.*

5. **Line 146: SMD: Abbreviation meaning? Should be written out here and not first in Methods.**
   **Our response**: Thank you for pointing this out. We have added "standardized mean difference (SMD)" in the Results section in the updated manuscript, which reads

   *"After achieving the balance of standardized mean difference (SMD) by propensity score matching (****Section S11****) and conducting the difference-in-difference model with NCO calibration,* ***Figure 3*** *shows the racial/ethnic difference attributable to COVID-19 in PASC symptoms and conditions by the severity of COVID-19 after the NCO calibration."*

6. **Line 146: Figure 3 – should be renamed Figure 1, if it is the first figure?**

**Our response**: Thank you for your suggestion. Based on feedback from you and other reviewers, we have added two key figures prior to the results section—one outlining the selection criteria and the other explaining the rationale for the main analysis. As a result, the figure presenting the racial/ethnic differences remains as **Figure 3**.

7. **Line 158: "in any condition" does that correspond to what is called "At least one condition" in Figure 3"? It is not clear what this refers to.**

   **Our response**: We apologize for the confusion. Yes, "in any condition" corresponds to "at least one condition" as mentioned in Figure 3. We have updated the manuscript to reflect this change. That is,

   *"Hispanic patients showed no increase in at least one condition (RR 0.99, 95% CI 0.91 to 1.08, P=0.804)."*

8. **Line 175: CI includes 1.00 and you have a significant p-value? Can that be right? Is the result significant?**

   **Our response**: The lower bound of CI was written as 1.00 because we kept only two decimal places. The CI with three decimal places is 1.004. We added the following words to avoid confusion,

   *"[…] respiratory signs and symptoms (RR 1.11, P=0.036) with 95% CI 1.00 to 1.23 (significant; rounded to two decimals) compared to NHW patients."*

9. **Line 186-193: Are you referring to the correct figures? Figure S2, do you mean Section S2? The text does not correspond to Figure S3 but rather to Figure S13 in Section S3.**

   **Our response**: Thank you for the question. We reordered the Figure numbers and Sections. Figure S2 should be Section S2 and Figure S3 should be Figure S14 in Section S3. We corrected these in the updated manuscript, which reads

   *"Section S2 showed the results of the negative control outcome experiments and estimated systematic error, such as the unmeasured confounder bias. Figure S14 in Section S3 showed the racial/ethnic differences after SARS-CoV-2 infection stratified by severity of COVID-19, using standard regression models only."*

10. **Line 195-207: You are in the text referring to figure S13 in supplementary as I read it. You have not written the reference anywhere?**

    **Our response**: Thank you, we reordered the Figure S13 to be Figure S14. We have added this paragraph referring to Figure S14 in the updated manuscript, which referred to as

    *"In Figure S14, among COVID-19 patients with severe illness during the acute infection, […]"*

**Discussion:**

11. **It would be interesting, on top of existing literature, to read your reflections on why we see these differences in PASC? And what implications does this have on a more general public health level in order to care for this patient group?**

**Our response**: Please find our point-to-point response below:

- "Why do we see these differences in PASC?"

  The observed differences in PASC can be attributed to differential access to health professionals, particularly pediatric subspecialists, which vary significantly across different regions of the country according to a recent National Academy of Medicine report.[3] These variations are not due to biological reasons but are more closely linked to socio-economic factors, healthcare infrastructure, and regional healthcare policies.

  *Reference:*
  *[3] Rivara FP, Gonzalez-del-Rey J, Forrest CB. The pediatric subspecialty physician workforce. JAMA pediatrics. 2024 Feb 1;178(2):107-8.*

- "What implications does this have on a more general public health level in order to care for this patient group?"

  This study is among one of the first studies examining PASC differences, providing valuable insights into the clinical relevance of these disparities. Understanding these differences is crucial for developing targeted interventions to ensure equitable healthcare access and outcomes for all pediatric populations affected by PASC.

  The implications of this study on a broader public health level are significant. By utilizing data from 13 institutions, representing approximately 10% of US children and encompassing both urban and suburban health systems, this study offers a comprehensive national sample. Public health strategies must prioritize the reduction of healthcare access disparities and the enhancement of specialized pediatric care availability across all regions. This will help in better managing and treating PASC in children, ensuring that all affected patients receive the necessary care regardless of their geographic location or socio-economic status. Furthermore, these findings underscore the importance of continuous surveillance, research, and tailored public health policies to address the evolving needs of this patient group.

  In response, we added the following sentences to the Discussion part,

  *"While the observed racial/ethnic differences in risk ratios are subtle, they remain significant when applied to large populations and may reflect broader systemic issues. For example, the observed differences in PASC can primarily be attributed to disparities in healthcare access and the availability of specialized pediatric care, which vary significantly across different regions of the country according to a national report[4]. These variations are not due to*

*biological reasons but are more closely linked to socio-economic factors, healthcare infrastructure, and regional healthcare policies.*

*This study is among one of the first studies examining PASC racial/ethnic differences, providing valuable insights into the clinical relevance of these disparities. Understanding these differences is crucial for developing targeted interventions to ensure equitable healthcare access and outcomes for all pediatric populations affected by PASC.*

*The increased risk among minor racial/ethnic pediatric patients may necessitate targeted follow-up care and support for the minor racial/ethnic population. Healthcare providers should be aware of the potential for varied PASC presentations across racial/ethnic groups in pediatric populations. This knowledge can inform more targeted screening and follow-up protocols, potentially improving early detection and management of PASC symptoms in different racial/ethnic patient populations.*

*The implications of this study on a broader public health level are significant. By utilizing data from 13 institutions, representing approximately 10% of US children and encompassing both urban and suburban health systems, this study offers a comprehensive national sample. Public health strategies must prioritize the reduction of healthcare access disparities and the enhancement of specialized pediatric care availability across all regions. This will help in better managing and treating PASC in children, ensuring that all affected patients receive the necessary care regardless of their geographic location or socio-economic status. Furthermore, these findings underscore the importance of continuous surveillance, research, and tailored public health policies to address the evolving needs of this patient group."*

*Reference:*
*[4] Rivara FP, Gonzalez-del-Rey J, Forrest CB. The pediatric subspecialty physician workforce. JAMA pediatrics. 2024 Feb 1;178(2):107-8.*

**Methods:**

12. **Line 285: What is OCHIN?**

    **Our response**: OCHIN stands for Oregon Community Health Information Network, which serves as a national network of community health organizations. We have updated the manuscript, which referred to

    *"Participating institutions in this study included: […], Oregon Community Health Information Net (OCHIN), […]"*

13. **Line 293-294: "[…] we included the patients who had positive polymerase-chain-reaction (PCR), serology, or antigen tests or diagnoses of COVID-19, or post-acute sequelae of SARS-CoV-2 (PASC), which we defined as documented SARS-CoV-2 infection": how do you define PASC without a prior COVID-19 diagnosis? Can you clarify that sentence?**

**Our response**: Thank you for the clarification. The diagnosis of post-acute sequelae of SARS-CoV-2 (PASC) is identified using the ICD code U09.9, which is designated for conditions occurring after COVID-19. If a patient has this diagnosis code, it indicates that the patient likely had COVID-19 infection, although the positive test or initial diagnosis may not have been recorded in the healthcare system. We have revised the sentence accordingly in the updated manuscript:

*"For COVID-19-positive patients, we included the patients who had positive polymerase-chain-reaction (PCR), serology, or antigen tests or diagnoses of COVID-19, or diagnoses of PASC (U09.9), which we defined as documented SARS-CoV-2 infection."*

14. **Line 300: Figure 1 should be renamed Figure 2**

**Our response**: Thank you for your suggestion. We have moved the paragraph about the selection criteria into the Results section to improve readability. As a result, the figure depicting the cohort selection criteria is now labeled as Figure 1 in the main manuscript.

This reads in the main manuscript,
*"The study involved 225,723 children and adolescents with COVID-19 (COVID-19-positive cohort) from March 2020 to October 2022 (cohort entry month). […] The COVID-19-positive cohort comprised 109,022 NHW patients (48.3%), 45,823 NHB patients (20.3%), 60,012 Hispanic patients (26.6%), and 10,866 AAPI patients (4.8%) with SARS-CoV-2 infection from March 2020 to October 2022 (**Figure 1**)."*

15. **312-313: You should refer to Supplementary, Table S3 in Section S1, E.**

**Our response**: Thank you for the suggestion. We have added a referral sentence to Table S3 in the Supplementary Appendix in the updated manuscript:

*"Systematic and syndromic conditions related to PASC were grouped by the 24 potential PASC symptoms and conditions, which are detailed in Table S3."*

16. **Line 322: '25' right after the parenthesis, is that at reference?**

**Our response**: Sorry for the confusion. We have updated the format for this reference in the updated manuscript, which reads

"*The primary exposure was race/ethnicity, categorized into NHW, NHB, Hispanic, and AAPI. […] a chronic condition indicator as defined by the Pediatric Medical Complexity Algorithm (PMCA, no chronic condition, non-complex chronic condition, and complex chronic condition)*[28]"

17. **325-329: The severity of COVID-19, is that based on health care records, hospital contacts, and survey data, or how did you measure COVID severity?**

**Our response:** Thank you for the question. The severity of COVID-19 was measured using healthcare records, and the definition for each level of severity can be referred to in reference 2 in the response letter.

The severity of COVID-19 at cohort entry was stratified into the following categories: asymptomatic, mild (symptomatic), moderate (involving moderately severe COVID-19-related conditions like gastroenteritis, dehydration, and pneumonia), and severe (comprising unstable COVID-19-related conditions, ICU admission, or mechanical ventilation) [2]. In this paper, we categorized patients exhibiting either asymptomatic or mild symptoms as belonging to the "non-severe" group, while all other patients were classified as part of the "severe" group. We included a footnote to Table 1 in the updated manuscript, which reads as follows:

*"Footnote:*
*1: According to the definition of the severity of COVID-19 [2] at cohort entry, the positive patients with moderate or severe COVID-19 were grouped into the severe group.*
*2: Similarly, the positive patients with asymptomatic or mild COVID-19 were grouped into the non-severe group."*

*Reference:*
*[2] Forrest CB, Burrows EK, Mejias A, et al. Severity of acute COVID-19 in children< 18 years old March 2020 to December 2021. Pediatrics. 2022;149(4):e2021055765.*

**18. Line 329: '26' right after the parenthesis, is that at reference?**

**Our response**: Sorry for the confusion. We have updated the format for this reference in the updated manuscript.

*"The severity of COVID-19 at the cohort entry date was stratified into the following categories: asymptomatic, mild (symptomatic), moderate (involving moderately severe COVID-19-related conditions like gastroenteritis, dehydration, and pneumonia), and severe (comprising unstable COVID-19-related conditions, ICU admission, or mechanical ventilation)[27]"*

**19. Line 354: Figure 2 should be renamed Figure 3**

**Our response**: Thank you. To accommodate the nature format, we added one paragraph briefly talking about the difference-in-difference model in the Results section. Thus, this figure remains as **Figure 2** in the updated manuscript.

*"**Figure 2** illustrates the rationale and methodology of our difference-in-differences (DiD) with NCO calibration approach."*

# Response to Reviewer 2:

**Thank you very much for the opportunity to review this manuscript. Overall, that's an interesting and well-written article about Racial/Ethnic differences in Long-COVID-associated symptoms among a pediatric population. Strength includes the multi-center approach, the large data set, and the thorough statistical analysis.**

**Our response**: Thank you for reviewing our manuscript. We appreciate your feedback. We have addressed specific comments in the Introduction, Methods, Results, Discussion, and Conclusion sections to enhance clarity and quality. Thank you again for your valuable input.

**Concerns/questions**

1. **Data sources: Only described very rudimentary.**
   - **What data were received from each institution?**
   - **Was the data set standardized, e.g., by using the same data abstraction form?**
   - **Was only EHR used? Is there patient- or clinician-reported information?**
   - **How were the data from different institutions harmonized?**

   **Please describe this much more in detail. An overview of variables collected as an appendix would be helpful.**

   **Our response**: Please find our point-to-point response below:

   - "What data were received from each institution?"

     Data received from each institution is electronic health records (EHR) data, which includes clinical data such as diagnoses and treatments, laboratory and test results, and administrative data including patient demographics and billing information.

   - "Was the data set standardized, e.g., by using the same data abstraction form?"

     Yes, the data used in this study were structured and standardized EHR entries made by healthcare providers within hospital settings.

   - "Were only EHR used? Is there patient- or clinician-reported information?"

     Yes, only EHR data were used. No self-reported data were used in this study.

   - "How were the data from different institutions harmonized?"

     The National Institutions of Health (NIH) launched the new RECOVER initiative in 2021 to leverage EHR data to better identify and characterize patients with PASC. RECOVER obtains EHRs from three large national healthcare networks within the United States. The RECOVER employs PEDSnet common data model (CDM) to standardize data formats and structures across participating institutions and harmonized data are stored in a centralized repository.

More details have been included in **Section S1 A.** in the Supplementary Appendix as blow:

*"The real-world data utilized in our analysis is derived from electronic health records (EHRs), covering a wide range of healthcare interaction information routinely collected and stored by hospitals. This includes clinical data such as diagnoses and treatments, laboratory and test results, and administrative data including patient demographics and billing information. The hospital-based EHR data from the Researching COVID to Enhance Recovery (RECOVER) Initiative COVID-19 Database served as the basis for defining and determining exposure, outcomes, and covariates. Unlike General Practitioner (GP) data, self-reported data, or external data sources, our study used the structured, standardized EHR entries made by healthcare providers within hospital settings. EHR data provides a more detailed and integrated view of a patient's health status, medical history, and healthcare interactions across various providers and settings."*

2. **Classification of ethnic/racial groups: Description of the group construction and the underlying rationale is described too superficially.**
   - **Some basic information for all ethnic/racial groups (raw data) should be presented.**
   - **Based on this, the process of categorization of ethnic/racial groups used in this article should be described.**
   - **The flowchart should describe what ethnic/racial groups were excluded, including the number of patients for each group.**
   - **What is considered a small sample size? What was the cutoff for exclusion of ethnic/racial groups based on too small numbers?**

**Our response**: Thank you for the clarification. Please find our point-to-point response below:

- "Some basic information for all ethnic/racial groups (raw data) should be presented."

The raw data for all ethnic/racial groups are from the "person" CDM table from the PEDSnet CDM in the EHR data. The ethnic and racial information is identified from columns "ethnicity_concept_id" and "race_concept_id". Details of categorical definitions are as follows:

**-American Indian or Alaska Native**: A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.
**-Asian**: A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
**-Black or African American**: A person having origins in any of the black racial groups of Africa.
**-Native Hawaiian or Other Pacific Islander**: A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
**-White**: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

For patients with multiple races (i.e. biracial), race is considered a single concept, meaning there is only one race slot. If there are multiple races in the source system, concatenate all races into one race_source_value (see below) and use concept_id code as 'Multiple Race.'

Predefined values for race are:

- American Indian/Alaska Native: concept_id = 8657
- Asian: concept_id = 8515
- Black or African American: concept_id = 8516
- Native Hawaiian or Other Pacific Islander: concept_id = 8557
- White: concept_id = 8527
- Multiple Race: concept_id = 44814659 (vocabulary_id='PCORNet')
- Refuse to answer: concept_id = 44814660 (vocabulary_id='PCORNet')
- No Information: concept_id = 44814650 vocabulary_id='PCORNet')
- Unknown: concept_id = 44814653
- Other: concept_id = 44814649

For PEDSnet, a person with Hispanic ethnicity is defined as "A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race." Please include valid concept ids (consistent with OMOP CDMv5).
Predefined value set for ethnicity are:

- Hispanic: concept_id = 38003563
- Not Hispanic: concept_id = 38003564
- No Information: concept_id = 44814650 (vocabulary_id='PCORNet')
- Unknown: concept_id = 44814653 (vocabulary_id='PCORNet')
- Other: concept_id = 44814649 (vocabulary_id='PCORNet')

We have added the racial/ethnic information to **Table S2** and **Section S3** in the updated supplementary material. The Section S3 in the Supplementary Material now includes following

"*The raw data for all ethnic/racial groups are from the "person" CDM table from the PEDSnet CDM in the EHR data. The ethnic and racial information is identified from columns "ethnicity_concept_id" and "race_concept_id". Details of categorical definitions are as follows:*
***American Indian or Alaska Native**: A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.*
***Asian**: A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.*
***Black or African American**: A person having origins in any of the black racial groups of Africa. Native Hawaiian or Other Pacific Islander: A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.*

***White**: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.*
*For patients with **multiple races** (i.e. biracial), race is considered a single concept, meaning there is only one race slot. If there are multiple races in the source system, concatenate all races into one race_source_value (see below) and use concept_id code as 'Multiple Race.' […] "*

The **Table S2** about the race/ethnicity presents below:

**Table S2**. Confounding Variables used in the propensity model and other variables for eligible criteria.

| Variable | Functional form | Values | Detail | Codes/references |
|---|---|---|---|---|
| **Confounding variables** | | | | |
| **Race/Ethnicity** | 4 categories | NHW<br>NHB<br>Hispanic<br>AAPI | Based on the race_concept_id and ethnicity_concept id in the person domain. | See https://github.com/PEDSnet/Data_Models_Public/blob/master/PEDSnet/docs/Conventions%20Docs/v5.5_PEDSnet_CDM_ETL_Conventions.md |

- "Based on this, the process of categorization of ethnic/racial groups used in this article should be described."

  All patients with "ethnicity_concept_id" 38003563 are categorized as "Hispanic". Then based on "race_concept_id", a value of 8516 indicated "Non-Hispanic Black (NHB)", 8515 and 8557 indicated "Asian American/Pacific Islanders (AAPI)", 8527 indicated "Non-Hispanic White (NHB)", 44814659 indicating "Multiple", and other values indicated "Other/Unknown".

- "The flowchart should describe what ethnic/racial groups were excluded, including the number of patients for each group."

  We have added the number of patients of Multiple and Other/Unknown groups in the updated **Figure 1**.

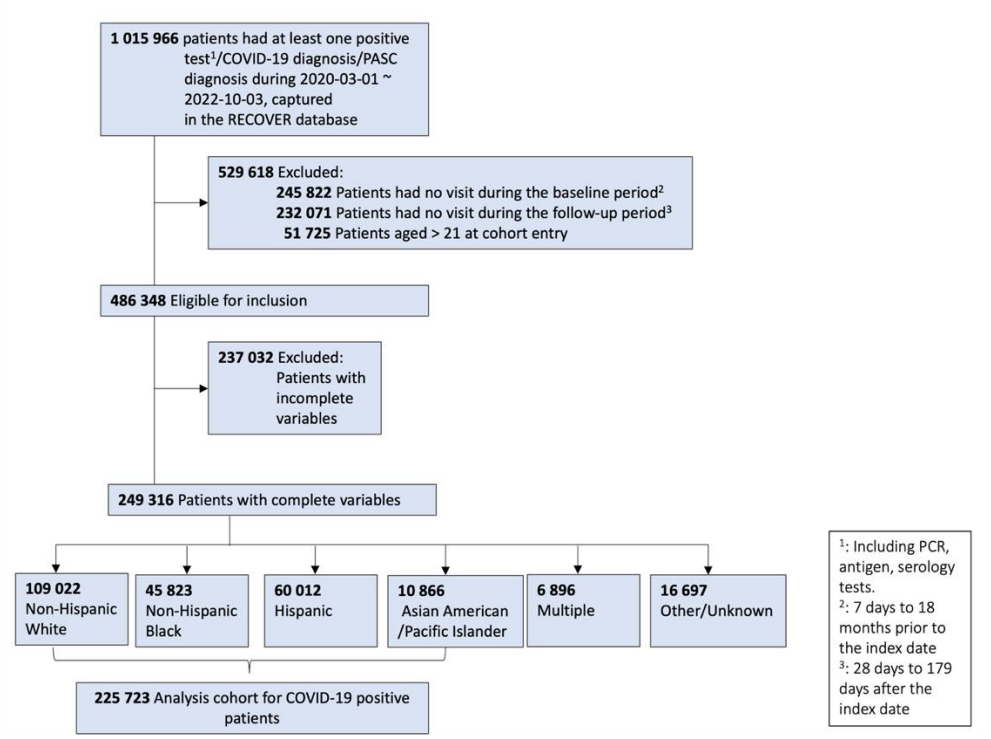



***Figure 1****. Selection of participants for COVID-19-positive patients. The "patients with complete variables" refer to patients with complete records of BMI, race/ethnicity, and vaccine records.*

- "What is considered a small sample size? What was the cutoff for exclusion of ethnic/racial groups based on too small numbers?"

  We excluded the race/ethnicity based on both the consideration of statistical power (sample size) and the interpretability of the study.

  In our study, we considered a sample size to be small for an ethnic/racial group if it comprised less than 3% of the population in any of the following categories:

  1. Severe COVID-19 positive patients
  2. Non-severe COVID-19 positive patients
  3. Severe COVID-19 negative patients
  4. Non-severe COVID-19 negative patients

  This 3% threshold was applied consistently across all categories to ensure adequate representation and statistical power for meaningful comparisons.

  Regarding the exclusion of the "Other/Unknown" categories from our analysis, this decision was made because:

1. The "Other" category often includes a heterogeneous mix of ethnicities that can't be reliably interpreted as a cohesive group.
2. The "Unknown" category represents missing data, which cannot be meaningfully analyzed in the context of racial/ethnic comparisons.

These exclusions help to maintain the integrity and interpretability of our racial/ethnic analyses, focusing on well-defined groups with sufficient representation in our dataset.

We have added this information to the methods section of our manuscript to provide readers with a clear understanding of our approach to handling racial/ethnic data in this study, which reads as follows

*"We used patient race/ethnicity data included in the EHR and categorized patients into six racial/ethnic groups: non-Hispanic White (NHW), non-Hispanic Black (NHB), Hispanic, Asian American/Pacific Islanders (AAPI), Multiple, Other or Unknown. Multiple comprising less than 3% of the population in any COVID-19 and severity category (**Table S4**) were considered small samples and excluded from analysis. The "Other or Unknown" categories were also excluded due to interpretability issues."*

We have added more description in the updated manuscript, and more details have been included in **Section S3** in the Supplementary Appendix and **Table S4**:

***Table S4**: Number of patients for COVID-19 positive and negative groups, stratified by severity status.*

| | ***COVID-19-positive*** | | | ***COVID-19-negative*** | | |
|---|---|---|---|---|---|---|
| | **Severe (N=18,489)** | **Non-severe (N=230,827)** | **Overall (N=249,316)** | **Severe (N=158,783)** | **Non-severe (N=594,263)** | **Overall (N=753,046)** |
| **Hispanic** | 3786 (20.5%) | 56226 (24.4%) | 60012 (24.1%) | 26637 (16.8%) | 123049 (20.7%) | 149686 (19.9%) |
| **Non-Hispanic White** | 9140 (49.4%) | 99882 (43.3%) | 109022 (43.7%) | 87222 (54.9%) | 277891 (46.8%) | 365113 (48.5%) |
| **Non-Hispanic Black/African American** | 3346 (18.1%) | 42477 (18.4%) | 45823 (18.4%) | 24024 (15.1%) | 106441 (17.9%) | 130465 (17.3%) |
| **Asian/Pacific Islander** | 590 (3.2%) | 10276 (4.5%) | 10866 (4.4%) | 5709 (3.6%) | 26475 (4.5%) | 32184 (4.3%) |
| **Multiple** | 584 (3.2%) | 6312 (2.7%) | 6896 (2.8%) | 5392 (3.4%) | 17960 (3.0%) | 23352 (3.1%) |
| **Other/Unknown** | 1043 (5.6%) | 15654 (6.8%) | 16697 (6.7%) | 9799 (6.2%) | 42447 (7.1%) | 52246 (6.9%) |

3. **Calculation of outcomes: Overall, it's unclear where the outcomes are coming from and how they were calculated.**
   - **Are the outcomes based on the EHR or another data collection tool?**
   - **How were they calculated? Using ICD-Codes?**
   - **Did all hospitals deliver information to outcomes?**

- **For what outcome variables data were missing?**

**Our response**: Thank you for the questions. We answered your questions for each point:

- "Are the outcomes based on the EHR or another data collection tool?"

The outcomes are based on the EHR diagnoses CDM table.

- "How were they calculated? Using ICD-Codes?"

They were captured using validated diagnostic codes (ICD and SNOMED) codes, with details of the code sets available in the updated **Table S2** in the Supplementary Appendix.

- Did all hospitals deliver information to outcomes?

Yes, all participating institutions delivered information on outcomes.

- For what outcome variables data were missing?

We addressed intermittent or sporadic missing data, particularly for covariates such as obesity or race, where values were absent in the EHR. In these cases, the EHR system would indicate a record with a corresponding concept_id, showing that a person's weight or race/ethnicity was missing. For outcome variables, the EHR recorded only diagnosis codes without indicating any missingness. We defined the outcome as present (value of 1) if the relevant ICD codes for the diagnosis were found during the follow-up or baseline period, and absent (value of 0) if these ICD codes were not present. As the outcome was determined solely based on the presence or absence of ICD codes, there is no missing data for the outcome definitions.

We updated the Supplement to show the definitions of the outcomes in **Table S2** as follows:

| Variable | Functional form | Values | Detail | Codes/references |
|---|---|---|---|---|
| **PASC symptoms and conditions** | 24 categories | abdominal pain, abnormal liver enzyme, acute kidney injury, acute respiratory distress syndrome, arrhythmias, cardiovascular signs and symptoms, changes in taste and smell, chest pain, cognitive functions, fatigue and malaise, fever and chills, fluid and electrolyte, generalized pain, hair loss, headache, | Based on the condition occurrence and visit occurrence domains. A value of 1 to the outcome if the relevant ICD codes for the diagnosis were found during the follow-up or baseline period, and a value of 0 if these ICD codes were not present during the corresponding period. There is no intermittent or sporadic missing data concerning the outcome definitions. | Razzaghi H, Forrest CB, Hirabayashi K, Wu Q, Allen AJ, Rao S, Chen Y, Bunnell HT, Chrischilles EA, Cowell LG, Cummins MR. Vaccine effectiveness against long COVID in children. Pediatrics. 2024 Apr 1;153(4):e2023064446. |

| | |
|---|---|
| | heart disease, mental health disorders, musculoskeletal pain, myocarditis, myositis, Postural Orthostatic Tachycardia Syndrome (POTS) or dysautonomia, respiratory signs and symptoms, skin symptoms, and thrombophlebitis and thromboembolism |

4. **Results: The racial/ethnic differences are very small – even with the high number of patients included.**

**Our response**: Thank you for your observation regarding the racial/ethnic differences in our results. We agree that the differences appear small, and we appreciate the opportunity to elaborate on their significance and implications.

The racial/ethnic differences we observed are indeed subtle, but it's important to note that they are presented in the scale of risk ratios. Even small differences in risk ratios can translate to substantial impacts when applied to large populations. In addition, these small differences are particularly noteworthy in the context of healthcare disparities, where seemingly minor variations can reflect broader systemic issues. We believe these differences may be attributed to several factors, including:

1. Differential access to specialists
2. Disparities in healthcare access and quality
3. Socioeconomic factors that influence health outcomes

We have added a discussion of these points in our manuscript to contextualize our findings and emphasize their relevance despite their small magnitude.

*"While the observed racial/ethnic differences in risk ratios are subtle, they remain significant when applied to large populations and may reflect broader systemic issues. For example, the observed differences in PASC can primarily be attributed to disparities in healthcare access and the availability of specialized pediatric care, which vary significantly across different regions of the country according to a national report[30]. These variations are not due to biological reasons but are more closely linked to socio-economic factors, healthcare infrastructure, and regional healthcare policies."*

**Other comments**

**5. Discussion underdeveloped.**

**Our response**: Thank you for the suggestion. We have enriched the Discussion section with the following aspects:

- The reason we see these differences in PASC. We updated the manuscript as follows
  *"While the observed racial/ethnic differences in risk ratios are subtle, they remain significant when applied to large populations and may reflect broader systemic issues. For example, the observed differences in PASC can primarily be attributed to disparities in healthcare access and the availability of specialized pediatric care, which vary significantly across different regions of the country according to a national report[29]. These variations are not due to biological reasons but are more closely linked to socio-economic factors, healthcare infrastructure, and regional healthcare policies."*

- The implications on a more general public health level. We updated the manuscript as follows, "*This study is among one of the first studies examining PASC racial/ethnic differences, providing valuable insights into the clinical relevance of these disparities. Understanding these differences is crucial for developing targeted interventions to ensure equitable healthcare access and outcomes for all pediatric populations affected by PASC.*

  *The increased risk among minor racial/ethnic pediatric patients may necessitate targeted follow-up care and support for the minor racial/ethnic population. Healthcare providers should be aware of the potential for varied PASC presentations across racial/ethnic groups in pediatric populations. This knowledge can inform more targeted screening and follow-up protocols, potentially improving early detection and management of PASC symptoms in different racial/ethnic patient populations.*

  *The implications of this study on a broader public health level are significant. By utilizing data from 13 institutions, representing approximately 10% of US children and encompassing both urban and suburban health systems, this study offers a comprehensive national sample. Public health strategies must prioritize the reduction of healthcare access disparities and the enhancement of specialized pediatric care availability across all regions. This will help in better managing and treating PASC in children, ensuring that all affected patients receive the necessary care regardless of their geographic location or socio-economic status. Furthermore, these findings underscore the importance of continuous surveillance, research, and tailored public health policies to address the evolving needs of this patient group."*

- Strength of the statistical methods used in the main analysis: We highlighted three key strengths of our statistical approach: the use of matching, difference-in-differences (DiD) analysis, and calibration with negative control outcomes, which detailed as follows:
  *"The method utilized in our study has multiple strengths. First, we used propensity score matching methods instead of linear regression models in our adjustment of the confounders, which helped us reduce the non-linear effects of the confounders[30]. Second, we accounted for the pre-infection racial/ethnic difference in long-term COVID-19-related symptoms and conditions. This approach enabled us to quantify racial/ethnic differences attributable to*

*COVID-19 more accurately. The DiD model is particularly powerful in the context of studying racial/ethnic disparities, as it helps to isolate the effect of COVID-19 from pre-existing health disparities. By comparing the change in health outcomes before and after COVID-19 across racial/ethnic groups, we can more confidently attribute observed differences to the impact of the virus rather than to the pre-infection racial/ethnic differences.*

*A key strength of our study is the implementation of NCOs, which enhances the reliability and interpretability of our findings. We used the NCOs which are outcomes not expected to be affected by COVID-19 or show racial/ethnic differences due to COVID-19. By analyzing these alongside our primary outcomes, we enable the detection of systematic bias from unmeasured confounding variables. This approach allows us to calibrate results from the DiD model. Importantly, NCOs help mitigate the impact of unmeasured confounding that may change before and after the infection period. This is particularly crucial when studying racial/ethnic disparities, where many dynamic social and environmental factors may not be captured in our dataset."*

- The possible bias sources regarding the cohort: (1) **Socioeconomic Differences as a Mediator:** Socioeconomic disparities, particularly those related to race/ethnicity, may act as mediators influencing racial/ethnic differences in PASC outcomes. (2) **Health-Seeking Behavior and Healthcare Access (Ascertainment Bias):** Limited access to care and medical records for some minority racial groups could introduce ascertainment bias. However, the use of negative control outcomes (NCOs) in the study helps to account for these disparities, mitigating this potential bias. (3) **Confounding and Misclassification Bias:** Confounding is a major concern in observational studies. To counter this, the study employed propensity score matching, DiD analysis, and NCOs to address residual biases. Additionally, potential data completeness issues in EHRs could lead to misclassification. The misclassification bias may warrant future research. (4) **Impact of Asymptomatic Cases:** Asymptomatic COVID-19 cases are often not captured, which may impact findings related to non-severe cases. This study focused on diagnosed or treated cases, which may lead to conservative estimates in the DiD analysis due to misclassification bias. Future research should capture a broader range of COVID-19 severity. (5) **Selection Bias and Healthcare Visits:** The requirement for at least one healthcare visit within 28-179 days post-infection helped reduce bias from loss to follow-up. We updated the discussion section as follows

  *"Our study has several research directions that warrant future investigations. First, socioeconomic differences due to race/ethnicity may exacerbate racial/ethnic differences in PASC symptoms and conditions, thereby acting as a mediator effect in the causal pathway between race/ethnicity and clinical outcomes. Such influences have been suggested as risk factors for acute COVID-19 by Chisolm and colleagues in a RECOVER EHR study[31]. Future research on PASC outcomes is of interest to study such mediation effects.*

  *Secondly, health-seeking behavior or healthcare access is an important consideration, known as ascertainment bias[32]. It is possible that certain minority racial groups have more limited access to care and associated medical records and that this contributes to potential bias in the observed racial/ethnic differences. Related issues were recently described by Nasir et al. for*

*the ascertainment of PASC symptoms and conditions in adult populations through EHR[33]. Nevertheless, we tried to address the ascertainment bias by including healthcare utilization factors in propensity score model, such as the number of inpatient, outpatient, and ED visits during the baseline period (shown in **Table S2**).*

*Thirdly, confounding poses a significant bias threat in observational studies. To address this, we extensively adjusted for potential confounders using a propensity-score-based matching method and difference-in-differences analyses. We employed negative control outcomes to reduce the residual bias, such as unmeasured confounder bias. Additionally, EHR data completeness issues may lead to misclassification and loss-to-follow-up bias. Some attempts have been made to mitigate the impacts of these biases[34–37]. The analysis used a combined set of patients, but potential race/ethnicity bias may vary between outpatients and inpatients. Addressing these issues can help improve the reliability of evidence generated from these investigations.*

*Fourthly, asymptomatic cases are less likely to be captured as they are less likely to be tested, impacting the findings for non-severe COVID-19. This study focused on diagnosed or treated cases of Long COVID, reflecting children who sought healthcare services. The direction of bias can be bidirectional, depending on whether asymptomatic cases are disproportionately missed in minority or NHW groups. This could attenuate the DiD analysis for the non-severe COVID-19 group due to potential misclassification bias, leading to conservative or inflated results. Future research should aim to capture a more comprehensive spectrum of COVID-19 severity.*

*Fifthly, our study design, requiring at least one healthcare visit within 28-179 days post-index date, balanced selection bias against 'loss to follow-up' bias[38]. While this approach may over-represent frequent healthcare users, especially during 2020-2021, we mitigated this by incorporating diverse testing data to reduce COVID-19 status misclassification. Furthermore, in the sensitivity analysis (**Section S4**), the resulting selection bias likely included sicker patients in the COVID-19 negative group, potentially biasing our results towards the null. Nevertheless, we still observed a higher incidence of PASC symptoms and conditions for COVID-19-positive patients compared to COVID-19-negative patients. Future studies could explore alternative designs or data sources to enhance the external validity of these findings."*

Reference:

[29] Rivara FP, Gonzalez-del-Rey J, Forrest CB. The pediatric subspecialty physician workforce. *JAMA Pediatr*. 2024;178(2):107-108.

[30] Amoah J, Stuart EA, Cosgrove SE, et al. Comparing Propensity Score Methods Versus Traditional Regression Analysis for the Evaluation of Observational Data: A Case Study Evaluating the Treatment of Gram-Negative Bloodstream Infections. *Clin Infect Dis*. 2020;71(9):e497. doi:10.1093/CID/CIAA169

[31] Chisolm DJ, Webb R, Salamon K, et al. Associations between area-level health-related social factor indices and risk of acute COVID-19: An EHR-based cohort study from the RECOVER program. *medRxiv*. Published online 2022:2012-2022.

[32] Asch DA, Sheils NE, Islam MN, et al. Variation in US Hospital Mortality Rates for Patients Admitted With COVID-19 During the First 6 Months of the Pandemic. *JAMA Intern Med*. 2021;181(4):471-478. doi:10.1001/JAMAINTERNMED.2020.8193
[33] Nasir M, Cook N, Parras D, et al. Using Data Science and a Health Equity Lens to Identify Long-COVID Sequelae Among Medically Underserved Populations. *J Health Care Poor Underserved*. 2023;34(2):521-534. doi:10.1353/HPU.2023.0047
[34] Duan R, Cao M, Wu Y, et al. An Empirical Study for Impacts of Measurement Errors on EHR based Association Studies. *AMIA Annual Symposium Proceedings*. 2016;2016:1764.
[35] Chen Y, Wang J, Chubak J, Hubbard RA. Inflation of type I error rates due to differential misclassification in EHR-derived outcomes: empirical illustration using breast cancer recurrence. *Pharmacoepidemiol Drug Saf*. 2019;28(2):264-268.
[36] Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proceedings of the National Academy of Sciences*. 2018;115(11):2571-2577.
[37] Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. Interpreting observational studies: why empirical calibration is needed to correct p-values. *Stat Med*. 2014;33(2):209-218.
[38] Howe CJ, Cole SR, Lau B, Napravnik S, Eron Jr JJ. Selection bias due to loss to follow up in cohort studies. *Epidemiology*. 2016;27(1):91-97.

6. **Figure numbers are not in order.**

   **Our response**: We have corrected the Figure numbers in the updated manuscript.

7. **Reference "Khullar D, Zhang Y, Zang C, et al. Racial/ethnic disparities in post-acute sequelae of SARS-CoV-2 infection in New York: an EHR-based cohort study from the RECOVER program. J Gen Intern Med. 2023;38(5):1127-1136." is listed twice (Reference 16 and 30)**

   **Our response**: We have corrected this reference in the updated manuscript.

## Response to Reviewer 3:

**I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.**

**Our response**: Thank you very much. We greatly appreciate your input to this manuscript.

## Response to Reviewer 4:

**This manuscript explores ethnic/racial differences in the association between SARS-CoV-2 infection and a list of 24 long-COVID symptoms among young individuals up to the age of 21 in a US setting. The focus of the manuscript is well introduced. However, some key details of the methods and results appear missing and should be included to improve the readability and understanding of what has been done. The discussion section should be developed further. Since the format of a manuscript for the Nature journals means the methods section is presented at the end of the manuscript, generally, a few pointers in the introduction and at the start of the results section will be helpful to the readers.**

**Our response**: Thank you very much for your valuable feedback. We have added pointers in the introduction and at the start of the results section to improve the readability. We have enriched the method and discussion sections to help the readers better understand what has been done in this study. We also renumbered all figures and tables and included definitions for all abbreviations. We have also addressed specific comments in the Introduction, Methods, Results, Discussion, and Conclusion sections to enhance clarity and quality.

INTRODUCTION

1. **- Page 4 - Lines 119-120 - The last sentence of the introduction is confusing. The authors refer to the "relationship between COVID-19, racial/ethnic differences and potential PASC symptoms and conditions" when it might be more appropriate to rephrase as "racial/ethnic differences in the relationship between COVID-19 and potential PASC symptoms and conditions"?**

   **Response**: Thank you for the suggestion. The sentence suggested made it more clear, and we have incorporated this sentence in the revised manuscript:

   *"[…] racial/ethnic differences in the relationship between COVID-19 and PASC symptoms and conditions."*

2. **Is the word "potential" necessary? Is your endpoint "potential symptoms" or "symptoms"? If these symptoms/conditions are recorded in the hospital, are there likely to be real symptoms/conditions? By potential, do you mean they might not be related to PASC? If not related to PASC then why would they be included in the list of PASC symptoms/conditions for this study?**

   **Our response**: Thank you for pointing this out. These symptoms and conditions are recorded in the hospital, and are diagnosed or treated Long COVID cases. We have removed the word "potential" to avoid further confusion in the revised manuscript. The updated manuscript reads as follows:

   *"We use the term 'PASC symptoms and conditions' to include a broad spectrum of health issues observed following COVID-19 infection*[25,26]*"*

*[25] Rao S, Lee GM, Razzaghi H, et al. Clinical features and burden of postacute sequelae of SARS-CoV-2 infection in children and adolescents. JAMA Pediatr. 2022;176(10):1000-1009.*
*[26]Razzaghi H, Forrest CB, Hirabayashi K, et al. Vaccine effectiveness against long COVID in children. Pediatrics. 2024;153(4):e2023064446.*

RESULTS

3. **- Page 4 – lines 125-129 – The population analyzed is made of both those with COVID-19 (COVID-19 positive cohort) and those without (COVID-19 negative cohort). First, to improve readability, you can introduce both "COVID-19 positive cohort" and "COVID-19 negative cohort" in the first sentence of your results section as follows:**
**"The study involved 225,723 children and adolescents with COVID-19 (COVID-19 positive cohort) and 677,448 without COVID-19 (COVID-19 negative cohort)."**

**Our response**: Thank you for the suggestion. We acknowledge the readability issues related to introducing the "COVID-19 negative cohort" in the crude analysis comparing the incidence between the COVID-19 positive and negative cohorts. Based on feedback from other reviewers, we have moved the crude analysis to the supplementary materials (**Section S4**) to avoid distracting from the main results focused on racial/ethnic differences in PASC symptoms and conditions within the "COVID-19 positive cohort". As our primary analysis used a difference-in-differences approach focused on COVID-19 positive patients, comparing outcomes before and after COVID-19 infection, we revised the first sentence of the Results section to reflect this focus. The updated manuscript now reads:

*"The study involved 225,723 children and adolescents with COVID-19 (COVID-19 positive cohort) from March 2020 to October 2022 (cohort entry month)."*

*"To confirm the defined twenty-four PASC symptoms and conditions, we conducted a crude analysis for the incidence of the PASC symptoms and conditions compared to the COVID-19-negative cohort (**Section S4**). A total of 677,448 COVID-19-negative patients were included in the crude analysis. A random negative test was chosen as the index date for COVID-19-negative patients and the selection criteria for the COVID-19-negative cohort are shown in **Figure S13**. The results indicated that the incidence of all PASC symptoms and conditions was higher in the COVID-19-positive patients compared to the COVID-19-negative patients (**Table S6**)."*

4. **Second, a description of the characteristics of the "COVID-19 negative cohort" should be included in this paragraph following that or along with the description of the characteristics of the "COVID-19 positive cohort". Descriptive statistics of the "COVID-19 negative cohort" should be included in Table 1 as well. You should have each cohort as a column and maybe decide to keep the breakdown of either ethnicity or COVID-19 severity as a column if you think this is useful and the other could go as in rows like age.**

**Our response**: Thank you for the suggestion. We agreed to describe the characteristics of the "COVID-19 negative cohort". Descriptive statistics of the COVID-19-negative cohort have been included in the updated supplementary materials **Table S5** as we moved this incidence table into the supplement table per other reviewers' comments to not distract the main results from the racial/ethnic difference in PASC symptoms and conditions.

The updated **Table S5** now reads as follows:

**Table S5**. Baseline characteristics of COVID-19 negative patients, by race/ethnicity and severity status.

| | ***Severe*** | | | | ***Non-severe*** | | | |
|---|---|---|---|---|---|---|---|---|
| | **NHW (N=87,222)** | **AAPI (N=5,709)** | **Hispanic (N=26,637)** | **NHB (N=24,024)** | **NHW (N=277,891)** | **AAPI (N=26,475)** | **Hispanic (N=123,049)** | **NHB (N=106,441)** |
| **Age categories (%)** | | | | | | | | |
| <1 years | 10895 (12.5%) | 711 (12.5%) | 3461 (13.0%) | 3978 (16.6%) | 24652 (8.9%) | 2313 (8.7%) | 11795 (9.6%) | 10748 (10.1%) |
| 1 to <5 years | 24363 (27.9%) | 1692 (29.6%) | 8024 (30.1%) | 7646 (31.8%) | 75138 (27.0%) | 7739 (29.2%) | 32990 (26.8%) | 31213 (29.3%) |
| 5 to <12 years | 23450 (26.9%) | 1727 (30.3%) | 7382 (27.7%) | 5815 (24.2%) | 82081 (29.5%) | 8493 (32.1%) | 41893 (34.0%) | 33954 (31.9%) |
| 12 to <16 years | 14085 (16.1%) | 771 (13.5%) | 4087 (15.3%) | 3175 (13.2%) | 45195 (16.3%) | 3128 (11.8%) | 19255 (15.6%) | 15311 (14.4%) |
| 16 to <21 years | 14429 (16.5%) | 808 (14.2%) | 3683 (13.8%) | 3410 (14.2%) | 50825 (18.3%) | 4802 (18.1%) | 17116 (13.9%) | 15215 (14.3%) |
| **Sex** | | | | | | | | |
| Female | 38451 (44.1%) | 2374 (41.6%) | 11735 (44.1%) | 10568 (44.0%) | 138595 (49.9%) | 13015 (49.2%) | 60993 (49.6%) | 52848 (49.7%) |
| Male | 48771 (55.9%) | 3335 (58.4%) | 14902 (55.9%) | 13456 (56.0%) | 139296 (50.1%) | 13460 (50.8%) | 62056 (50.4%) | 53593 (50.4%) |
| **Hospital** | | | | | | | | |
| A | 9036 (10.4%) | 344 (6.0%) | 872 (3.3%) | 2286 (9.5%) | 44879 (16.1%) | 1639 (6.2%) | 4010 (3.3%) | 15472 (14.5%) |
| B | 13650 (15.6%) | 1180 (20.7%) | 2933 (11.0%) | 5301 (22.1%) | 52929 (19.0%) | 4833 (18.3%) | 10083 (8.2%) | 23891 (22.4%) |
| C | 10948 (12.6%) | 524 (9.2%) | 6719 (25.2%) | 1210 (5.0%) | 16505 (5.9%) | 838 (3.2%) | 10727 (8.7%) | 2402 (2.3%) |
| D | 3475 (4.0%) | 271 (4.7%) | 927 (3.5%) | 1678 (7.0%) | 13149 (4.7%) | 1343 (5.1%) | 5287 (4.3%) | 8610 (8.1%) |
| E | 6184 (7.1%) | 154 (2.7%) | 532 (2.0%) | 492 (2.0%) | 11805 (4.2%) | 604 (2.3%) | 1543 (1.3%) | 1500 (1.4%) |
| F | 5702 (6.5%) | 692 (12.1%) | 4498 (16.9%) | 1704 (7.1%) | 7353 (2.6%) | 911 (3.4%) | 5971 (4.9%) | 2114 (2.0%) |
| G | 7301 (8.4%) | 434 (7.6%) | 623 (2.3%) | 1150 (4.8%) | 25525 (9.2%) | 2997 (11.3%) | 2245 (1.8%) | 3658 (3.4%) |
| H | 676 (0.8%) | 10 (0.2%) | 21 (0.1%) | 111 (0.5%) | 9893 (3.6%) | 292 (1.1%) | 554 (0.5%) | 1458 (1.4%) |
| I | 14095 (16.2%) | 860 (15.1%) | 1368 (5.1%) | 4838 (20.1%) | 34132 (12.3%) | 3274 (12.4%) | 5465 (4.4%) | 19021 (17.9%) |
| J | 9492 (10.9%) | 436 (7.6%) | 4810 (18.1%) | 3527 (14.7%) | 26348 (9.5%) | 1667 (6.3%) | 15282 (12.4%) | 10672 (10.0%) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| K | 430 (0.5%) | 42 (0.7%) | 701 (2.6%) | 189 (0.8%) | 17024 (6.1%) | 4133 (15.6%) | 50784 (41.3%) | 11677 (11.0%) |
| L | 1616 (1.9%) | 567 (9.9%) | 1764 (6.6%) | 412 (1.7%) | 7631 (2.7%) | 3183 (12.0%) | 8160 (6.6%) | 2507 (2.4%) |
| M | 4617 (5.3%) | 195 (3.4%) | 869 (3.3%) | 1126 (4.7%) | 10718 (3.9%) | 761 (2.9%) | 2938 (2.4%) | 3459 (3.2%) |
| **Entry time** | | | | | | | | |
| 03/2020 - 06/2020 | 6598 (7.6%) | 358 (6.3%) | 1611 (6.0%) | 1404 (5.8%) | 10273 (3.7%) | 714 (2.7%) | 3540 (2.9%) | 3039 (2.9%) |
| 07/2020 - 10/2020 | 13240 (15.2%) | 771 (13.5%) | 3371 (12.7%) | 2909 (12.1%) | 37038 (13.3%) | 2343 (8.8%) | 11387 (9.3%) | 9104 (8.6%) |
| 11/2020 - 02/2021 | 12002 (13.8%) | 698 (12.2%) | 3125 (11.7%) | 2802 (11.7%) | 46000 (16.6%) | 2981 (11.3%) | 14354 (11.7%) | 12656 (11.9%) |
| 03/2021 - 06/2021 | 13330 (15.3%) | 815 (14.3%) | 3909 (14.7%) | 3443 (14.3%) | 38195 (13.7%) | 3134 (11.8%) | 14268 (11.6%) | 14430 (13.6%) |
| 07/2021 - 10/2021 | 12747 (14.6%) | 902 (15.8%) | 3996 (15.0%) | 3962 (16.5%) | 52742 (19.0%) | 5802 (21.9%) | 25692 (20.9%) | 22646 (21.3%) |
| 11/2021 - 02/2022 | 11798 (13.5%) | 781 (13.7%) | 3822 (14.3%) | 3452 (14.4%) | 46495 (16.7%) | 5083 (19.2%) | 22561 (18.3%) | 19908 (18.7%) |
| 03/2022 - 06/2022 | 10277 (11.8%) | 779 (13.6%) | 3939 (14.8%) | 3408 (14.2%) | 28142 (10.1%) | 3927 (14.8%) | 17856 (14.5%) | 14132 (13.3%) |
| 07/2022 – 10/2022 | 7230 (8.3%) | 605 (10.6%) | 2864 (10.8%) | 2644 (11.0%) | 19006 (6.8%) | 2491 (9.4%) | 13391 (10.9%) | 10526 (9.9%) |
| **Obesity** | | | | | | | | |
| 0 | 59812 (68.6%) | 4299 (75.3%) | 16878 (63.4%) | 14958 (62.3%) | 190501 (68.6%) | 17846 (67.4%) | 54986 (44.7%) | 59931 (56.3%) |
| 1 | 27410 (31.4%) | 1410 (24.7%) | 9759 (36.6%) | 9066 (37.7%) | 87390 (31.4%) | 8629 (32.6%) | 68063 (55.3%) | 46510 (43.7%) |
| **PMCA** | | | | | | | | |
| 0 | 48327 (55.4%) | 3241 (56.8%) | 15290 (57.4%) | 13804 (57.5%) | 186253 (67.0%) | 19569 (73.9%) | 91266 (74.2%) | 71779 (67.4%) |
| 1 | 16614 (19.0%) | 1029 (18.0%) | 4902 (18.4%) | 4602 (19.2%) | 51283 (18.5%) | 3952 (14.9%) | 18525 (15.1%) | 20753 (19.5%) |
| 2 | 22281 (25.5%) | 1439 (25.2%) | 6445 (24.2%) | 5618 (23.4%) | 40355 (14.5%) | 2954 (11.2%) | 13258 (10.8%) | 13909 (13.1%) |
| **Negative tests prior entry** | | | | | | | | |
| 0 | 77433 (88.8%) | 5035 (88.2%) | 23626 (88.7%) | 20892 (87.0%) | 242041 (87.1%) | 23083 (87.2%) | 108613 (88.3%) | 92542 (86.9%) |
| 1 | 6553 (7.5%) | 455 (8.0%) | 2083 (7.8%) | 2141 (8.9%) | 25358 (9.1%) | 2423 (9.2%) | 10649 (8.7%) | 10170 (9.6%) |
| 2 | 1816 (2.1%) | 117 (2.0%) | 522 (2.0%) | 543 (2.3%) | 6537 (2.4%) | 596 (2.3%) | 2537 (2.1%) | 2460 (2.3%) |
| >=3 | 1420 (1.6%) | 102 (1.8%) | 406 (1.5%) | 448 (1.9%) | 3955 (1.4%) | 373 (1.4%) | 1250 (1.0%) | 1269 (1.2%) |
| **Vaccine dosage** | | | | | | | | |
| 0 | 79029 (90.6%) | 4866 (85.2%) | 23943 (89.9%) | 22208 (92.4%) | 243699 (87.7%) | 21074 (79.6%) | 105725 (85.9%) | 96004 (90.2%) |
| 1 | 1387 (1.6%) | 149 (2.6%) | 512 (1.9%) | 401 (1.7%) | 5670 (2.0%) | 923 (3.5%) | 3101 (2.5%) | 2242 (2.1%) |
| >=2 | 6806 (7.8%) | 694 (12.2%) | 2182 (8.2%) | 1415 (5.9%) | 28522 (10.3%) | 4478 (16.9%) | 14223 (11.6%) | 8195 (7.7%) |

5. **- Page 4 – line 128 – "March 2020 to October 2022" is introduced in this paragraph. This is later on referred to as "entry month". This could be simply mentioned here or further explained if preferred.**

   **Our response**: Thank you for your suggestion. We added the text right after the month, which reads

   *"The study involved 225,723 children and adolescents with COVID-19 (COVID-19 positive cohort) from March 2020 to October 2022 (cohort entry month)."*

6. **- Page 4 – lines 134-143 – the first sentence of this paragraph introduces a first key result of the manuscript i.e. incidence. Table 2 presents incidence for both cohorts and stratified by severity status.**
   - **First, the "severity status" is mentioned but has not yet been introduced to the reader. Perhaps a small description into parenthesis of what is this "severity status" would be useful.**
   - **Second, some statistical tests seem to have been performed between the 2 cohorts, but this is not clear between what and what. A p-value is provided in the text but not in Table 2.**
   - **Third, the color code might be wrongly used since the titles are presented in blue in Table 2 when this should represent lower incidence in the COVID-19 positive cohort compared to the COVID-19 negative cohort.**
   - **Table 2 should provide more information on the tests and comparison performed, possibly in footnotes. If it is red for a condition, is this red colour code applicable to "all", "severe" and "non-severe"?**
   - **Legend is unclear: "COVID-19 positive higher incidence" and might need to state what it is compared to.**
   - **Finally, to improve the structure of Table 2, the way the conditions are listed could be reorganized to have "systemic" and all the conditions that are included within "systemic" and then "syndromic" and the listing of all the conditions that are included in "syndromic".**

   **Our response**: Thank you for the suggestion. This table have been moved to supplementary as Table S6, and please see our response to each question as follows:

   - **"**First, the "severity status" is mentioned but has not yet been introduced to the reader. Perhaps a small description into parenthesis of what is this "severity status" would be useful."

     The severity of COVID-19 at cohort entry was stratified into the following categories: asymptomatic, mild (symptomatic), moderate (involving moderately severe COVID-19-related conditions like gastroenteritis, dehydration, and pneumonia), and severe (comprising unstable COVID-19-related conditions, ICU admission, or mechanical ventilation) [1]. In this paper, we categorized patients exhibiting either asymptomatic or mild symptoms as belonging to the

"non-severe" group, while all other patients were classified as part of the "severe" group. We included a footnote to this table in the updated manuscript, which reads:

*"Footnote:*
*1: According to the definition of the severity of COVID-19 [1] at cohort entry, the positive patients with moderate or severe COVID-19 were grouped into the severe group.*
*2: Similarly, the positive patients with asymptomatic or mild COVID-19 were grouped into the non-severe group."*

*Reference:*
*[1] Forrest CB, Burrows EK, Mejias A, et al. Severity of acute COVID-19 in children< 18 years old March 2020 to December 2021. Pediatrics. 2022;149(4):e2021055765.*

- "Second, some statistical tests seem to have been performed between the 2 cohorts, but this is not clear between what and what. A p-value is provided in the text but not in Table 2."

**Our response**: Thank you, we have added a p-value of two-sided Z test and footnotes to illustrate the p-value between which and which in the updated supplementary materials **Table S6**. This reads as follows:
***"****Footnotes:*
*[1]P-value for all: A two-sided Z-test for proportions comparing the COVID-19 positive cohort and the COVID-19 negative cohort.*
*[2]P-value for severe: A two-sided Z-test for proportions comparing the COVID-19 positive and negative cohorts among patients with severe cases during the acute phase.*
*[3]P-value for non-severe: A two-sided Z-test for proportions comparing the COVID-19 positive and negative cohorts among patients with non-severe cases during the acute phase.****"***

- "Third, the color code might be wrongly used since the titles are presented in blue in **Table 2** when this should represent lower incidence in the COVID-19 positive cohort compared to the COVID-19 negative cohort."

**Our response**: Thank you for the suggestion. We have decided to remove colors in this table in the updated supplementary materials.

- "Table 2 should provide more information on the tests and comparison performed, possibly in footnotes. If it is red for a condition, is this red colour code applicable to "all", "severe" and "non-severe"?"

**Our response**: Thank you for the suggestion. Originally, the red color code was applicable to column "All"; however, we decided to remove the color code, because for every outcome we are interested in, the COVID-19 positive group has a lower incidence compared to the COVID-19 negative group without considering the groups stratified by severity status. The updated table reads

- "Legend is unclear: "COVID-19 positive higher incidence" and might need to state what it is compared to."

  **Our response**: Thank you, the COVID-19 positive higher incidence was compared to COVID-19 negative patients. As this conclusion is consistent across all outcomes we are interested in, we removed this legend to avoid confusion.

- "Finally, to improve the structure of Table 2, the way the conditions are listed could be reorganized to have "systemic" and all the conditions that are included within "systemic" and then "syndromic" and the listing of all the conditions that are included in "syndromic"."

  **Our response**: Thank you for your suggestion. We have reorganized the condition list to group all these conditions into "systemic" and "syndromic", to improve the structure of this table. Finally we showed the updated Table S6:

  **Table S6**. Raw incidence (%) of PASC symptoms and conditions comparing COVID-19 positive and negative patients.

| | **All** | | | **Severe** | | | **Non-severe** | | |
|---|---|---|---|---|---|---|---|---|---|
| | Pos (%) | Neg (%) | p-value for all[1] | Pos (%) | Neg (%) | p-value for severe[2] | Pos (%) | Neg (%) | p-value for non-severe[3] |
| | (N=225,723) | (N=677,448) | | (N=16,862) | (N=143,592) | | (N=208,861) | (N=533,856) | |
| **At least one condition** | 26.86 | 21.44 | 0.27 | 34.17 | 19.80 | 0.34 | 26.42 | 21.84 | 0.26 |
| **Systematic conditions** | 2.69 | 1.87 | 0.03 | 7.46 | 2.93 | 0.07 | 2.37 | 1.61 | 0.02 |
| Acute kidney injury | 0.22 | 0.15 | 0.00 | 1.32 | 0.42 | 0.01 | 0.13 | 0.08 | 0.00 |
| Acute respiratory distress syndrome | 0.03 | 0.01 | 0.00 | 0.24 | 0.03 | 0.00 | 0.01 | 0.01 | 0.00 |
| Arrhythmias | 1.43 | 0.95 | 0.01 | 4.41 | 1.56 | 0.04 | 1.21 | 0.79 | 0.01 |
| Fluid and electrolyte | 0.56 | 0.39 | 0.01 | 3.33 | 0.92 | 0.03 | 0.36 | 0.24 | 0.00 |
| Heart disease | 0.37 | 0.28 | 0.00 | 1.67 | 0.62 | 0.02 | 0.27 | 0.20 | 0.00 |
| Myocarditis | 0.03 | 0.01 | 0.00 | 0.17 | 0.02 | 0.00 | 0.02 | 0.00 | 0.00 |
| Myositis | 0.02 | 0.02 | 0.00 | 0.08 | 0.02 | 0.00 | 0.02 | 0.01 | 0.00 |
| POTS/dysautonomia | 1.03 | 0.71 | 0.01 | 1.34 | 0.72 | 0.01 | 1.01 | 0.71 | 0.01 |
| Thrombophlebitis and thromboembolism | 0.13 | 0.11 | 0.00 | 1.02 | 0.33 | 0.01 | 0.06 | 0.05 | 0.00 |
| **Syndromic conditions** | 26.15 | 20.83 | 0.26 | 32.48 | 18.77 | 0.32 | 25.75 | 21.35 | 0.26 |
| Abdominal pain | 3.04 | 2.27 | 0.03 | 3.66 | 2.15 | 0.04 | 2.99 | 2.31 | 0.03 |
| Abnormal liver enzyme | 0.32 | 0.24 | 0.00 | 1.17 | 0.43 | 0.01 | 0.25 | 0.19 | 0.00 |
| Cardiovascular signs and symptoms | 1.16 | 0.88 | 0.01 | 1.69 | 0.86 | 0.02 | 1.12 | 0.89 | 0.01 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Changes in the taste and smell | 0.16 | 0.04 | 0.00 | 0.08 | 0.02 | 0.00 | 0.17 | 0.05 | 0.00 |
| Chest pain | 1.39 | 0.74 | 0.01 | 1.69 | 0.68 | 0.02 | 1.37 | 0.75 | 0.01 |
| Cognitive functions | 0.68 | 0.66 | 0.01 | 1.00 | 0.62 | 0.01 | 0.66 | 0.68 | 0.01 |
| Fatigue and malaise | 1.74 | 1.30 | 0.02 | 3.22 | 1.74 | 0.03 | 1.63 | 1.18 | 0.02 |
| Fever and chills | 5.67 | 3.73 | 0.06 | 9.15 | 3.31 | 0.09 | 5.41 | 3.84 | 0.05 |
| Generalized pain | 1.26 | 0.91 | 0.01 | 1.78 | 1.02 | 0.02 | 1.22 | 0.88 | 0.01 |
| Hair loss | 0.24 | 0.14 | 0.00 | 0.48 | 0.14 | 0.00 | 0.22 | 0.14 | 0.00 |
| Headache | 2.20 | 1.51 | 0.02 | 2.33 | 1.23 | 0.02 | 2.19 | 1.59 | 0.02 |
| Mental health | 6.32 | 5.45 | 0.06 | 8.04 | 5.05 | 0.08 | 6.19 | 5.55 | 0.06 |
| Musculoskeletal pain | 3.53 | 2.79 | 0.04 | 4.03 | 2.72 | 0.04 | 3.50 | 2.81 | 0.03 |
| Respiratory signs and symptoms | 9.68 | 7.25 | 0.10 | 15.02 | 6.77 | 0.15 | 9.31 | 7.38 | 0.09 |
| Skin symptoms | 3.62 | 2.70 | 0.04 | 4.99 | 2.66 | 0.05 | 3.51 | 2.71 | 0.04 |

Footnotes:

[1]P-value for all: A two-sided Z-test for proportions comparing the COVID-19 positive cohort and the COVID-19 negative cohort.

[2]P-value for severe: A two-sided Z-test for proportions comparing the COVID-19 positive and negative cohorts among patients with severe cases during the acute phase.

[3]P-value for non-severe: A two-sided Z-test for proportions comparing the COVID-19 positive and negative cohorts among patients with non-severe cases during the acute phase.

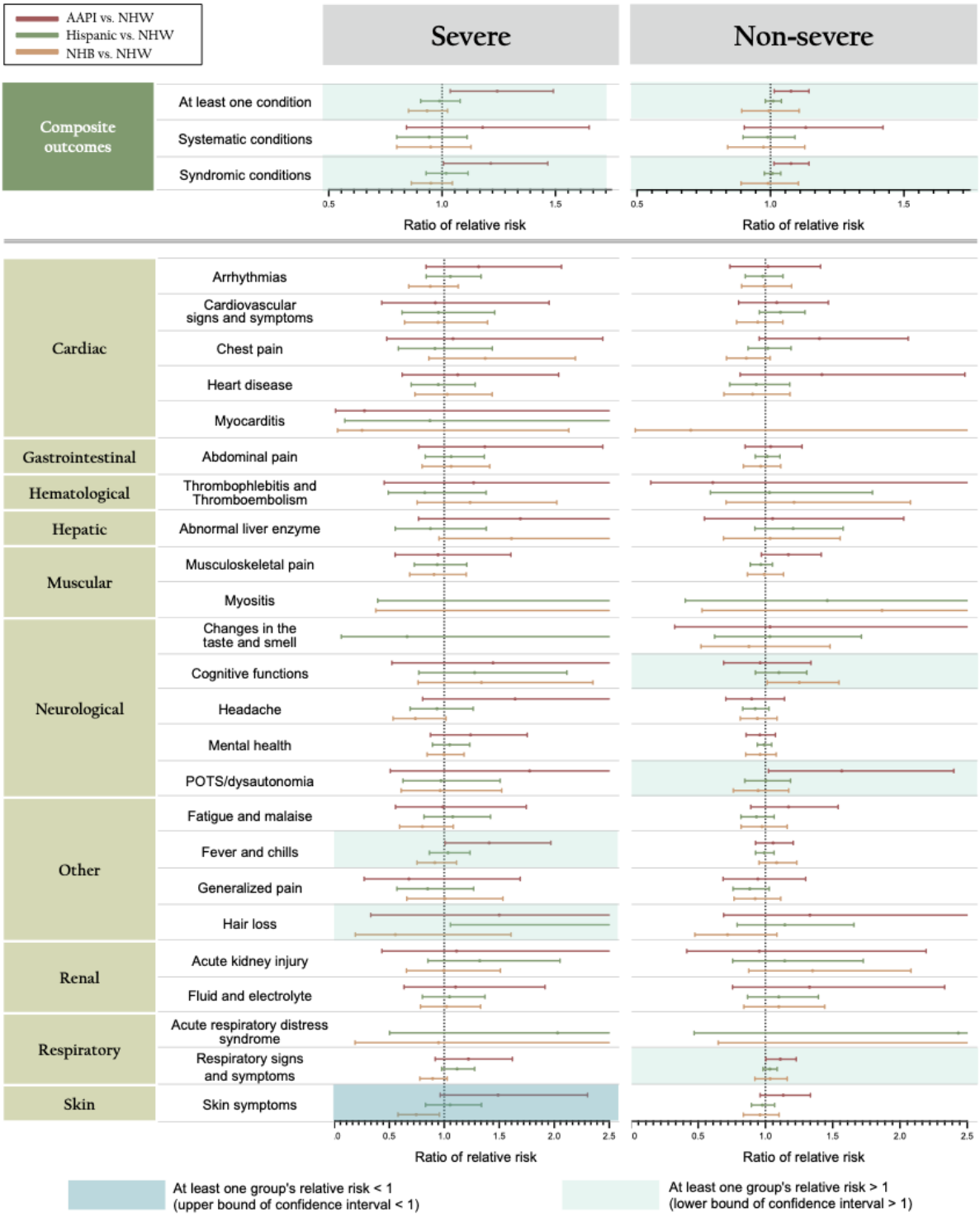
7. **- Page 4, line 146 – What does SMD stand for? Write in full. Readability will be improved if what has been done is clarified in the first sentence of this paragraph, even though this might have been detailed in methods/supplementary material.**

   **Our response**: Thank you, we have added standardized mean difference (SMD) and added a brief sentence to clarify what has been done at the beginning of this paragraph in the Results section in the updated manuscript:

   *"After achieving the balance of SMD by propensity score matching (Section S11) […]"*

8. **- Figure 3. It is unusual to have the legend of the figure in the middle of the figure. Maybe "at least one group's ratio of relative risk <1" and "at least one group's ratio of relative risk > 1" should be moved to the bottom of the figure. The legend is misleading because the interpretation and colour system are based on whether the confidence interval (CI) includes 1. So, it is not just about the value of the risk ratio (RR) here. Maybe the legend needs to be rephrased to be more accurate and mention the role of the CI in judging whether the RR is below or above 1.**

   **Our response**: Thank you for the suggestion. We have moved the legend at the bottom of the figure, and rephrased the legend as "at least one group's relative risk < 1 (upper bound of confidence interval < 1)" and "at least one group's relative risk > 1 (lower bound of confidence interval > 1)".




9. **- page 5 – line 158 and line 169 – the text says "Hispanic patients showed no increase in any of the conditions" but later says there is an increase in "hair loss" in that same group. There is a problem with logic and terminology. This "any of the conditions" terminology also**

**referred to "At least one condition" is an outcome that is not clear and might need to be replaced by another wording to ensure a clear interpretation of the findings.**

**Our response**: Sorry for the confusion. Yes, and we have updated "any of the conditions" to "at least one condition" in the updated manuscript.

As shown in **Figure S21**, it is possible for a population to exhibit no increase in the "at least one condition" composite outcome while still showing an increase in a specific condition. For instance, in the severe COVID-19 cohort comparing Hispanic and non-Hispanic White (NHW) patients, there is almost no difference between the post-infection and pre-infection periods for the "at least one condition" outcome. However, we observe a significant difference for hair loss between the post- and pre-infection periods. To further illustrate this point, we provide such figure detailed in Supplementary **Section S13**.

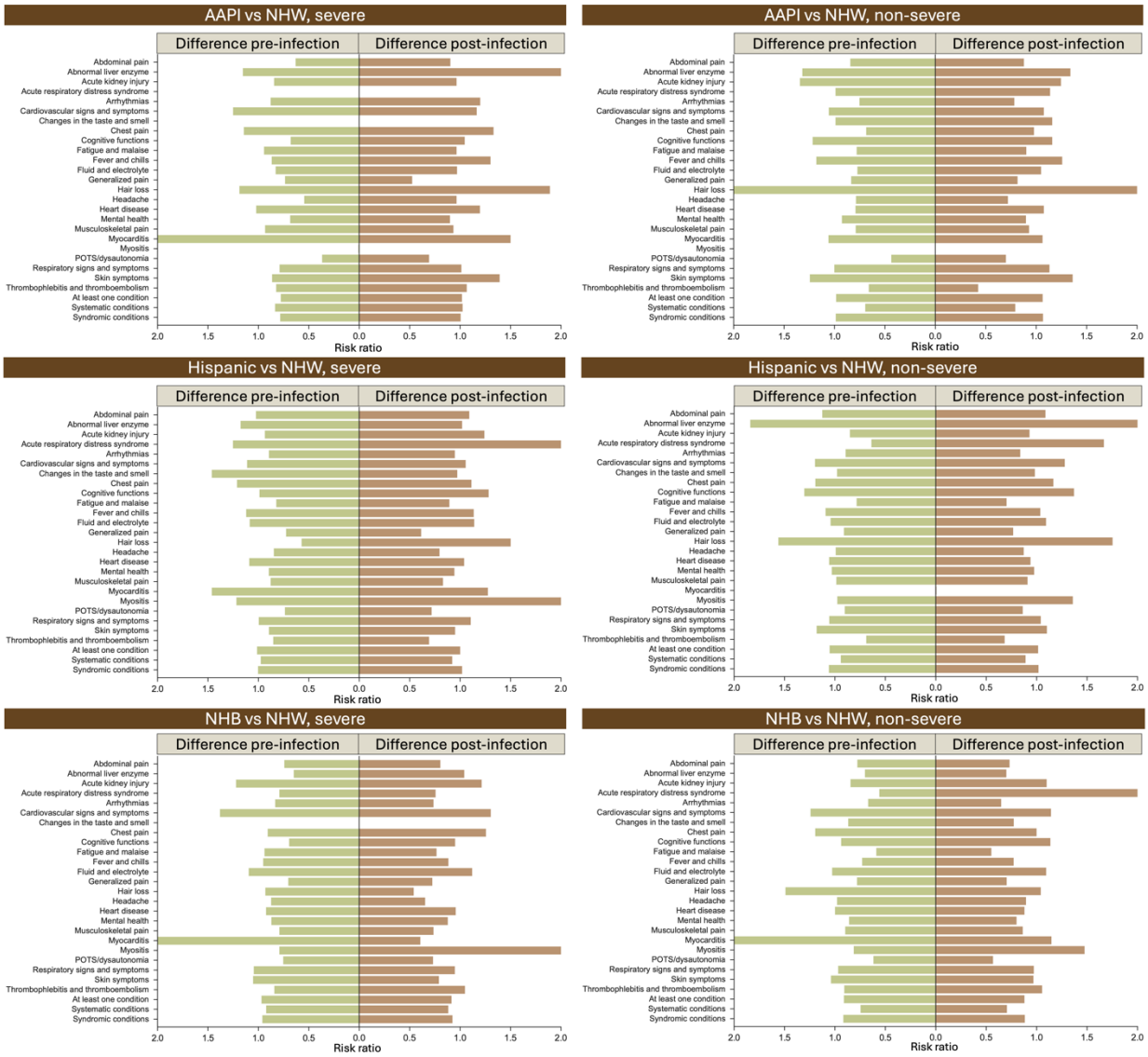



**Figure S21.** Relative risk (RR) differences for 24 individual outcomes and 3 composite outcomes

among COVID-19-positive patients, stratified by comparison groups and severity status across both pre- and post-infection periods. The RR results for racial/ethnic differences attributable to COVID-19 are derived from difference-in-difference analyses, calculated by taking the differences between post-infection (red lines) and pre-infection (green lines) periods.

10. **- Page 6, lines 224-225 - the supplementary file is large but some of the extra analysis performed are vaguely mentioned with no interpretation of the key findings. The supplementary file might need to be reviewed to include what is really necessary.**

    **Our response**: Thank you for the suggestion. We have added a paragraph discussing the key findings from the sensitivity analysis in the updated manuscript, reviewed and kept the items necessary in the updated supplementary materials, and enriched necessary interpretation of the key findings for each sensitivity analysis in the updated supplementary materials.

    *"**Section S9** shows the results of subgroup analysis by age group. No significant differences were observed for patients aged <5; For patients aged from 5 to 11, non-severe NHB patients showed a higher increase in any of the conditions (RR 1.08, 95% CI 1.02 to 1.15) and any of syndromic conditions (RR 1.08, 95% CI 1.02 to 1.15) compared to NHW group; And for adolescents aged 12 to 20, differences among severe and non-severe patients had a similar pattern. **Section S10** shows the results of stratified analysis by virus variants. The Hispanic cohort showed less increase among severe and non-severe groups during the pre-Delta period, differences were consistent in both sets during the Delta period, and all minor groups showed increased risk in any of the conditions and syndromic conditions compared to NHW patients during the Omicron period. **Section S12** provided the results of adding socio-economic confounder Area Deprivation Index (ADI) into the analysis, and the differences detected were consistent in both sets of analyses."*

11. **- Section S3 and Figure S13 are not referred to in the text of the manuscript. The title of Figure S13 might be misleading. Is it about using regression analysis only (to compare to DiD), or using both DiD and regression? What is the purpose of Figure S13?**

    **Our response**: We apologize for the confusion. What was previously Figure S13 is now Figure S14 in Section S3, which focuses solely on regression analysis. The purpose of this figure is to compare the results of the regression model with those of the DiD analysis. We have made these corrections in the updated manuscript.
    *"Figure S14 in Section S3 showed the racial/ethnic differences after SARS-CoV-2 infection stratified by severity of COVID-19, using standard regression models only."*

DISCUSSION

12. **The discussion is short and a few limitations/strengths could be added. Other points that could be added include:**
    - **Explain how the findings might be generalizable to the rest of the population.**
    - **Is the sample size robust enough to generate strong evidence? For example, it is interesting that there is a lack of racial/ethnic differences found in Figure 3 for most**

**racial/ethnic groups apart from AAPI compared to the NHW group for the combined outcomes. In such cases, we tend to think of an issue of small sample size leading to not being able to distinguish ethnic differences. However, AAPI is the smallest racial/ethnic group and for this group, the results show evidence of differences compared to NHW. Therefore, this strengthens the findings for the other groups. This point could be reflected in the discussion section.**

- **Discuss bias due to the way the cohorts have been selected.**
- **Asymptomatic cases are less likely to be captured because less likely to be tested and therefore this will have an impact in a specific direction to the findings for the non-severe COVID-19. Discuss how this affects your results.**
- **How could you capture more information on socio-economic differences to strengthen your analysis? As the discussion explains, socio-economic differences are a key element in understanding ethnic/racial differences in health. If you have individuals' addresses, could these be linked to geographical level data that can inform on an area-based measure of deprivation for example?**

**Our response**: Thank you for the suggestion. Please see our point-to-point response below:

- "Explain how the findings might be generalizable to the rest of the population."

We agreed it is vitally important to explain how the findings might be generalizable to the rest of the population. We highlighted the generalizability of the findings, emphasizing that the study's data, drawn from 13 institutions representing about 10% of U.S. children, provides a broad national sample. In addition, we illustrated the need for public health strategies to focus on reducing healthcare access disparities and improving pediatric care across all regions to better manage PASC in children. It also calls for ongoing surveillance and research to inform tailored public health policies for this population. We updated the manuscript and now it reads

*"The implications of this study on a broader public health level are significant. By utilizing data from 13 institutions, representing approximately 10% of US children and encompassing both urban and suburban health systems, this study offers a comprehensive national sample. Public health strategies must prioritize the reduction of healthcare access disparities and the enhancement of specialized pediatric care availability across all regions. This will help in better managing and treating PASC in children, ensuring that all affected patients receive the necessary care regardless of their geographic location or socio-economic status. Furthermore, these findings underscore the importance of continuous surveillance, research, and tailored public health policies to address the evolving needs of this patient group."*

- "Is the sample size robust enough to generate strong evidence? For example, it is interesting that there is a lack of racial/ethnic differences found in Figure 3 for most racial/ethnic groups apart from AAPI compared to the NHW group for the combined outcomes. In such cases, we tend to think of an issue of small sample size leading to not being able to distinguish ethnic differences. However, AAPI is the smallest racial/ethnic group and for this group, the results show evidence of differences compared to NHW. Therefore, this strengthens the findings for the other groups. This point could be reflected in the discussion section."

**Our response**: Thank you for highlighting this important point. We have incorporated this point into the Discussion section in the revised manuscript:

*"It is worth noting that AAPI is the smallest racial/ethnic group, and for this group, the results show evidence of differences compared to NHW. Therefore, this strengthens the findings of racial/ethnic differences across PASC symptoms and conditions for the other minor groups compared to NHW.*

*While the observed racial/ethnic differences in risk ratios are subtle, they remain significant when applied to large populations and may reflect broader systemic issues. For example, the observed differences in PASC can be attributed to differential access to health professionals, particularly pediatric subspecialists, which vary significantly across different regions of the country according to a recent National Academy of Medicine report.[2] These variations are not due to biological reasons but are more closely linked to socio-economic factors, healthcare infrastructure, and regional healthcare policies."*

*Reference:*
*[2] Rivara FP, Gonzalez-del-Rey J, Forrest CB. The pediatric subspecialty physician workforce. JAMA pediatrics. 2024 Feb 1;178(2):107-8.*

- "Discuss bias due to the way the cohorts have been selected."

**Our response**: Thank you for the suggestion. We have added a paragraph in the Discussion section regarding the bias due to the way that cohorts have been selected:

*"Fifthly, our study design, requiring at least one healthcare visit within 28-179 days post-index date, balanced selection bias against 'loss to follow-up' bias[39]. While this approach may over-represent frequent healthcare users, especially during 2020-2021, we mitigated this by incorporating diverse testing data to reduce COVID-19 status misclassification. Furthermore, in the sensitivity analysis (**Section S4**), the resulting selection bias likely included sicker patients in the COVID-19 negative group, potentially biasing our results towards the null. Nevertheless, we still observed a higher incidence of PASC symptoms and conditions for COVID-19-positive patients compared to COVID-19-negative patients. Future studies could explore alternative designs or data sources to enhance the external validity of these findings."*

*Reference:*
*[3] Howe CJ, Cole SR, Lau B, Napravnik S, Eron Jr JJ. Selection bias due to loss to follow up in cohort studies. Epidemiology. 2016;27(1):91-97.*

- "Asymptomatic cases are less likely to be captured because less likely to be tested and therefore this will have an impact in a specific direction to the findings for the non-severe COVID-19. Discuss how this affects your results."

**Our response**: Thank you for the suggestion. This study focused on diagnosed or treated cases of Long COVID. This data do not reflect the population experience; they reflect children who obtained care from health systems. We agree that asymptomatic cases can be less likely to be captured, which includes the potential for bias due to undocumented SARS-CoV-2 infections. Nevertheless, if this potential lack of data is evenly distributed across racial/ethnic groups, it could potentially attenuate the DiD analysis for the non-severe COVID-19 group, leading to conservative or inflation results. For example, suppose asymptomatic cases that are not captured have a lower rate of PASC symptoms and conditions after COVID infection, compared to captured cases. If more asymptomatic cases are missed in Hispanic than NHW groups, our conclusion will be more conservative. Conversely, more missed asymptomatic cases in NHW than Hispanic groups would result in inflated bias in our conclusions. Our inclusion of previous negative COVID-19 tests as a confounder aims to balance the probability of testing between different racial/ethnic groups which could partially adjust for this factor.

We updated our manuscript as follows:

*"Fourthly, asymptomatic cases are less likely to be captured as they are less likely to be tested, impacting the findings for non-severe COVID-19. This study focused on diagnosed or treated cases of Long COVID, reflecting children who sought healthcare services. The direction of bias can be bidirectional, depending on whether asymptomatic cases are disproportionately missed in minority or non-Hispanic White groups. This could attenuate the DiD analysis for the non-severe COVID-19 group due to potential misclassification bias, leading to conservative or inflated results. Future research should aim to capture a more comprehensive spectrum of COVID-19 severity."*

- "How could you capture more information on socio-economic differences to strengthen your analysis. As the discussion explains, socio-economic differences are a key element in understanding ethnic/racial differences in health. If you have individuals' addresses, could these be linked to geographical level data that can inform on an area-based measure of deprivation for example?"

**Our response**: Thank you for pointing this out. The data utilized in this study can be linked to geographical level data, i.e., Area Deprivation Index (ADI), but some institutions do not have this information. We conducted a sensitivity analysis including ADI as a measured confounder, which excluded patients who did not have this information. More details are included in **Table S27-28** in **Supplementary Section S10**. We also added a sentence summarizing the findings from this analysis:

*"Section S12 provided the results of adding socio-economic confounder Area Deprivation Index (ADI) into the analysis, and the differences detected were consistent in both sets of analyses."*

13. **- Page 6, line 229 – In the first summary sentence of the discussion, it would be useful to specify this is about children and younger individuals, so it flows better with the reference to evidence in adults appearing in a later sentence.**

    **Our response**: Thank you for the suggestion. We have revised the first summary sentence of the Discussion section in the updated manuscript as follows:

    *"We examined racial/ethnic variations in long-term consequences of documented SARS-CoV-2 infection across thirteen health institutions in the RECOVER study for 225,723 children and adolescents."*

14. **- Page 6, lines 233-234, there is a bit of redundancy between these two sentences. The statement that findings in younger individuals align with previous findings in adults should be mentioned once only.**

    **Our response**: We condensed these two sentences into one in our updated manuscript as follows:

    *"We observed specific symptom differences among racial/ethnic groups, despite no evidence indicating disparities in composite outcomes for NHB or Hispanic populations compared to NHW. For example, NHB patients showed a smaller change in skin symptoms compared to NHW patients, consistent with previous findings in adults by Khullar et al."*

15. **- Page 7, lines 255-259 – You can state the type of bias as "ascertainment bias" in this paragraph.**

    **Our response**: Thank you for the suggestion. We have revised the first sentence of this paragraph in the Discussion section:

    *"Secondly, health-seeking behavior or healthcare access is an important consideration, known as ascertainment bias."*

METHODS

16. **- Page 7, lines 278-285. Even though a link is provided, a few sentences should be added to describe what the data is about. Right now, this is unclear what the data contains and what are the exact sources. Is this about hospitalization data only (as the list of sites suggests)? Is it linked to other information from other sources (death, community testing, GP data, vaccination data, ...)? What type of data sources are used? Electronic Health Records or self-reported survey information or a mix of both?**

    **Our response**: Thank you for pointing this out. Data received from each institution is electronic health records (EHR) data, which includes clinical data such as diagnoses and treatments, laboratory and test results, and administrative data including patient demographics and billing information. Unlike General Practitioner (GP) data, self-reported data, and other data sources, the data used in this study were structured and standardized EHR entries made by healthcare providers

within hospital settings. More details have been included in Section S1 in the Supplementary Appendix, and we added a sentence at the end of this paragraph:

*"More details are available in the Supplementary Materials Section S1."*

And in **Section S1 A.** we added the following words

*"The real-world data utilized in our analysis is derived from electronic health records (EHRs), covering a wide range of healthcare interaction information routinely collected and stored by hospitals. This includes clinical data such as diagnoses and treatments, laboratory and test results, and administrative data including patient demographics and billing information. The hospital-based EHR data from the Researching COVID to Enhance Recovery (RECOVER) Initiative COVID-19 Database served as the basis for defining and determining exposure, outcomes, and covariates. Unlike General Practitioner (GP) data, self-reported data, or external data sources, our study used the structured, standardized EHR entries made by healthcare providers within hospital settings. EHR data provides a more detailed and integrated view of a patient's health status, medical history, and healthcare interactions across various providers and settings."*

17. **- Page 7, lines 292-293 – From which data source is the COVID-19 information collected/captured?**

**Our response**: The COVID-19 information was collected from observation, visit occurrence, and condition occurrence domains. A documented SARS-CoV-2 infection was defined as a polymerase-chain-reaction (PCR), serology, or antigen test positive for COVID-19, or diagnoses of COVID-19, post-acute sequelae of SARS-CoV-2 (PASC) regardless of the presence of symptoms. More details have been included in Table S2 in the Supplementary Appendix.

**Table S2**. Confounding Variables used in the propensity model and other variables for eligible criteria.

| Variable | Functional form | Values | Detail | Codes/references |
|---|---|---|---|---|
| **Documented SARS-CoV-2 infection** | Indicator | Yes/No | Based on the observation and visit occurrence, and condition occurrence domains. Defined as a polymerase-chain-reaction (PCR), serology, or antigen test positive for COVID-19, or diagnoses of COVID-19, post-acute sequelae of SARS-CoV-2 (PASC) regardless of the presence of symptoms. | See https://github.com/PEDSnet/PASC/tree/main/observation_derivation_recover_ml_phenotype/specs for detailed codes. |

18. **- Page 8, lines 296-298 – For the "COVID-19 negative patients", from which data source are they selected from? Have they all had a visit in the hospital? Are they meant to be**

**representative of the general population or a specific subpopulation of the general population? Could you clarify?**

**Our response**: The COVID-19 negative patients should also be a member of the health care organization during the previous 18 months and have at least one visit 28 to 179 days after the index date, as specified in the flowchart. These included subjects come from ED visits or other visits (i.e., routine, physical), primary care, office-visits, and others. Primary care settings typically provide a more comprehensive representation of the general population due to their broader patient demographics and the nature of routine and preventive care they offer. Thus, the presence of these major primary care-driven systems in our data source bolsters the external validity and generalizability of our results to the broader child and adolescent population. More details have been included in Section S3 in the Supplementary Appendix as follows:

*"For COVID-19-negative patients, we included patients who had neither a documented SARS-CoV-2 infection nor a diagnosis of multisystem inflammatory syndrome in children (MIS-C) within the same study period, and who had at least one negative COVID-19 test result. A random negative test was chosen as the index date for COVID-19-negative patients. The selection of participants for COVID-19-negative in real-world data is summarized in* ***Figure S13****. The COVID-19 negative patients were required to be members of the healthcare organization during the previous 18 months and have at least one visit 28 to 179 days after the index date, as specified in the flowchart. These included subjects come from ED visits or other visits (i.e., routine, physical), primary care, office visits, and others.*

*For COVID-19-negative patients, we determined severity status based on the symptoms and conditions observed during the acute phase (seven days before to thirteen days after the index date). Primary care settings typically provide a more comprehensive representation of the general population due to their broader patient demographics and the nature of routine and preventive care they offer. Thus, the presence of these major primary care-driven systems in our data source bolsters the external validity and generalizability of our results to the broader child and adolescent population."*

**19. Reference to figure 1 is not enough to explain the selection of the cohorts. This should be explained in a couple of sentences in the text as well.**

**Our response**: Thank you for pointing this out. We have revised this figure and added explanations for each selection criteria in the updated manuscript as follows:

*"We conducted a retrospective study from March 1, 2020, to October 3, 2022, with at least 6 months of follow-up time. We included patients under the age of 21 who had at least one visit within 18 months to 7 days before the index date (defined as the baseline period) and at least one encounter within 28 days and 179 days after the index date (defined as the follow-up period). For COVID-19-positive patients, we included the patients who had positive polymerase-chain-reaction (PCR), serology, or antigen tests or diagnoses of COVID-19, or diagnoses of PASC (U09.9), which*

*we defined as documented SARS-CoV-2 infection. The index date for these patients was defined as the first time of SARS-CoV-2 infection. We used patient race/ethnicity data included in the EHR and categorized patients into six racial/ethnic groups: non-Hispanic White (NHW), non-Hispanic Black (NHB), Hispanic, Asian American/Pacific Islanders (AAPI), Multiple, Other or Unknown. Multiple comprising less than 3% of the population in any COVID-19 and severity category (Table S4) were considered small samples and excluded from analysis. The "Other or Unknown" categories were also excluded due to interpretability issues. The selection of participants for COVID-19-positive patients in real-world data is summarized in* ***Figure 1****."*

20. **- Page 8, reference to figure 1 appears after reference to figures 2-3, figures will need to be renumbered**

    **Our response**: Thank you, we have renumbered this Figure in the updated manuscript.

21. **- Page 8, line 320, entry month is "March 2020 to October 2022". A clear range of information is collected at baseline (entry month). However, this is not clear whether patients are followed over time to collect information on PASC symptoms/conditions and over which period. More information on follow-up, censoring, median follow-up time for example should be described.**

    **Our response**: Thank you for the suggestion. In the inclusion-exclusion criteria, we include the patients who have at least one visit from 28 to 179 days after their entry. This can prevent the loss-to-follow-up bias [3]. If the patient does not have the follow-up during the 28 days to 179 days, we will exclude them from the cohort. In other words, the follow up time for each patient is the same. We have clarified this in the updated manuscript:

    *"We conducted a retrospective study from March 1, 2020, to October 3, 2022, with at least 6 months of follow-up time. We included patients under the age of 21 who had at least one visit within 18 months to 7 days before the index date (defined as the baseline period) and at least one encounter within 28 days and 179 days after the index date (defined as the follow-up period)."*

    *Reference:*
    *[3] Howe CJ, Cole SR, Lau B, Napravnik S, Eron Jr JJ. Selection bias due to loss to follow up in cohort studies. Epidemiology. 2016;27(1):91-97.*

22. **- Page 8, line 337. Incidence here is likely a measure of the number of individuals with new symptoms/conditions developed after baseline over all individuals, all followed over a specific period of time. Therefore, it might be more appropriate to use person-year as a denominator, especially if the follow-up period varies across individuals.**

    **Our response**: We decided to keep the term "incidence" as the follow-up period is the same for all individuals, that is, 28 to 179 days after entry. Since each person has the same follow-up period, using the person-year as the denominator and using the number of participants will lead to similar conclusions.

**23. - Page 9, line 356 – the end of the statistical methodology does not develop on how pre-infection racial/ethnic differences are accounted for and using which methodology. This should be described at the end of this paragraph, possibly after mention of Poisson regression.**

**Our response**: Thank you for the suggestion. The pre-infection racial/ethnic differences are accounted for by using difference-in-difference (DiD) analysis. Briefly speaking, the pre- and post-infection racial/ethnic differences in PASC symptoms and conditions are measured with the same metrics, and the racial/ethnic differences in PASC symptoms and conditions attributable to COVID-19 – the subject we are interested in – were calculated by the difference of pre- and post-infection outcomes. We have added "Statistical Method Overview" subsection in the Results part, and we also added "Difference-in-differences models" subsection in the Method part in the revised manuscript as follows:

*"We measured the pre-infection racial/ethnic differences by comparing PASC symptoms and conditions between minor racial/ethnic groups and the NHW group during the pre-infection period. We then measured post-infection differences using the same metrics. The racial/ethnic differences attributable to COVID-19 were determined by calculating the difference between these post-infection and pre-infection differences."*

*"Figure 2a provides a visual representation of the difference-in-difference method used to estimate racial/ethnic differences in the increased prevalence of PASC symptoms and conditions related to COVID-19. We introduce the notation to illustrate the difference-in-difference method. For each minor racial/ethnic group (AAPI, NHB, or Hispanic) matched with NHW cohort, we use a binary indicator R to denote racial/ethnic groups (R = 0 for NHW, R = 1 for NHB, Hispanic, or AAPI). We define T as the time period, where T = 1 represents the post-infection period (28 to 179 days after COVID-19 infection) and T = 0 represents the pre-infection period (28 to 179 days before COVID-19 infection). Y is a binary indicator denoting the occurrence of PASC symptoms and conditions, where Y = 1 indicates the patient has been diagnosed with the specific PASC symptoms and conditions, and Y = 0 indicates the absence of such diagnosis. For each participant, we observed the outcome Y in both pre-infection (T = 0) and post-infection (T = 1) periods.*

*We fitted a Poisson regression model by regressing each PASC symptom and condition on racial/ethnic groups, period, and the interaction between racial/ethnic groups and time period:*

$$\log(E[Y|R,T]) = \beta_0 + \beta_1 R + \beta_2 T + \beta_3 R \times T. \quad (1)$$

*where* $\log(\cdot)$ *is the log link function. From Equation (1), the pre-infection racial/ethnic difference is found by the coefficient,* $\beta_1$*, shown in the green part in Figure 2a. The post-infection racial/ethnic difference is denoted by the sum of the coefficients,* $\beta_1+\beta_3$*. Thus, the coefficient of the interaction term,* $\beta_3$*, represents the differential increase in racial/ethnic differences attributable to COVID-19, shown in the blue part in Figure 2a.*

*To illustrate the difference-in-difference model removes the bias from the unmeasured confounder when the parallel trends assumption holds, Figure 2b shows the directed acyclic graph (DAG) for the causal relationship when conducting the difference-in-difference model. We use* $Y_0$ *for the outcome during the pre-infection period,* $Y_1$ *for the outcome during the post-infection period, and*

*U for the unmeasured confounding variables. For simplicity, we present the DAG for the matched cohort of each minor race-NHW pair after adjusting for measured confounding variables X using the matching method."*

24. **- Page 9, line 379, when mention of "31 negative control" is given, a pointer to where to find these conditions in supplementary material should also be given more precisely (i.e. Table S1).**

**Our response**: Thank you, we have added a reference for the 31 negative control outcomes to Table S1 in the supplementary materials in the updated manuscript:

*"To mitigate such bias, we collected 31 negative control outcomes (detailed in **Table S1**), as prespecified by pediatric physicians, that should not exhibit racial/ethnic differences due to COVID-19."*

## Response to Reviewer 5:

**The study utilized data from RECOVER and examined differences in PASC across racial and ethnic groups in a pediatric population. The research question is very important, and the authors have sufficient data resources to address it. However, the paper requires significant improvements in clarity and is missing several key pieces of information. In its current form, the importance of the study is diminished due to unclear writing and a lack of focus on clinical questions.**

**Our response**: We sincerely appreciate your thoughtful review and recognition of the importance of our research question. We acknowledge the concerns you've raised regarding clarity and missing information, and we are committed to addressing these issues to enhance the impact of our study.

In response to your feedback, we have undertaken the following revisions:

1. Clarity improvements: We have thoroughly revised the manuscript based on your comments to improve clarity and readability, ensuring that our methodology, results, and discussions are presented in a more accessible manner.

2. Focus on clinical questions: We noticed there were some questions on the racial/ethnic differences in PASC symptoms and conditions between the COVID-19 positive and negative patients. Our main aim is to find the racial/ethnic differences attributable to COVID-19, that is, our focus is on the racial/ethnic PASC in pediatric populations across racial and ethnic groups for the COVID-19 positive patients.

3. Additional information: We have added several key pieces of information that were missing in the original manuscript. This includes the comparison of the difference-in-difference method and adjusting for the baseline PASC symptoms and conditions, the clarification on the negative control outcomes for the difference-in-difference analyses.

4. Restructuring: We have reorganized sections of the paper to improve the logical flow and enhance the emphasis on our main findings and their clinical relevance. For example, we added two paragraphs in the beginning of the subsection “Racial/ethnic differences in PASC symptoms and conditions” for the Results section.

We believe these revisions significantly strengthen the paper and address the concerns you've raised. We are grateful for the opportunity to improve our work and hope that these changes will more effectively communicate the importance of our findings to the scientific and clinical communities.

We welcome any further suggestions you may have for improving the manuscript.

1. **The study provided a comparison of outcomes across racial and ethnic groups, as well as the unadjusted rates in COVID-19 and control groups within the same racial and ethnic groups. However, it lacks the adjusted rates of outcomes (which represent PASC) by comparison between the COVID-19 and control groups. The incidence of PASC in this population itself**

**holds major clinical importance. Moreover, it is impossible to estimate the difference in PASC without determining whether the outcome qualifies as PASC in this population.**

**Our response**: Thank you for your insightful question. We acknowledge the clinical importance of PASC incidence in this population. However, our primary focus is on racial differences in PASC symptoms and conditions attributable to COVID-19. Specifically, within the COVID-19-positive cohort, we investigate racial/ethnic differences between pre-infection and post-infection periods.

Adjusted analysis for PASC symptoms and conditions falls outside our research scope. Several publications from the RECOVER consortium, using the same data as our study, have addressed PASC incidence among COVID-19 and control groups from adjusted analysis. Notably:

- Suchitra et al. [1] investigated the incidence and risk burdens for symptoms and conditions in children and adolescents, using similar selection criteria and COVID-19 status definitions. Their adjusted analysis confirmed the codeset for PASC symptoms and conditions.
- Razzaghi et al. [2] utilized these codesets as the outcomes to study vaccine effectiveness for long-COVID in children.
- Other publications (Zhou et al. [3] and Wu et al. [4]) have also employed these codesets.

We adopted the same codesets from [1] to [4]. To confirm the relevance of our specified PASC symptoms and conditions, we included a crude analysis as a sensitivity analysis (Section S4). Our comparison groups are different racial groups rather than COVID-19 and control groups. Therefore, we did not conduct adjusted analyses comparing COVID-19 positive and negative patients. To avoid confusion about the major clinical question, we moved the crude analysis to the supplementary materials, which reads

*"We calculated the incidence of PASC symptoms and conditions in both COVID-19 positive and negative cohorts stratified by severity. For each PASC symptom or condition, we calculated its incidence by dividing the number of patients who experienced the symptom or condition during the follow-up period but not at baseline by the total number of patients."*

*Reference:*

*[1] Rao, S., Lee, G. M., Razzaghi, H., Lorman, V., Mejias, A., Pajor, N. M., ... & Forrest, C. B. (2022). Clinical features and burden of postacute sequelae of SARS-CoV-2 infection in children and adolescents. JAMA pediatrics, 176(10), 1000-1009.*

*[2] Razzaghi H, Forrest CB, Hirabayashi K, Wu Q, Allen AJ, Rao S, Chen Y, Bunnell HT, Chrischilles EA, Cowell LG, Cummins MR. Vaccine effectiveness against long COVID in children. Pediatrics. 2024 Apr 1;153(4):e2023064446.*

*[3] ZHOU, T., Zhang, B., Zhang, D., Wu, Q., Becich, M., Blecker, S. B., ... & Chen, Y. (2024). Overweight and Obesity as Predictors of Post-acute Sequelae of SARS-Cov-2 Infection: Findings from the RECOVER Initiative. medRxiv, 2024-06.*

*[4] Wu Q, Zhang B, Tong J, Bailey LC, Bunnell HT, Chen J, Chrischilles EA, Christakis DA, Downs SM, Hirabayashi K, Jhaveri R. Real-world Effectiveness and Causal Mediation Study of BNT162b2 on Long COVID Risks in Children and Adolescents. medRxiv. 2024:2024-02.*

2. **The baseline characteristics of the non-COVID group, as well as all groups after matching, need to be presented. Additionally, please include the number of participants matched for each matching criterion.**

**Our response**: Thank you for your valuable suggestion. We have addressed your concerns by making the following additions to our manuscript:

1. Baseline characteristics of the non-COVID group: We have included these in Table S5 of the updated supplementary materials.
   The Table S5 in the Supplement reads as:

**Table S5**. Baseline characteristics of COVID-19 negative patients, by race/ethnicity and severity status.

| | ***Severe*** | | | | ***Non-severe*** | | | |
|---|---|---|---|---|---|---|---|---|
| | **NHW (N=87,222)** | **AAPI (N=5,709)** | **Hispanic (N=26,637)** | **NHB (N=24,024)** | **NHW (N=277,891)** | **AAPI (N=26,475)** | **Hispanic (N=123,049)** | **NHB (N=106,441)** |
| **Age categories (%)** | | | | | | | | |
| <1 years | 10895 (12.5%) | 711 (12.5%) | 3461 (13.0%) | 3978 (16.6%) | 24652 (8.9%) | 2313 (8.7%) | 11795 (9.6%) | 10748 (10.1%) |
| 1 to <5 years | 24363 (27.9%) | 1692 (29.6%) | 8024 (30.1%) | 7646 (31.8%) | 75138 (27.0%) | 7739 (29.2%) | 32990 (26.8%) | 31213 (29.3%) |
| 5 to <12 years | 23450 (26.9%) | 1727 (30.3%) | 7382 (27.7%) | 5815 (24.2%) | 82081 (29.5%) | 8493 (32.1%) | 41893 (34.0%) | 33954 (31.9%) |
| 12 to <16 years | 14085 (16.1%) | 771 (13.5%) | 4087 (15.3%) | 3175 (13.2%) | 45195 (16.3%) | 3128 (11.8%) | 19255 (15.6%) | 15311 (14.4%) |
| 16 to <21 years | 14429 (16.5%) | 808 (14.2%) | 3683 (13.8%) | 3410 (14.2%) | 50825 (18.3%) | 4802 (18.1%) | 17116 (13.9%) | 15215 (14.3%) |
| **Sex** | | | | | | | | |
| Female | 38451 (44.1%) | 2374 (41.6%) | 11735 (44.1%) | 10568 (44.0%) | 138595 (49.9%) | 13015 (49.2%) | 60993 (49.6%) | 52848 (49.7%) |
| Male | 48771 (55.9%) | 3335 (58.4%) | 14902 (55.9%) | 13456 (56.0%) | 139296 (50.1%) | 13460 (50.8%) | 62056 (50.4%) | 53593 (50.4%) |
| **Hospital** | | | | | | | | |
| A | 9036 (10.4%) | 344 (6.0%) | 872 (3.3%) | 2286 (9.5%) | 44879 (16.1%) | 1639 (6.2%) | 4010 (3.3%) | 15472 (14.5%) |
| B | 13650 (15.6%) | 1180 (20.7%) | 2933 (11.0%) | 5301 (22.1%) | 52929 (19.0%) | 4833 (18.3%) | 10083 (8.2%) | 23891 (22.4%) |
| C | 10948 (12.6%) | 524 (9.2%) | 6719 (25.2%) | 1210 (5.0%) | 16505 (5.9%) | 838 (3.2%) | 10727 (8.7%) | 2402 (2.3%) |
| D | 3475 (4.0%) | 271 (4.7%) | 927 (3.5%) | 1678 (7.0%) | 13149 (4.7%) | 1343 (5.1%) | 5287 (4.3%) | 8610 (8.1%) |
| E | 6184 (7.1%) | 154 (2.7%) | 532 (2.0%) | 492 (2.0%) | 11805 (4.2%) | 604 (2.3%) | 1543 (1.3%) | 1500 (1.4%) |
| F | 5702 (6.5%) | 692 (12.1%) | 4498 (16.9%) | 1704 (7.1%) | 7353 (2.6%) | 911 (3.4%) | 5971 (4.9%) | 2114 (2.0%) |
| G | 7301 (8.4%) | 434 (7.6%) | 623 (2.3%) | 1150 (4.8%) | 25525 (9.2%) | 2997 (11.3%) | 2245 (1.8%) | 3658 (3.4%) |
| H | 676 (0.8%) | 10 (0.2%) | 21 (0.1%) | 111 (0.5%) | 9893 (3.6%) | 292 (1.1%) | 554 (0.5%) | 1458 (1.4%) |
| I | 14095 (16.2%) | 860 (15.1%) | 1368 (5.1%) | 4838 (20.1%) | 34132 (12.3%) | 3274 (12.4%) | 5465 (4.4%) | 19021 (17.9%) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| J | 9492 (10.9%) | 436 (7.6%) | 4810 (18.1%) | 3527 (14.7%) | 26348 (9.5%) | 1667 (6.3%) | 15282 (12.4%) | 10672 (10.0%) |
| K | 430 (0.5%) | 42 (0.7%) | 701 (2.6%) | 189 (0.8%) | 17024 (6.1%) | 4133 (15.6%) | 50784 (41.3%) | 11677 (11.0%) |
| L | 1616 (1.9%) | 567 (9.9%) | 1764 (6.6%) | 412 (1.7%) | 7631 (2.7%) | 3183 (12.0%) | 8160 (6.6%) | 2507 (2.4%) |
| M | 4617 (5.3%) | 195 (3.4%) | 869 (3.3%) | 1126 (4.7%) | 10718 (3.9%) | 761 (2.9%) | 2938 (2.4%) | 3459 (3.2%) |
| **Entry time** | | | | | | | | |
| 03/2020 - 06/2020 | 6598 (7.6%) | 358 (6.3%) | 1611 (6.0%) | 1404 (5.8%) | 10273 (3.7%) | 714 (2.7%) | 3540 (2.9%) | 3039 (2.9%) |
| 07/2020 - 10/2020 | 13240 (15.2%) | 771 (13.5%) | 3371 (12.7%) | 2909 (12.1%) | 37038 (13.3%) | 2343 (8.8%) | 11387 (9.3%) | 9104 (8.6%) |
| 11/2020 - 02/2021 | 12002 (13.8%) | 698 (12.2%) | 3125 (11.7%) | 2802 (11.7%) | 46000 (16.6%) | 2981 (11.3%) | 14354 (11.7%) | 12656 (11.9%) |
| 03/2021 - 06/2021 | 13330 (15.3%) | 815 (14.3%) | 3909 (14.7%) | 3443 (14.3%) | 38195 (13.7%) | 3134 (11.8%) | 14268 (11.6%) | 14430 (13.6%) |
| 07/2021 - 10/2021 | 12747 (14.6%) | 902 (15.8%) | 3996 (15.0%) | 3962 (16.5%) | 52742 (19.0%) | 5802 (21.9%) | 25692 (20.9%) | 22646 (21.3%) |
| 11/2021 - 02/2022 | 11798 (13.5%) | 781 (13.7%) | 3822 (14.3%) | 3452 (14.4%) | 46495 (16.7%) | 5083 (19.2%) | 22561 (18.3%) | 19908 (18.7%) |
| 03/2022 - 06/2022 | 10277 (11.8%) | 779 (13.6%) | 3939 (14.8%) | 3408 (14.2%) | 28142 (10.1%) | 3927 (14.8%) | 17856 (14.5%) | 14132 (13.3%) |
| 07/2022 – 10/2022 | 7230 (8.3%) | 605 (10.6%) | 2864 (10.8%) | 2644 (11.0%) | 19006 (6.8%) | 2491 (9.4%) | 13391 (10.9%) | 10526 (9.9%) |
| **Obesity** | | | | | | | | |
| 0 | 59812 (68.6%) | 4299 (75.3%) | 16878 (63.4%) | 14958 (62.3%) | 190501 (68.6%) | 17846 (67.4%) | 54986 (44.7%) | 59931 (56.3%) |
| 1 | 27410 (31.4%) | 1410 (24.7%) | 9759 (36.6%) | 9066 (37.7%) | 87390 (31.4%) | 8629 (32.6%) | 68063 (55.3%) | 46510 (43.7%) |
| **PMCA** | | | | | | | | |
| 0 | 48327 (55.4%) | 3241 (56.8%) | 15290 (57.4%) | 13804 (57.5%) | 186253 (67.0%) | 19569 (73.9%) | 91266 (74.2%) | 71779 (67.4%) |
| 1 | 16614 (19.0%) | 1029 (18.0%) | 4902 (18.4%) | 4602 (19.2%) | 51283 (18.5%) | 3952 (14.9%) | 18525 (15.1%) | 20753 (19.5%) |
| 2 | 22281 (25.5%) | 1439 (25.2%) | 6445 (24.2%) | 5618 (23.4%) | 40355 (14.5%) | 2954 (11.2%) | 13258 (10.8%) | 13909 (13.1%) |
| **Negative tests prior entry** | | | | | | | | |
| 0 | 77433 (88.8%) | 5035 (88.2%) | 23626 (88.7%) | 20892 (87.0%) | 242041 (87.1%) | 23083 (87.2%) | 108613 (88.3%) | 92542 (86.9%) |
| 1 | 6553 (7.5%) | 455 (8.0%) | 2083 (7.8%) | 2141 (8.9%) | 25358 (9.1%) | 2423 (9.2%) | 10649 (8.7%) | 10170 (9.6%) |
| 2 | 1816 (2.1%) | 117 (2.0%) | 522 (2.0%) | 543 (2.3%) | 6537 (2.4%) | 596 (2.3%) | 2537 (2.1%) | 2460 (2.3%) |
| >=3 | 1420 (1.6%) | 102 (1.8%) | 406 (1.5%) | 448 (1.9%) | 3955 (1.4%) | 373 (1.4%) | 1250 (1.0%) | 1269 (1.2%) |
| **Vaccine dosage** | | | | | | | | |
| 0 | 79029 (90.6%) | 4866 (85.2%) | 23943 (89.9%) | 22208 (92.4%) | 243699 (87.7%) | 21074 (79.6%) | 105725 (85.9%) | 96004 (90.2%) |
| 1 | 1387 (1.6%) | 149 (2.6%) | 512 (1.9%) | 401 (1.7%) | 5670 (2.0%) | 923 (3.5%) | 3101 (2.5%) | 2242 (2.1%) |
| >=2 | 6806 (7.8%) | 694 (12.2%) | 2182 (8.2%) | 1415 (5.9%) | 28522 (10.3%) | 4478 (16.9%) | 14223 (11.6%) | 8195 (7.7%) |

2. Post-matching baseline characteristics: Tables S8-10 present the baseline characteristics after matching for different racial/ethnic groups within the COVID-19-positive cohort. These tables demonstrate that the baseline characteristics are well-balanced between the minoritized racial/ethnic groups and the non-Hispanic White (NHW) group.

**Table S8**. Baseline characteristics of COVID-19-positive patients after matching, by severity status, AAPI vs NHW.

| | ***Severe*** | | ***Non-severe*** | |
|---|---|---|---|---|
| | **NHW (N=3,834)** | **AAPI (N=575)** | **NHW (N=19,056)** | **AAPI (N=10,119)** |
| **Age categories (%)** | | | | |
| <1 years | 730 (19.0%) | 122 (21.2%) | 2044 (10.7%) | 965 (9.5%) |
| 1 to <5 years | 1374 (35.8%) | 190 (33.0%) | 5047 (26.5%) | 2791 (27.4%) |
| 5 to <12 years | 857 (22.4%) | 129 (22.4%) | 5888 (30.9%) | 3139 (30.8%) |
| 12 to <16 years | 388 (10.1%) | 59 (10.3%) | 2565 (13.5%) | 1429 (14.0%) |
| 16 to <21 years | 485 (12.7%) | 75 (13.0%) | 3512 (18.4%) | 1875 (18.4%) |
| **Sex** | | | | |
| Female | 1649 (43.0%) | 247 (43.0%) | 9368 (49.2%) | 4941 (48.4%) |
| Male | 2185 (57.0%) | 328 (57.0%) | 9688 (50.8%) | 5258 (51.6%) |
| **Hospital** | | | | |
| A | 366 (9.5%) | 37 (6.4%) | 1193 (6.3%) | 530 (5.2%) |
| B | 741 (19.3%) | 95 (16.5%) | 3940 (20.7%) | 1960 (19.2%) |
| C | 376 (9.8%) | 57 (9.9%) | 494 (2.6%) | 242 (2.4%) |
| D | 201 (5.2%) | 29 (5.0%) | 918 (4.8%) | 471 (4.6%) |
| E | 104 (2.7%) | 13 (2.3%) | 559 (2.9%) | 322 (3.2%) |
| F | 530 (13.8%) | 71 (12.3%) | 490 (2.6%) | 239 (2.3%) |
| G | 198 (5.2%) | 28 (4.9%) | 1549 (8.1%) | 742 (7.3%) |
| H | 14 (0.4%) | 1 (0.2%) | 216 (1.1%) | 109 (1.1%) |
| I | 640 (16.7%) | 85 (14.8%) | 2023 (10.6%) | 1098 (10.8%) |
| J | 239 (6.2%) | 33 (5.7%) | 1034 (5.4%) | 526 (5.2%) |
| K | 103 (2.7%) | 25 (4.3%) | 4478 (23.5%) | 2524 (24.7%) |
| L | 179 (4.7%) | 73 (12.7%) | 1615 (8.5%) | 1118 (11.0%) |
| M | 143 (3.7%) | 28 (4.9%) | 547 (2.9%) | 318 (3.1%) |
| **Entry time** | | | | |
| 03/2020 - 06/2020 | 84 (2.2%) | 17 (3.0%) | 278 (1.5%) | 185 (1.8%) |
| 07/2020 - 10/2020 | 111 (2.9%) | 16 (2.8%) | 777 (4.1%) | 433 (4.2%) |
| 11/2020 - 02/2021 | 290 (7.6%) | 47 (8.2%) | 2332 (12.2%) | 1240 (12.2%) |
| 03/2021 - 06/2021 | 251 (6.5%) | 32 (5.6%) | 900 (4.7%) | 445 (4.4%) |
| 07/2021 - 10/2021 | 261 (6.8%) | 36 (6.3%) | 1658 (8.7%) | 818 (8.0%) |
| 11/2021 - 02/2022 | 1224 (31.9%) | 188 (32.7%) | 7065 (37.1%) | 3726 (36.5%) |
| 03/2022 - 06/2022 | 794 (20.7%) | 117 (20.3%) | 3390 (17.8%) | 1914 (18.8%) |
| 07/2022 – 10/2022 | 819 (21.4%) | 122 (21.2%) | 2656 (13.9%) | 1438 (14.1%) |
| **Obesity** | | | | |
| 0 | 2704 (70.5%) | 411 (71.5%) | 11825 (62.1%) | 6112 (59.9%) |
| 1 | 1130 (29.5%) | 164 (28.5%) | 7231 (37.9%) | 4087 (40.1%) |
| **PMCA** | | | | |
| 0 | 2082 (54.3%) | 305 (53.0%) | 14456 (75.9%) | 7548 (74.0%) |
| 1 | 515 (13.4%) | 79 (13.7%) | 2663 (14.0%) | 1551 (15.2%) |
| 2 | 1237 (32.3%) | 191 (33.2%) | 1937 (10.2%) | 1100 (10.8%) |
| **Negative tests prior entry** | | | | |
| 0 | 2625 (68.5%) | 400 (69.6%) | 14738 (77.3%) | 7848 (76.9%) |
| 1 | 619 (16.1%) | 94 (16.3%) | 2651 (13.9%) | 1428 (14.0%) |
| 2 | 243 (6.3%) | 31 (5.4%) | 922 (4.8%) | 518 (5.1%) |
| >=3 | 347 (9.1%) | 50 (8.7%) | 745 (3.9%) | 405 (4.0%) |
| **Vaccine dosage** | | | | |

| | | | | |
|---|---|---|---|---|
| 0 | 3137 (81.8%) | 459 (79.8%) | 13882 (72.8%) | 7238 (71.0%) |
| 1 | 128 (3.3%) | 20 (3.5%) | 840 (4.4%) | 467 (4.6%) |
| >=2 | 569 (14.8%) | 96 (16.7%) | 4334 (22.7%) | 2494 (24.5%) |

**Table S9**. Baseline characteristics of COVID-19-positive patients after matching, by severity status, NHB vs NHW.

| | ***Severe*** | | ***Non-severe*** | |
|---|---|---|---|---|
| | **NHW (N=5,035)** | **NHB (N=2,911)** | **NHW (N=62,071)** | **NHB (N=39,156)** |
| **Age categories (%)** | | | | |
| <1 years | 1230 (24.4%) | 716 (24.6%) | 7520 (12.1%) | 4483 (11.4%) |
| 1 to <5 years | 1446 (28.7%) | 822 (28.2%) | 13844 (22.3%) | 8519 (21.8%) |
| 5 to <12 years | 849 (16.9%) | 477 (16.4%) | 17901 (28.8%) | 11822 (30.2%) |
| 12 to <16 years | 663 (13.2%) | 386 (13.3%) | 11149 (18.0%) | 7073 (18.1%) |
| 16 to <21 years | 847 (16.8%) | 510 (17.5%) | 11657 (18.8%) | 7259 (18.5%) |
| **Sex** | | | | |
| Female | 2311 (45.9%) | 1374 (47.2%) | 30815 (49.6%) | 19664 (50.2%) |
| Male | 2724 (54.1%) | 1537 (52.8%) | 31256 (50.4%) | 19492 (49.8%) |
| **Hospital** | | | | |
| A | 730 (14.5%) | 361 (12.4%) | 9711 (15.6%) | 5741 (14.7%) |
| B | 833 (16.5%) | 541 (18.6%) | 15704 (25.3%) | 9208 (23.5%) |
| C | 291 (5.8%) | 148 (5.1%) | 1628 (2.6%) | 826 (2.1%) |
| D | 336 (6.7%) | 221 (7.6%) | 4382 (7.1%) | 3796 (9.7%) |
| E | 96 (1.9%) | 41 (1.4%) | 2041 (3.3%) | 1021 (2.6%) |
| F | 373 (7.4%) | 212 (7.3%) | 1082 (1.7%) | 615 (1.6%) |
| G | 268 (5.3%) | 141 (4.8%) | 2559 (4.1%) | 1288 (3.3%) |
| H | 65 (1.3%) | 32 (1.1%) | 1383 (2.2%) | 698 (1.8%) |
| I | 766 (15.2%) | 449 (15.4%) | 7123 (11.5%) | 5169 (13.2%) |
| J | 684 (13.6%) | 421 (14.5%) | 5680 (9.2%) | 3242 (8.3%) |
| K | 92 (1.8%) | 57 (2.0%) | 6320 (10.2%) | 4974 (12.7%) |
| L | 131 (2.6%) | 82 (2.8%) | 1481 (2.4%) | 967 (2.5%) |
| M | 370 (7.3%) | 205 (7.0%) | 2977 (4.8%) | 1611 (4.1%) |
| **Entry time** | | | | |
| 03/2020 - 06/2020 | 113 (2.2%) | 83 (2.9%) | 746 (1.2%) | 679 (1.7%) |
| 07/2020 - 10/2020 | 136 (2.7%) | 89 (3.1%) | 2526 (4.1%) | 1621 (4.1%) |
| 11/2020 - 02/2021 | 412 (8.2%) | 238 (8.2%) | 8711 (14.0%) | 4974 (12.7%) |
| 03/2021 - 06/2021 | 417 (8.3%) | 264 (9.1%) | 5427 (8.7%) | 3686 (9.4%) |
| 07/2021 - 10/2021 | 675 (13.4%) | 416 (14.3%) | 9507 (15.3%) | 6039 (15.4%) |
| 11/2021 - 02/2022 | 1702 (33.8%) | 1001 (34.4%) | 22555 (36.3%) | 14807 (37.8%) |
| 03/2022 - 06/2022 | 799 (15.9%) | 407 (14.0%) | 6315 (10.2%) | 3559 (9.1%) |
| 07/2022 – 10/2022 | 781 (15.5%) | 413 (14.2%) | 6284 (10.1%) | 3791 (9.7%) |
| **Obesity** | | | | |
| 0 | 3174 (63.0%) | 1744 (59.9%) | 36791 (59.3%) | 20816 (53.2%) |
| 1 | 1861 (37.0%) | 1167 (40.1%) | 25280 (40.7%) | 18340 (46.8%) |
| **PMCA** | | | | |
| 0 | 2596 (51.6%) | 1490 (51.2%) | 43344 (69.8%) | 26630 (68.0%) |
| 1 | 806 (16.0%) | 473 (16.2%) | 11336 (18.3%) | 7651 (19.5%) |
| 2 | 1633 (32.4%) | 948 (32.6%) | 7391 (11.9%) | 4875 (12.5%) |
| **Negative tests prior entry** | | | | |
| 0 | 3549 (70.5%) | 2057 (70.7%) | 47779 (77.0%) | 30172 (77.1%) |
| 1 | 727 (14.4%) | 416 (14.3%) | 8987 (14.5%) | 5694 (14.5%) |
| 2 | 320 (6.4%) | 186 (6.4%) | 3045 (4.9%) | 1928 (4.9%) |
| >=3 | 439 (8.7%) | 252 (8.7%) | 2260 (3.6%) | 1362 (3.5%) |
| **Vaccine dosage** | | | | |
| 0 | 4537 (90.1%) | 2642 (90.8%) | 53628 (86.4%) | 34251 (87.5%) |
| 1 | 107 (2.1%) | 59 (2.0%) | 1627 (2.6%) | 1057 (2.7%) |
| >=2 | 391 (7.8%) | 210 (7.2%) | 6816 (11.0%) | 3848 (9.8%) |

**Table S10**. Baseline characteristics of COVID-19-positive patients after matching, by severity status, Hispanic vs NHW.

| | *Severe* | | *Non-severe* | |
|---|---|---|---|---|
| | **NHW (N=5,128)** | **Hispanic (N=3,392)** | **NHW (N=75,966)** | **Hispanic (N=33,064)** |
| **Age categories (%)** | | | | |
| <1 years | 950 (18.5%) | 702 (20.7%) | 8435 (11.1%) | 3865 (11.7%) |
| 1 to <5 years | 1643 (32.0%) | 1067 (31.5%) | 17687 (23.3%) | 7593 (23.0%) |
| 5 to <12 years | 1035 (20.2%) | 686 (20.2%) | 21630 (28.5%) | 9912 (30.0%) |
| 12 to <16 years | 700 (13.7%) | 427 (12.6%) | 13400 (17.6%) | 5946 (18.0%) |
| 16 to <21 years | 800 (15.6%) | 510 (15.0%) | 14814 (19.5%) | 5748 (17.4%) |
| **Sex** | | | | |
| Female | 2359 (46.0%) | 1544 (45.5%) | 38046 (50.1%) | 16476 (49.8%) |
| Male | 2769 (54.0%) | 1848 (54.5%) | 37920 (49.9%) | 16588 (50.2%) |
| **Hospital** | | | | |
| A | 378 (7.4%) | 154 (4.5%) | 6914 (9.1%) | 1224 (3.7%) |
| B | 622 (12.1%) | 315 (9.3%) | 18943 (24.9%) | 4280 (12.9%) |
| C | 1010 (19.7%) | 743 (21.9%) | 4404 (5.8%) | 3484 (10.5%) |
| D | 313 (6.1%) | 155 (4.6%) | 4607 (6.1%) | 2123 (6.4%) |
| E | 115 (2.2%) | 60 (1.8%) | 5278 (6.9%) | 1094 (3.3%) |
| F | 702 (13.7%) | 550 (16.2%) | 1904 (2.5%) | 1596 (4.8%) |
| G | 165 (3.2%) | 92 (2.7%) | 3958 (5.2%) | 768 (2.3%) |
| H | 27 (0.5%) | 16 (0.5%) | 1687 (2.2%) | 283 (0.9%) |
| I | 313 (6.1%) | 182 (5.4%) | 7016 (9.2%) | 1791 (5.4%) |
| J | 862 (16.8%) | 613 (18.1%) | 8151 (10.7%) | 5500 (16.6%) |
| K | 105 (2.0%) | 132 (3.9%) | 6783 (8.9%) | 7606 (23.0%) |
| L | 180 (3.5%) | 207 (6.1%) | 1940 (2.6%) | 2082 (6.3%) |
| M | 336 (6.6%) | 173 (5.1%) | 4381 (5.8%) | 1233 (3.7%) |
| **Entry time** | | | | |
| 03/2020 - 06/2020 | 115 (2.2%) | 105 (3.1%) | 854 (1.1%) | 983 (3.0%) |
| 07/2020 - 10/2020 | 167 (3.3%) | 133 (3.9%) | 4367 (5.7%) | 2350 (7.1%) |
| 11/2020 - 02/2021 | 428 (8.3%) | 313 (9.2%) | 12189 (16.0%) | 5043 (15.3%) |
| 03/2021 - 06/2021 | 379 (7.4%) | 263 (7.8%) | 5656 (7.4%) | 2120 (6.4%) |
| 07/2021 - 10/2021 | 553 (10.8%) | 361 (10.6%) | 9859 (13.0%) | 3959 (12.0%) |
| 11/2021 - 02/2022 | 1776 (34.6%) | 1184 (34.9%) | 26353 (34.7%) | 11411 (34.5%) |
| 03/2022 - 06/2022 | 927 (18.1%) | 538 (15.9%) | 8615 (11.3%) | 3429 (10.4%) |
| 07/2022 – 10/2022 | 783 (15.3%) | 495 (14.6%) | 8073 (10.6%) | 3769 (11.4%) |
| **Obesity** | | | | |
| 0 | 3280 (64.0%) | 2085 (61.5%) | 45211 (59.5%) | 16514 (49.9%) |
| 1 | 1848 (36.0%) | 1307 (38.5%) | 30755 (40.5%) | 16550 (50.1%) |
| **PMCA** | | | | |
| 0 | 2573 (50.2%) | 1729 (51.0%) | 52184 (68.7%) | 21931 (66.3%) |
| 1 | 749 (14.6%) | 474 (14.0%) | 13575 (17.9%) | 6385 (19.3%) |
| 2 | 1806 (35.2%) | 1189 (35.1%) | 10207 (13.4%) | 4748 (14.4%) |
| **Negative tests prior entry** | | | | |
| 0 | 3529 (68.8%) | 2379 (70.1%) | 57339 (75.5%) | 25655 (77.6%) |
| 1 | 741 (14.5%) | 480 (14.2%) | 11315 (14.9%) | 4557 (13.8%) |
| 2 | 317 (6.2%) | 200 (5.9%) | 4001 (5.3%) | 1623 (4.9%) |
| >=3 | 541 (10.6%) | 333 (9.8%) | 3311 (4.4%) | 1229 (3.7%) |
| **Vaccine dosage** | | | | |
| 0 | 4482 (87.4%) | 2991 (88.2%) | 63375 (83.4%) | 27436 (83.0%) |
| 1 | 122 (2.4%) | 75 (2.2%) | 2177 (2.9%) | 1040 (3.1%) |
| >=2 | 524 (10.2%) | 326 (9.6%) | 10414 (13.7%) | 4588 (13.9%) |

3. Number of matched participants: We have added information on the number of participants matched within the COVID-19-positive cohort in Table S7. It's important to note that our primary research question focuses on racial/ethnic differences within the COVID-19-positive cohort. Therefore, we have not included matching results for the COVID-19-negative cohort, as this falls outside the scope of our main clinical inquiry.

**Table S7**. Number of participants matched for each racial/ethnic group.

| | | Matching ratio | # of matched participants of minority groups | # of all participants of minority groups | # of matched participants in NHW group | # of all participants in NHW group |
|---|---|---|---|---|---|---|
| **Hispanic vs. NHW** | Severe | 2 | 3,392 | 3,786 | 5,128 | 9,140 |
| | Non-severe | 5 | 33,064 | 56,226 | 75,966 | 99,882 |
| **NHB vs. NHW** | Severe | 2 | 2,911 | 3,346 | 5,035 | 9,140 |
| | Non-severe | 2 | 39,156 | 42,477 | 62,071 | 99,882 |
| **AAPI vs. NHW** | Severe | 8 | 575 | 590 | 3,834 | 9,140 |
| | Non-severe | 2 | 10,199 | 10,276 | 1,9056 | 99,882 |

These additions provide a more comprehensive view of our study population and matching process, enhancing the transparency and interpretability of our results. We believe these changes address your concerns while maintaining our focus on the primary research question.

3. **The study calculated the incidence rate by dividing the number of patients experiencing an incident outcome by the total number of patients (lines 336-337). Such a comparison of incidence rates presents a clear bias when comparing groups with different follow-up times and when comparing groups with different baseline prevalences of outcomes. Unless I have overlooked it, I did not see any effort mentioned in the manuscript to balance the follow-up or to adjust for the baseline prevalence of outcomes.**

**Our response**: Thank you for your important question regarding our incidence rate calculation and potential biases. We appreciate the opportunity to clarify our methodology. Our study design addresses the concerns you've raised in the following ways. For the follow-up period, we observed PASC symptoms and conditions for a consistent follow-up period of five months (28 to 179 days after cohort entry) for each individual. This standardization ensures that all participants have equal observation time, mitigating bias from different follow-up durations. For incidence calculation, in our crude analysis, we calculated incidence by dividing the number of new cases within the follow-up period (numerator) by the number of people at risk for that PASC symptoms or conditions (denominator). This approach is appropriate given our standardized follow-up period. In our main analysis, we employed a difference-in-difference (DiD) method on the COVID-19 positive cohort. This approach inherently adjusts for the baseline prevalence of outcomes by calculating the difference in outcomes between the post- and pre-infection periods, as illustrated in the DAG in the response of Q5. This method addresses your concern about potential bias arising from differing baseline prevalences across groups.

By using these methodological approaches, we have taken steps to balance follow-up time and account for baseline prevalences, thus minimizing the biases you've identified. We have updated our discussion section to more explicitly describe these important aspects of our study design:

*"A key strength of our study is the implementation of NCOs, which enhances the reliability and interpretability of our findings. We used the NCOs which are outcomes not expected to be affected by COVID-19 or show racial/ethnic differences due to COVID-19. By analyzing these alongside our primary outcomes, we enable the detection of systematic bias from unmeasured confounding variables. This approach allows us to calibrate results from the DiD model. Importantly, NCOs help mitigate the impact of unmeasured confounding that may change before and after the infection period. This is particularly crucial when studying racial/ethnic disparities, where many dynamic social and environmental factors may not be captured in our dataset. Furthermore, we carefully addressed potential bias related to differences in follow-up times. To ensure comparability, we standardized the follow-up period for all participants, observing PASC symptoms and conditions over a consistent timeframe of five months (28 to 179 days post-cohort entry). This approach mitigates any bias that might arise from variations in follow-up duration across individuals."*

4. **The 'DID' analyses presented by the authors are very unclear. According to the manuscript, the study only conducted pairwise matching and did not match between the COVID and non-COVID groups. A valid comparison of PASC between two racial groups would require balanced baseline characteristics both between the COVID and non-COVID groups, and between racial groups within each COVID-19 status. This would require multiple arms matching, which does not appear to be conducted. Given the differences in baseline characteristics between the COVID and non-COVID groups, the beta2 and beta3 (which, beta3 are the main results of the study) in the author's 'DID' formula could be biased by confounding effects.**
   **Please further justify the methodology and provide a clear explanation of how confounding effects between COVID and non-COVID were adjusted for. Additionally, justify the decision to conduct matching separately instead of employing other statistical methods such as weighting or multiple treatment matching, which could provide comparisons within the same target population.**

   **Our response**: Thank you for the question. In the difference-in-difference (DiD) analysis, only the COVID-19-positive cohort was included, and the DiD is considering racial/ethnic differences in PASC before-and-after COVID infection for the COVID-19 population. The COVID-19-negative group was only used in the crude analysis, not the main DiD analysis.

   We noticed a misspecification for the $\beta_2$ in the formula shown in the Supplementary Materials. Before, we used the $\beta_2$ to denote the difference between the COVID-19 positive versus COVID-19 negative patients, which is an error.

We made the following correction: we used $\beta_2$ to denote the difference between the post-infection period versus pre-infection period for the reference group (major racial/ethnic group, or NHW group) of COVID-19-positive patients.

To clarify the DiD model, we used the following notations. For each minor racial/ethnic group(AAPI, NHB, or Hispanic) matched with NHW group within the COVID-19 positive cohort, we use a binary indicator R to denote racial/ethnic groups ( $R = 0$ for NHW, $R = 1$ for NHB, Hispanic, or AAPI). We define T as the time period, where $T = 1$ represents the post-infection period (28 to 179 days after COVID-19 infection) and $T = 0$ represents the pre-infection period (28 to 179 days before COVID-19 infection). Y is a binary indicator denoting the occurrence of PASC symptoms and conditions, where $Y = 1$ indicates the patient has been diagnosed with the specific PASC symptoms and conditions, and $Y = 0$ indicates the absence of such diagnosis. For each participant, we observed the outcome in both pre-infection ($T = 0$) and post-infection ($T = 1$) periods.

**Figure 2a** in the revised version showed the rationale of the DiD model. We fitted a Poisson regression model by regressing each PASC symptom and condition on racial/ethnic groups, period, and the interaction between racial/ethnic groups and time period:

$$\log(E[Y|R,T]) = \beta_0 + \beta_1 R + \beta_2 T + \beta_3 R \times T. \quad (1)$$

where $\log(\cdot)$ is the log link function. From Equation (1), the pre-infection racial/ethnic difference is found by the coefficient, $\beta_1$, shown in the green part in **Figure 2a**. The post-infection racial/ethnic difference is denoted by the sum of the coefficients, $\beta_1 + \beta_3$. Thus, the coefficient of the interaction term, $\beta_3$, represents the differential increase in racial/ethnic differences attributable to COVID-19, shown in the blue part in **Figure 2a**.

Within the COVID-19 positive cohort, we matched the minor racial/ethnic group to the NHW group on the measured confounding variables during the baseline period (2 years to 180 days before the cohort entry date). The matching on the baseline confounding variables strategy before applying difference-in-difference model has been used in [1], [2], and [3]. As we only used the COVID-19 positive cohort, the confounding bias for the difference between the COVID-19-negative and positive cohort will not be introduced.

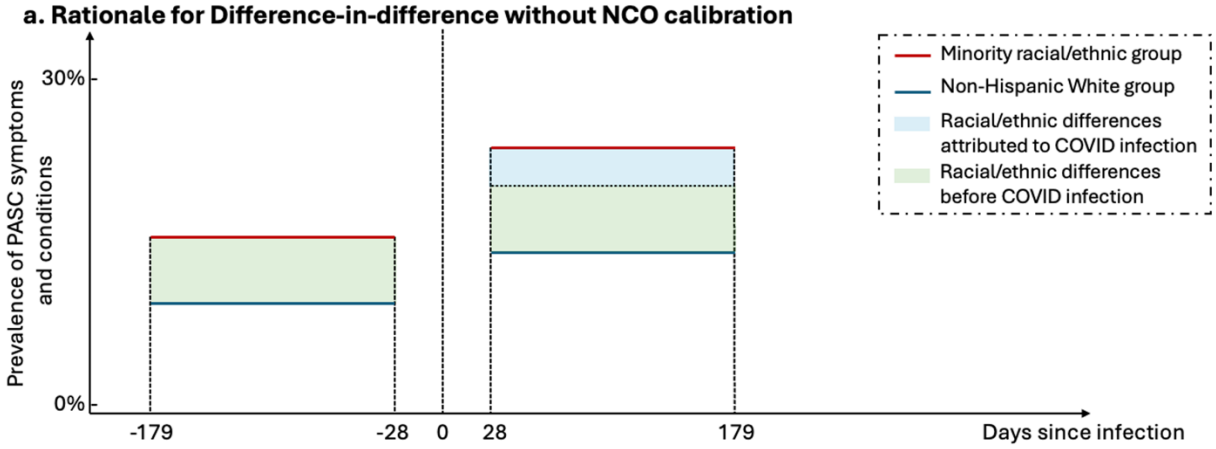


**Figure 2**. (a) shows the rationale for the difference-in-difference without the NCO calibration. The red line represents the prevalence for the PASC symptoms and conditions, and the blue line represents the non-Hispanic White group's PASC symptoms and conditions.

We agreed that multi-arm matching is a powerful tool to compare different racial/ethnic groups within the COVID-19 positive group. Before using pairwise matching in the statistical protocol, we conducted a literature search comparing pairwise matching versus multi-arm matching. First, we noticed that after multi-arm matching, the number in each treatment group would be determined by the smallest group among the minor racial/ethnic groups [4]. We observed that AAPI is the smallest at 5% of the group, which would harm statistical power. However, with pairwise matching, the number would be decided by the smaller one for each pair [5]. Second, the matching rate for multi-arm matching without replacement is low, which would further harm statistical power. Third, we noticed that the clinical interpretation for pairwise matching versus multi-arm matching differs. Pairwise matching can be used to draw clinical conclusions for each minor group compared to the NHW group. However, comparisons between different minor racial groups cannot be directly inferred from each minor-to-NHW comparison because the NHW population for each pairwise matching is different [4]. We will add this limitation to the discussion section.

Moreover, when comparing the method of matching with the weighting method, we also had literature research. First, matching creates directly comparable groups that mimic a randomized controlled trial, making results easier to interpret and communicate [5], [6]. Matching with calipers can handle extreme propensity scores better than weighting, which can produce extreme weights leading to high variance in estimates [6], [7]. Last, [8] found that in certain contexts, propensity score matching demonstrated a slightly superior ability to reduce covariate imbalance compared to inverse probability of treatment weighting.

We changed main manuscript accordingly in the Methods Section:

"*Figure 2a provides a visual representation of the difference-in-difference method used to estimate racial/ethnic differences in the increased prevalence of PASC symptoms and conditions related to COVID-19. We introduce the notation to illustrate the difference-in-difference method. For each minor racial/ethnic group (AAPI, NHB, or Hispanic) matched with NHW cohort, we use a binary indicator R to denote racial/ethnic groups (R = 0 for NHW, R = 1 for NHB, Hispanic, or AAPI). We define T as the time period, where T = 1 represents the post-infection period (28 to 179 days after COVID-19 infection) and T = 0 represents the pre-infection period (28 to 179 days before COVID-19 infection). Y is a binary indicator denoting the occurrence of PASC symptoms and conditions, where Y = 1 indicates the patient has been diagnosed with the specific PASC symptoms and conditions, and Y = 0 indicates the absence of such diagnosis. For each participant, we observed the outcome Y in both pre-infection (T = 0) and post-infection (T = 1) periods.*

*We fitted a Poisson regression model by regressing each PASC symptom and condition on racial/ethnic groups, period, and the interaction between racial/ethnic groups and time period:*

$$\log(E[Y|R,T]) = \beta_0 + \beta_1 R + \beta_2 T + \beta_3 R \times T. \quad (1)$$

*where* $\log(\cdot)$*is the log link function. From Equation (1), the pre-infection racial/ethnic difference is found by the coefficient,* $\beta_1$*, shown in the green part in Figure 2a. The post-infection racial/ethnic difference is denoted by the sum of the coefficients,* $\beta_1+\beta_3$*. Thus, the coefficient of the interaction term,* $\beta_3$*, represents the differential increase in racial/ethnic differences attributable to COVID-19, shown in the blue part in Figure 2a."*

*Reference:*

*[1] Burden BC, Canon DT, Mayer KR, Moynihan DP. Election laws, mobilization, and turnout: The unanticipated consequences of election reform. Am J Pol Sci 2014; 58: 95–109.*

*[2] Stuart EA, Huskamp HA, Duckworth K, et al. Using propensity scores in difference-in-differences models to estimate the effects of a policy change. Health Serv Outcomes Res Methodol 2014;* ***14****: 166–82.*

*[3] Imai K, Kim IS, Wang EH. Matching methods for causal inference with time-series cross-sectional data. Am J Pol Sci 2023; 67: 587–605.*

*[4] Lopez, M. J., & Gutman, R. (2017). Estimation of causal effects with multiple treatments: a review and new ideas. Statistical Science, 432-454.*

*[5] Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics, 25(1), 1.*

*[6] Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate behavioral research, 46(3), 399-424.*

*[7] Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge university press.*

*[8] Austin, P. C. (2009). Type I error rates, coverage of confidence intervals, and variance estimation in propensity-score matched analyses. The international journal of biostatistics, 5(1).*

5. **Figure 2 shows the advantage of the study in examining the difference in outcomes between two groups during follow-up after removing the differences that existed at baseline. Should this be considered a basic concept of causal inference? How is it more advanced compared to studies that balance the baseline prevalence of outcomes through the propensity score method or g-estimation? Or compared to studies that estimate purely incident outcomes in populations without the outcome at baseline?**

**Our response:** Thank you for your insightful questions regarding the DiD model. We appreciate the opportunity to clarify how our study advances the methodology in causal inference.

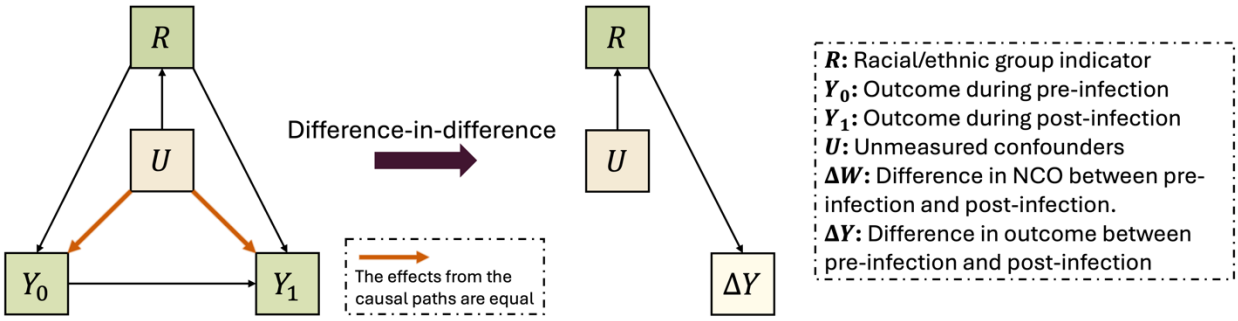
DiD has indeed played a crucial role within the causal inference framework [9]. To illustrate the advantages of the DiD model, we have used a directed acyclic graph (DAG) [10] to represent the causal mechanism between the racial/ethnic indicator (R), outcome during the pre-infection period ($Y_0$), outcome during the post-infection period ($Y_1$), and unmeasured confounding variables (U). For simplicity, we present the DAG for the matched cohort of each minor race-NHW pair after adjusting for measured confounding variables (X) using the matching method.

Scenario (b) shows the DAG underlying the DiD model. Arrows between variables denote causal relationships. For instance, the arrow from R to $Y_1$ represents the racial/ethnic difference during the post-infection period, while the arrow from R to $Y_0$ represents the difference during the pre-infection period. The DiD model aims to identify the racial/ethnic difference in the outcome change between the post-infection and the pre-infection period, ΔY (where $\Delta Y = Y_1 - Y_0$), as shown by the arrow from R to ΔY.

The key advantage of the DiD model is its ability to adjust for unmeasured confounding variables U under the parallel trends assumption [11]. The parallel trend assumption in our study states that the racial/ethnic difference does not change over time under the absence of COVID-19 infection. This assumption implies that the effect of U remains constant over time, meaning the effect of U on $Y_1$ and $Y_0$ is equal. In the DAG, this is represented by $U \rightarrow Y_0$ being equal to $U \rightarrow Y_1$. Under this assumption, the DiD model, by taking the difference between $Y_1$ and $Y_0$, effectively blocks the path from U to ΔY, allowing us to isolate the effect of R on ΔY without interference from U.

In contrast, simply adjusting for pre-infection outcomes $Y_0$ would effectively block the path $R \rightarrow Y_0 \rightarrow Y_1$. However, it fails to address the backdoor path $R \leftarrow U \rightarrow Y_1$. Additionally, it introduces new bias through the collider path $R \rightarrow Y_0 \leftarrow U \rightarrow Y_1$. Collider bias occurs when two variables, such as R and U, influence a common effect, such as $Y_0$, and controlling for this effect inadvertently creates a spurious association between the variables, leading to biased results. This is why the DiD method is more accurate than the model adjusting for the outcome during the pre-infection period.

**b. DAG for Difference-in-difference without NCO calibration**




*Reference:*
*[9] Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.*
*[10] Pearl, J. (2009). Causality. Cambridge university press.*
*[11] Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies, 72(1), 1-19.*

In the Methods Section of the main manuscript, we added the following paragraphs
*"To illustrate the difference-in-difference model removes the bias from the unmeasured confounder when the parallel trends assumption holds,* ***Figure 2b*** *shows the directed acyclic graph (DAG) for*

*the causal relationship[44] when conducting the difference-in-difference model. We use $Y_0$ for the outcome during the pre-infection period, $Y_1$ for the outcome during the post-infection period, and U for the unmeasured confounding variables. For simplicity, we present the DAG for the matched cohort of each minor race-NHW pair after adjusting for measured confounding variables X using the matching method.*

*Arrows between variables denote causal relationships. For instance, the arrow from R to $Y_1$ represents the racial/ethnic difference during the post-infection period, while the arrow from R to $Y_0$ represents the difference during the pre-infection period. The DiD model aims to identify the racial/ethnic difference in the outcome change between the post-infection and the pre-infection period, $\Delta Y$ (where $\Delta Y = Y_1 - Y_0$), as shown by the arrow from R to $\Delta Y$.*

*The key advantage of the DiD model is its ability to adjust for unmeasured confounding variables U under the parallel trends assumption. This assumption states that the effect of U remains constant over time, meaning the effect of U on $Y_0$ and $Y_1$ is equal. In the DAG, this is represented by the path $U \rightarrow Y_0$ being equal to $U \rightarrow Y_1$.*

*Under this assumption, the DiD model, by taking the difference between $Y_1$ and $Y_0$, effectively blocks the path from U to $\Delta Y$, allowing us to isolate the effect of R on $\Delta Y$ without interference from U. It is worth to note that a specified model with adjusting for the pre-infection outcomes $Y_0$ cannot adjust for the bias from the unmeasured confounder U. The model effectively blocks the path $R \rightarrow Y_0 \rightarrow Y_1$. Additionally, it introduces new bias through the collider path $R \rightarrow Y_0 \leftarrow U \rightarrow Y_1$. Collider bias occurs when two variables, such as R and U, influence a common effect, such as $Y_0$, and controlling for this effect inadvertently creates a spurious association between the variables, leading to biased results. This is why the DiD method is more accurate than the model adjusting for the outcome during the pre-infection period."*

6. **More clarity is needed to understand how this study represents an advancement in methodology. Moreover, given the lack of clarity, it is also unclear how the application of negative controls used in the study is superior to existing methods around empirical calibration.**

**Our response**: Thank you for your question. We appreciate the opportunity to clarify how our study advances the methodology and how our application of negative controls offers advantages over existing empirical calibration methods.

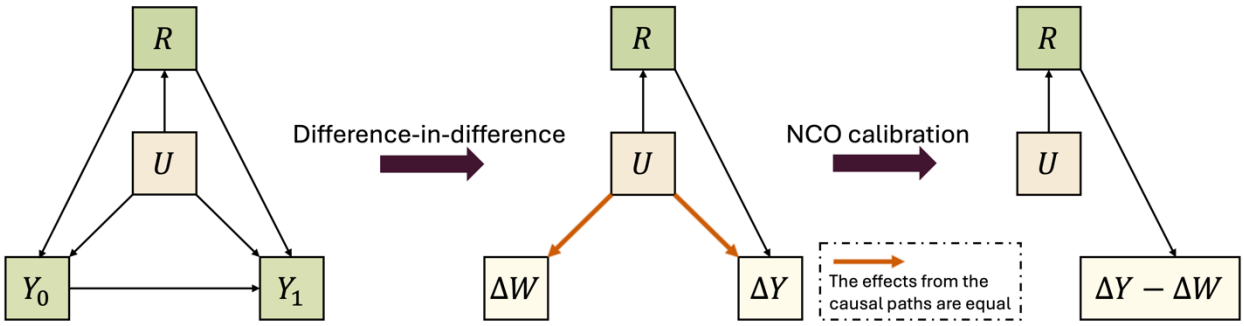
This study represents a methodological advancement by combining DiD analysis with negative control outcome (NCO). To explicitly demonstrate why the DiD approach with NCO calibration is superior in handling systematic bias compared to methods that adjust for baseline prevalence of outcomes, we have expanded on the DAG presented in our response to Comment 5.

Figure Scenario (c), illustrates this expanded DAG. As discussed in Comment 5, the standard DiD approach without NCO calibration assumes that the effect of unmeasured confounders (U) on the pre-infection outcome ($Y_0$) is equal to its effect on the post-infection outcome ($Y_1$) - this is known

as the parallel trends assumption. However, when this assumption is violated, a backdoor path from U to ΔY (where $\Delta Y = Y_1 - Y_0$) can emerge, potentially biasing our estimates.

To address this limitation, we incorporate NCOs into our analysis. Due to the influence of U, we observe a path from U to ΔW (where $\Delta W = W_1 - W_0$, and W represents the NCO). We then make the crucial assumption that the effect of U on ΔW is equivalent to its effect on ΔY. By leveraging this assumption, we can effectively block the path from U to $\Delta Y - \Delta W$. This allows us to estimate the racial/ethnic difference in PASC symptoms and conditions with greater accuracy and robustness to unmeasured confounding and violation of parallel trends assumption.

**c. DAG for Difference-in-difference with NCO calibration**




Below are statistical details to incorporate M NCOs to estimate the systematic bias b from the unmeasured confounder U (shown in the path U → ΔW within the model. We first apply the DiD model using the mth NCO ($m = 1, \ldots, M$) as the outcome, shown as,

$$log[E(W_m \mid R, T)] = \alpha_{0,m} + \alpha_{1,m}R + \alpha_{2,m}T + b_m T \times R. \quad (3)$$

The estimate is denoted by $\hat{b}_m$, and its asymptotic variance is denoted by $\hat{\sigma}_m^2$. We then use the hierarchical model $\hat{b}_m | b_{,m} \sim N(b_m, \hat{\sigma}_m^2)$ and $b_m \sim N(b, \sigma^2)$. We then estimate b and its variance $\sigma^2$ by maximizing the likelihood after integrating $b_m$ out, that is,

$$L(b, \sigma^2) = \prod_{m=1}^{M} \int p(\widehat{b_m} | b_m, \hat{\sigma}_m^2) p(b_m | b, \sigma^2) db_m.$$

The calibrated estimator for the racial/ethnic difference attributable to the COVID-19 infection thus is $\hat{\tau} = \hat{\beta}_3 - \hat{b}$.

In the main manuscript within the Methods Section, we added the following text:

*"While we used propensity score matching to account for the measured confounders and difference-in-differences analyses to address pre-infection racial/ethnic differences, the results can still be impacted by unmeasured confounder bias when the parallel trends assumption does not hold. To mitigate such bias, we collected thirty-one negative control outcomes, as prespecified by pediatric physicians, that should not exhibit racial/ethnic differences due to COVID-19. We fitted the model from Equation (1) with the outcome replaced by each NCO. The residual bias thus can*

*be estimated from the average of $\beta_3$s from the models for the NCOs. By using these NCOs, the study was able to calibrate the residual bias from unmeasured and systematic source.*

***Figure 2c** illustrates this by the expanded DAG. When the parallel trends assumption is violated, a backdoor path from U to ΔY can emerge, potentially biasing our estimates. To address this limitation, we incorporate NCOs into our analysis. Similarly, we denote the NCO during the post-infection period by $W_1$ and the NCO during the pre-infection period by $W_0$. We first apply the DiD model using the NCO as the outcome. Due to the influence of $U$, we observe a path from U to $\Delta W$ (where $\Delta W = W_1 - W_0$). We then assume that the effect of U on $\Delta W$ is equivalent to its effect on $\Delta Y$. By leveraging this assumption, we can effectively block the path from U to $\Delta Y - \Delta W$ by subtracting the $\Delta W$ from $\Delta Y$. This allows us to estimate the racial/ethnic difference in PASC symptoms and conditions with greater accuracy and robustness to unmeasured confounding."*

7. **Several terms used in the paper appear inappropriate:**

a. The term 'potential PASC' has been used in multiple places. It is unclear what this implies. Does it reflect the authors' lack of confidence in determining PASC, or does it suggest that the outcome measured may not be associated with COVID-19? I suggest removing the word 'potential.'
b. In lines 148 and 150, what do 'moderate' or 'strong evidence' mean? Are these based on effect size, P values, or another measure? Please clarify the quantification of 'moderate' and 'strong.'
c. The authors suggest they aim to account for time-varying unmeasured confounders. However, in this study, there are no time-varying confounders given the fixed exposures (racial/ethnic group and COVID status defined in this study). Please review the definition of a time-varying confounder.
d. In line 159, why is an RR of 0.93 (0.85, 1.24) described as a 'minor decrease'?
e. In line 164, what does 'almost no decrease' mean given an RR of 0.99 (0.89, 1.11)?
f. In lines 172-173, why is an RR of 0.74 described as 'increased risk'?
g. Line 134: There should be no PASC for participants without COVID-19."

**Our response:** Thank you very much for pointing this out. We have the following point-to-point response below.

- "a. The term 'potential PASC' has been used in multiple places. It is unclear what this implies. Does it reflect the authors' lack of confidence in determining PASC, or does it suggest that the outcome measured may not be associated with COVID-19? I suggest removing the word 'potential.'"

  Thank you for pointing this out. We have removed the word "potential" to avoid further confusion in the revised manuscript.

- "b. In lines 148 and 150, what do 'moderate' or 'strong evidence' mean? Are these based on effect size, P values, or another measure? Please clarify the quantification of 'moderate' and 'strong.'"

  Thank you for your question. In lines 148 and 150, the terms "moderate" and " strong evidence" are based on the effect sizes observed in our analysis. Specifically, "moderate evidence" refers to an effect size that is statistically significant with a P value less than 0.05 but not large enough (>1.2) to be considered a strong effect. "Strong evidence" indicates a larger effect size that is also statistically significant, typically with a P value much lower than 0.05 (e.g. <0.001).

  To avoid the confusion, we have deleted the "strong/ moderate" term. This reads

  *"Overall, we found evidence of an increase in composite outcomes, i.e., at least one condition and any of the syndromic conditions, after SARS-CoV-2 infection among the AAPI group in both severe and non-severe COVID-19 group, but there was no evidence of increased racial differences among Hispanic and NHB groups."*

- "c. The authors suggest they aim to account for time-varying unmeasured confounders. However, in this study, there are no time-varying confounders given the fixed exposures (racial/ethnic group and COVID status defined in this study). Please review the definition of a time-varying confounder."

  Thanks for your suggestion. We reviewed the definition of the time-varying confounder. The time-varying confounder refers to the value of the confounding variables that changed over time [12], [13]. The values of the measured confounder won't be changed over time.

  We initially used the word "time-varying" to refer to the systematic bias from the confounding variables that can vary across time. That is, when the parallel trends assumption for the DiD does not hold, the bias from the unmeasured confounders for the pre-infection period is different from the post-infection period. As pointed out in Comment 6, we used the NCOs observed in both time periods to solve this problem.

  We removed the "time-varying" from the supplementary materials to avoid confusion and it reads

  *"To address potential systematic bias between the pre-infection and post-infection periods in our study, we developed a novel approach to difference-in-differences analysis, incorporating calibration using negative control outcomes."*

  *Reference:*
  *[12] Mansournia, M. A., Etminan, M., Danaei, G., Kaufman, J. S., & Collins, G. (2017). Handling time varying confounding in observational research. bmj, 359.*
  *[13] Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. Epidemiology, 11(5), 550-560.*

- “d. In line 159, why is an RR of 0.93 (0.85, 1.24) described as a 'minor decrease'?”

  Thank you for bringing this to our attention. We acknowledge that our original description of the relative risk (RR) of 0.93 (95% CI: 0.85, 1.24) as a 'minor decrease' was imprecise. This characterization was suggested by one of our patient representatives during our collaborative review process, reflecting our commitment to incorporating patient perspectives in our research communication. Upon further consideration, we agree that this description needs clarification. The RR of 0.93 indicates a smaller change (decrease) in racial/ethnic difference after the infection compared to before infection, and as the confidence interval crosses 1, this result is not statistically significant. We have revised this sentence in the manuscript to accurately reflect the statistical interpretation:

  *“NHB patients showed a non-statistical significant decrease in at least one condition (RR 0.93, 95% CI 0.85 to 1.24, P=0.147) as compared to NHW patients.”*

- “e. In line 164, what does 'almost no decrease' mean given an RR of 0.99 (0.89, 1.11)?”

  We appreciate your question regarding our description of the relative risk (RR) of 0.99 (95% CI: 0.89, 1.11) as "almost no decrease." This phrasing was initially suggested by one of our patient representatives as part of our effort to include patient perspectives in our research communication.

  However, we recognize that this description lacks statistical precision. Upon reflection, we agree that a more accurate representation of these results is necessary. We have revised this sentence in the manuscript to read:

  *“The relative risk of 0.99 (95% CI: 0.89, 1.11) indicates no statistically significant change in risk.”*

- “f. In lines 172-173, why is an RR of 0.74 described as 'increased risk'?”

  Thank you for this important question. We apologize for any confusion our original phrasing may have caused. The interpretation of the relative risk (RR) in our DiD analysis requires some clarification.

  In our DiD analysis, the RR of the interaction term between racial/ethnic group and time period is interpreted differently than in a standard risk analysis. There are two scenarios to consider:

  1. When the RR of the interaction term is greater than 1, it indicates a greater change in PASC symptoms and conditions for the minority group compared to the reference group (NHW) after COVID-19 infection, or an increase in racial/ethnic disparity.
  2. When the RR is less than 1, as in this case (RR = 0.74), it indicates a smaller change in PASC symptoms and conditions for the minority group compared to the reference group after COVID-19 infection, or a decrease in racial/ethnic disparity.

Therefore, the RR of 0.74 suggests that the racial/ethnic difference in the risk of developing this particular PASC symptom or condition decreased after COVID-19 infection, compared to the pre-infection period. We have revised the sentence in the manuscript to more accurately reflect this interpretation:

*"NHB patients had a smaller change in skin symptoms (RR 0.74, 95% CI 0.58 to 0.96, P=0.021) than NHW patients following COVID-19 infection, compared to the pre-infection period."*

- "g. Line 134: There should be no PASC for participants without COVID-19."

Thank you for this important question. We appreciate the opportunity to clarify our terminology and methodology.

In line 134, the phrase "PASC symptoms and conditions" refers to categories of symptoms and conditions that could plausibly be considered part of the Post-Acute Sequelae of SARS-CoV-2 (PASC) spectrum. For example, fever and chills are symptoms that could be related to PASC, but they are not exclusive to post-COVID-19 conditions. These symptoms and conditions can occur in individuals regardless of their COVID-19 infection status.

It's crucial to distinguish between these general symptoms and conditions and a clinical diagnosis of PASC. A PASC diagnosis, typically denoted by the ICD-10 code U09.9, is a specific clinical diagnosis that can only be made after a confirmed COVID-19 infection.

Our study examines the prevalence and incidence of these PASC-related symptoms and conditions both before and after COVID-19 infection to understand how the infection may have altered their occurrence or severity across different racial/ethnic groups.

To address this potential confusion, we have added a clear definition of "PASC symptoms and conditions" in the Results section of our revised manuscript. This addition should help readers distinguish between the general symptoms we're tracking and the specific clinical diagnosis of PASC. We thank you for highlighting this need for clarification, as it allows us to improve the precision and clarity of our manuscript.

The added manuscript reads
*"Our definition of PASC symptoms and conditions included 24 symptoms and conditions as shown in Rao et al.[11], including abdominal pain, abnormal liver enzyme, acute kidney injury, acute respiratory distress syndrome, arrhythmias, cardiovascular signs and symptoms, changes in taste and smell, chest pain, cognitive functions, fatigue and malaise, fever and chills, fluid and electrolyte, generalized pain, hair loss, headache, heart disease, mental health disorders, musculoskeletal pain, myocarditis, myositis, Postural Orthostatic Tachycardia Syndrome (POTS) or dysautonomia, respiratory signs and symptoms, skin symptoms, and thrombophlebitis and thromboembolism. We used validated diagnostic codes (ICD and SNOMED) confirmed by board-certified pediatricians, with details of the code sets available*

*in the Supplementary Materials **Table S2**. Systematic and syndromic conditions related to PASC were grouped by the 24 PASC symptoms and conditions, which are detailed in **Table S3**."*

8. **The discussion fails to contextualize the study results in terms of their clinical interpretation and significance. The three strengths highlighted by the authors in the discussion (using matching, balancing pre-infection differences, and employing negative control outcomes) do not appear substantial enough to be emphasized as the central elements of the discussion of a methodological study nor for a clinical research study.**

**Our response:** We appreciate your insightful feedback regarding the discussion section of our manuscript. We acknowledge that our initial discussion may not have enough methodological aspects for the clinical interpretation and significance and methodological innovation. In response to your valuable comments, we have substantially revised our discussion to address these concerns. Here's how we've improved the discussion:

- Enhanced Clinical Interpretation and Significance: We have significantly expanded the clinical interpretation of the results. This now includes a detailed discussion of the clinical implications of our findings, particularly focusing on the racial/ethnic differences observed in PASC symptoms and conditions. We talk about the reason for the racial/ethnic difference and possible related socioeconomic factors. We discuss how these findings can inform the health system, and the healthcare providers for targeted screening, follow-up protocols, and treatment strategies for different racial/ethnic groups.
- Public Health Significance: We've added a section discussing the broader public health implications of our study. We highlight how our comprehensive national sample, representing approximately 10% of US children, can inform public health strategies to reduce healthcare access disparities and enhance specialized pediatric care availability across all regions.
- Balancing Clinical Findings and Methodological Advancement: While we still discuss our methodological approach, we have reframed it to emphasize how it enhances the reliability of our clinical findings for racial/ethnic differences attributable to COVID-19 rather than presenting it as the central focus. We explain how our use of propensity score matching, difference-in-differences analysis, and negative control outcomes strengthens the validity of our clinical conclusions.
- Enrich the discussion of three strengths: using matching, balancing pre-infection differences, and employing negative control outcomes.

Our adjustment in the manuscript for the discussion part is

*"This study is among one of the first studies examining PASC racial/ethnic differences, providing valuable insights into the clinical relevance of these disparities. Understanding these differences is crucial for developing targeted interventions to ensure equitable healthcare access and outcomes for all pediatric populations affected by PASC.*

*The increased risk among minor racial/ethnic pediatric patients may necessitate targeted follow-up care and support for the minor racial/ethnic population. Healthcare providers should be aware*

*of the potential for varied PASC presentations across racial/ethnic groups in pediatric populations. This knowledge can inform more targeted screening and follow-up protocols, potentially improving early detection and management of PASC symptoms in different racial/ethnic patient populations.*

*The implications of this study on a broader public health level are significant. By utilizing data from 13 institutions, representing approximately 10% of US children and encompassing both urban and suburban health systems, this study offers a comprehensive national sample. Public health strategies must prioritize the reduction of healthcare access disparities and the enhancement of specialized pediatric care availability across all regions. This will help in better managing and treating PASC in children, ensuring that all affected patients receive the necessary care regardless of their geographic location or socio-economic status. Furthermore, these findings underscore the importance of continuous surveillance, research, and tailored public health policies to address the evolving needs of this patient group.*

*The method utilized in our study has multiple strengths. First, we used propensity score matching methods instead of linear regression models in our adjustment of the confounders, which helped us reduce the non-linear effects of the confounders. Second, we accounted for the pre-infection racial/ethnic difference in long-term COVID-19-related symptoms and conditions. This approach enabled us to quantify racial/ethnic differences attributable to COVID-19 more accurately. The DiD model is particularly powerful in the context of studying racial/ethnic disparities, as it helps to isolate the effect of COVID-19 from pre-existing health disparities. By comparing the change in health outcomes before and after COVID-19 across racial/ethnic groups, we can more confidently attribute observed differences to the impact of the virus rather than to the pre-infection racial/ethnic differences.*

*A key strength of our study is the implementation of NCOs, which enhances the reliability and interpretability of our findings. We used the NCOs which are outcomes not expected to be affected by COVID-19 or show racial/ethnic differences due to COVID-19. By analyzing these alongside our primary outcomes, we enable the detection of systematic bias from unmeasured confounding variables. This approach allows us to calibrate results from the DiD model. Importantly, NCOs help mitigate the impact of unmeasured confounding that may change before and after the infection period. This is particularly crucial when studying racial/ethnic disparities, where many dynamic social and environmental factors may not be captured in our dataset."*

9. **The study requires participants, both infected and uninfected, to have at least one healthcare visit within 28-179 days after the index date. I am concerned that this may result in selection bias, as the majority of the study population may not require such frequent healthcare visits, especially those who were included during 2020 and 2021 and uninfected, due to hospital policies.**
   **Similarly, the study selected individuals who tested negative as the control group. Given that tests were not commonly conducted within the study population, does this suggest that the study is selecting a control group that is likely to have symptoms, or have an appointment for surgery or a procedure?**

**Our response**: Thank you for your insightful comments regarding potential selection bias in our study design. We appreciate the opportunity to address these important concerns. You are correct that requiring at least one healthcare visit within 28-179 days after the index date could potentially introduce selection bias. We acknowledge this limitation and have addressed it in our revised discussion section. However, this criterion was implemented to balance against the risk of "loss to follow-up" bias [14], which could significantly impact our results if we were unable to capture important health outcomes in participants who did not return for care.

In the main analysis of the DiD study, we only used the COVID-19-positive patients. However, we did use the COVID-19-negative patients in the crude analysis to compare the incidence of the PASC symptoms and conditions. We acknowledge that by using the test status to define the COVID-19-negative patients and to include the sick patients. We have incorporated various testing data into our studies to reduce the impact of miscapturing of COVID-19 negative status. Specifically, we have utilized PCR, antigen, and serology test results. We also want to let you know that the study captures the asymptomatic patients, as illustrated by the acute phase (the time period from seven days before to fourteen days after the negative test-related visit).

As mentioned above, the bias is toward the null for the incidence. However, we still show a greater incidence of PASC symptoms and conditions for the COVID-19-positive patients when compared to the COVID-19-negative patients.

We added to the discussion part, which reads:

"*Fifthly, our study design, requiring at least one healthcare visit within 28-179 days post-index date, balanced selection bias against 'loss to follow-up' bias [14]. While this approach may over-represent frequent healthcare users, especially during 2020-2021, we mitigated this by incorporating diverse testing data to reduce COVID-19 status misclassification. Furthermore, in the sensitivity analysis (**Section S4**), the resulting selection bias likely included sicker patients in the COVID-19 negative group, potentially biasing our results towards the null. Nevertheless, we still observed a higher incidence of PASC symptoms and conditions for COVID-19-positive patients compared to COVID-19-negative patients. Future studies could explore alternative designs or data sources to enhance the external validity of these findings."*

*Reference:*

*[14] Howe, C. J., Cole, S. R., Lau, B., Napravnik, S., & Eron Jr, J. J. (2016). Selection bias due to loss to follow up in cohort studies. Epidemiology (Cambridge, Mass.), 27(1), 91.*

**REVIEWER COMMENTS**

Reviewer #2 (Remarks to the Author):

Thank you very much for the hard work you put into the revision of this manuscript. All of my comments have been addressed, and I have no further comments.

**Our response**: Thank you for your positive feedback and for acknowledging our efforts in revising the manuscript. We appreciate your thorough review.

Reviewer #4 (Remarks to the Author):

This is a review of the first revised version of the manuscript titled "Racial/Ethnic Differences in Long-COVID-Associated Symptoms among Pediatrics Population: Findings from Difference-in-differences Analyses in RECOVER Program".

This revised version has hugely improved compared to the first submitted version of the manuscript and the efforts of the authors to respond to reviewer's comments are commendable.

**Our response**: Thank you for your positive feedback. We appreciate your detailed review and below are our point-by-point responses to your additional minor comments.

The authors have addressed my comments thoroughly and appropriately. A few minor points:
- Line 157: This seems to be the first mention of “NCO”. Make sure all acronyms are spelt out at first mention.
**Our response**: Thank you for pointing this out. We will ensure that "NCO" is spelled out as "Negative Control Outcome" at its first mention in the manuscript for clarity. The revised manuscript now reads in Line 161,
“*To adjust for pre-infection racial/ethnic differences, we applied a DiD model with negative control outcome (NCO) calibration to the matched cohort.*”

- Line 306: replace “the other minor groups” by “the other minority groups”. Double-check if other mentions of “minor group” appear in the text and replace by “minority group” which is a better accepted terminology.
**Our response**: Thank you for your suggestion. We have replaced “the other minor groups” with “the other minority groups” as recommended. We also thoroughly reviewed the manuscript to identify and update any other instances of “minor group” to “minority group” for consistency.

- Line 431: suggest using “who had at least one visit from 7 days before up to 18 months after the index date” rather than the other way around.
**Our response**: Thank you for your suggestion. To clarify, our study actually uses the criteria of “at least one visit from 18 months before to 7 days before the index date” rather than the proposed timeframe. This approach aligns with our study design and data requirements. To make it clearer, we rephrased it in Line 438 to
*“We included patients under the age of 21 who had at least one visit from 18 months before to 7 days before the index date (defined as the baseline period)”*

Reviewer #5 (Remarks to the Author):

I thank the authors for their efforts in revising the manuscript and for providing a detailed response. The manuscript has shown improvement through the clarification and correction of the previously highlighted errors. However, I still have a few remaining concerns.

1. The authors have now clarified that only participants with positive COVID-19 were used in their analyses, without comparison to those without COVID-19. This raises concerns about whether the effects measured are truly reflective of PASC. It is important to note that the references offered by the authors in their response used both those with and without COVID-19 to quantify PASC.
The authors could claim they are measuring the difference in PASC or long COVID between races without a COVID-19 negative control group only if: 1. They can demonstrate that these conditions did not occur within the COVID-19 negative group, or 2. They can show there is no difference in these outcomes across races within the COVID-19 negative group.
Otherwise, they are only measuring differences in diagnosis or healthcare utilization after COVID-19, which does not necessarily indicate PASC. While this remains an important research question, there is a key distinction compared to measuring PASC.
**Our response**: We thank you for your suggestions. To provide validation for PASC symptoms and conditions as a true reflection of PASC, we agreed to add more analyses on the COVID-19 negative cohort. To show this, we added a sensitivity analysis, which reran the main analysis, i.e., difference-in-difference (DiD) analysis with negative control outcome (NCO) on the COVID-19 negative cohort, as suggested here.

**Figure S14** presents the risk ratios (RRs) for racial/ethnic differences in PASC symptoms and conditions within the COVID-19-negative cohort, showing no significant disparities across any PASC symptoms or conditions. These findings indicate that any racial or ethnic differences observed in PASC symptoms among COVID-19-positive individuals are unlikely to stem from inherent differences in symptom incidence patterns. Instead, they may reflect factors specifically associated with COVID-19 infection or its aftereffects. This strengthens our confidence that the racial and ethnic disparities in PASC seen in the COVID-19-positive cohort are likely attributable to COVID-19 exposure itself, rather than to underlying population characteristics or unrelated health trends.

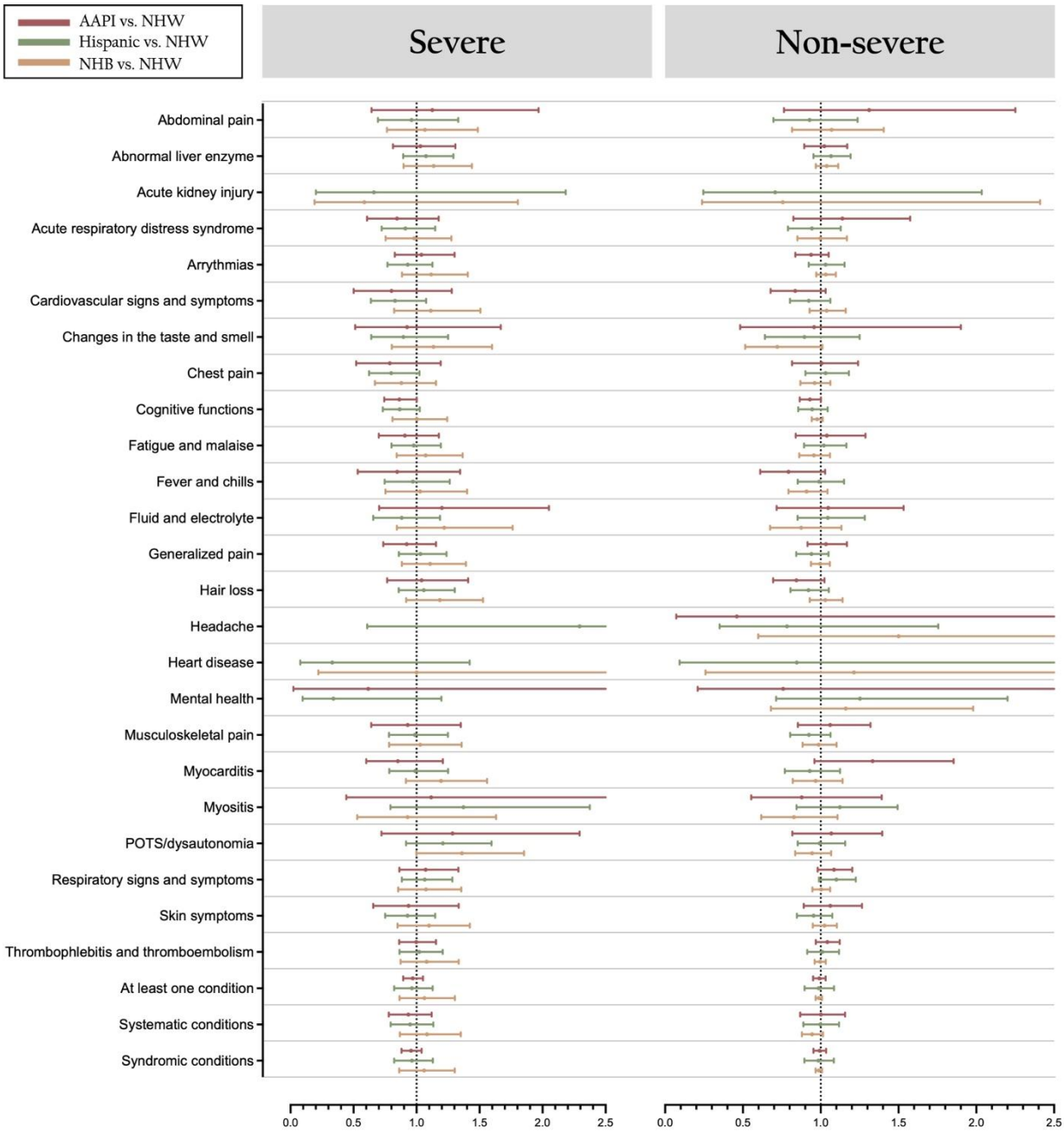



**Figure S14**. The racial/ethnic difference in risk ratio (RR) across PASC symptoms and conditions for COVID-19 negative cohort.

In the manuscript, we added the following text in the main manuscript in Line 264, which reads as

"*To further validate our definition of PASC symptoms and conditions, as well as to investigate potential racial and ethnic differences in PASC symptoms attributable to COVID-19, we applied the DiD approach with NCO calibration to the COVID-19-negative cohort (**Figure S14**). This analysis revealed no racial or ethnic differences in PASC symptoms and conditions attributable to COVID-19 within the COVID-19-negative cohort. These findings suggest that the observed*

*racial/ethnic differences in PASC symptoms and conditions among COVID-19-positive individuals are unlikely to be due to inherent differences in symptom and condition incidence patterns but rather may reflect factors specifically associated with COVID-19 infection or its sequelae. This analysis validates that the racial/ethnic differences in PASC observed in the COVID-19-positive cohort are related to COVID-19 infection, rather than underlying population differences or unrelated health trends.*"

2. I thank the authors for their clarification regarding their consideration and utilization of the DID approach. In addition to this, I suggest the authors clearly specify the two "differences" involved in the DID for this study early in the manuscript. Furthermore, while the authors state that DID can remove confounding effects from unmeasured confounders, could the authors please comment on the following scenario: if COVID-19 modifies the effect of an unmeasured confounder on the outcome (similar to how it modifies the effect of race on the outcome), would DID introduce bias in the presence of such unmeasured confounders?

**Our response**: Thank you for your suggestion. To clearly specify the two "differences" involved in the DID for this study early in the manuscript, we moved the illustration of **Figure 2a** (now **Figure 1a**) for the two differences in the last part of the introduction. Now it reads in Line 123,

*"Specifically, **Figure 1a** demonstrates how to use the DiD model to find the racial/ethnic differences attributable to COVID-19. We measured the pre-infection racial/ethnic differences by comparing PASC symptoms and conditions between minority racial/ethnic groups and the NHW group during the pre-infection period. We then measured post-infection differences using the same metrics. The racial/ethnic differences attributable to COVID-19 were determined by calculating the difference between these post-infection and pre-infection differences. When COVID-19 may modify the effect of unmeasured confounding variables, we applied a novel negative control method (**Figure 1c**) to adjust for the shifted impacts of these potential unmeasured confounders. By addressing these questions, our objective is to shed light on racial/ethnic differences in the relationship between COVID-19 and PASC symptoms and conditions."*

We agree that using the DiD method alone could produce biased results if COVID-19 modifies the effects of unmeasured confounding variables. However, in the manuscript, we incorporated an additional analysis step, using a negative control outcome (NCO) to empirically calibrate the bias from the shifted effects of these unmeasured confounders. Consequently, the DiD analysis with NCO calibration provides accurate results when COVID-19 affects unmeasured confounding variables.

**Figure 1c** illustrates the rationale for DiD with NCO calibration. In the left panel of **Figure 1c**, COVID-19 alters the effects of unmeasured confounding variables ($U$), causing a shift in the causal pathway from $U$ to the outcome before COVID-19 ($Y_0$), $U \rightarrow Y_0$, compared to the pathway after COVID-19 ($Y_1$), $U \rightarrow Y_1$. As shown in the middle panel of **Figure 1c**, this change creates a

potential backdoor path from $U$ to $\Delta Y$, introducing bias into estimates derived from the DiD method alone.

To address this bias, we incorporate NCOs into the analysis. We denote each NCO during the post-infection period as $W_1$ and NCO during the pre-infection period as $W_0$. Due to the influence of $U$ changed by COVID-19, there is a corresponding pathway from $U$ to $\Delta W$ (where $\Delta W = W_1 - W_0$). We assume that the effect of $U$ on $\Delta W$ is equivalent to its effect on $\Delta Y$. In the right panel of **Figure 1c**, we demonstrate that subtracting $\Delta W$ from $\Delta Y$ effectively blocks the path from $U$ to $\Delta Y$ under such assumption. This approach enables us to estimate racial/ethnic differences in PASC symptoms and conditions with greater accuracy and robustness, even when COVID-19 modifies unmeasured confounding factors.

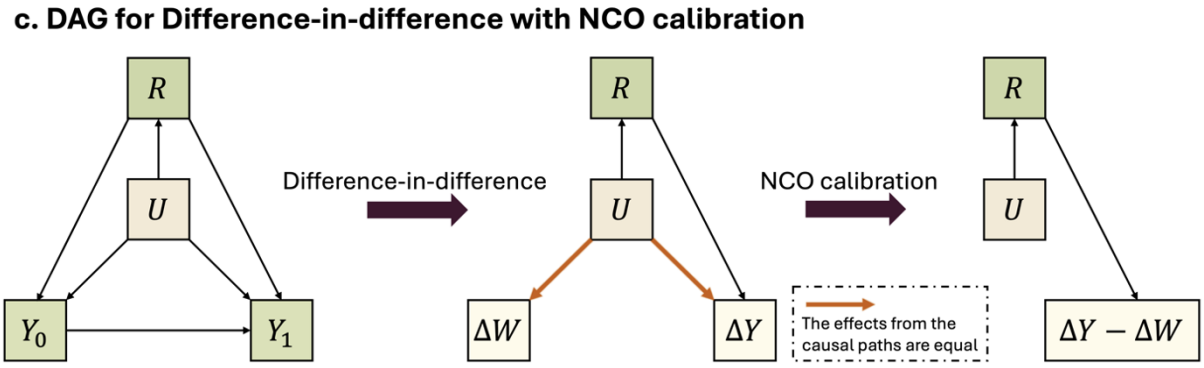



**Figure 1c**. The directed acyclic graph for the difference-in-difference (DiD) with negative control outcome (NCO) calibration.

We have updated the manuscript in the Introduction, Results, and Methods sections to clarify these points.

In the Introduction section, we added the following text in Line 128,
*"When COVID-19 may modify the effect of unmeasured confounding variables, we applied a novel negative control method (**Figure 1c**) to adjust for the shifted impacts of these potential unmeasured confounders"*

In the Results section, we added the following text in Line 170,
*"However, when the parallel trends assumption is violated—specifically if COVID-19 modifies the effect of unmeasured confounding variables—the outcome estimates from the DiD may become biased without the NCO calibration."*

In the Methods section, we wrote in Line 560

"***Figure 1c*** *illustrates this by the expanded DAG. When the parallel trends assumption is violated, a backdoor path from U to ΔY can emerge, potentially biasing our estimates. To address this limitation, we incorporate NCOs into our analysis. Similarly, we denote each NCO during the post-infection period by* $W_1$ *and during the pre-infection period by* $W_0$*. Due to the influence of* $U$*, we observe a path from U to ΔW (where* $\Delta W = W_1 - W_0$*). We then assume that the effect of U on ΔW is equivalent to its effect on ΔY. By leveraging this assumption, we can effectively block the path from U to* $\Delta Y - \Delta W$ *by subtracting ΔW from ΔY. This allows us to estimate the racial/ethnic difference in PASC symptoms and conditions with greater accuracy and robustness to unmeasured confounding."*

3. I also suggest the authors proofread and revise the manuscript throughout. For example, the latter portion of the revised sentence, "Hispanic patients showed almost no increase in at least one condition (RR 1.01, 95% CI 0.98 to 1.04, P=0.498), and the relative risk of 0.99 (95% CI: 0.89, 1.11) indicates no statistically significant change in risk," is unclear and needs revision.

**Our response**: We thank you for highlighting the unclear sentence. We have revised it to:

"*Hispanic patients showed a non-significant increase in at least one condition (RR 1.01, 95% CI 0.98 to 1.04, P = 0.498), while NHB patients had a similar non-significant risk (RR 0.99, 95% CI 0.89 to 1.11, P = 0.915) in at least one condition.*"

Additionally, we have conducted further proofreading throughout the manuscript, including the following revisions:

1. In Line 187, we revised the text to: "*Figure 3 shows the racial/ethnic differences attributable to COVID-19 in PASC symptoms and conditions, stratified by the severity of acute COVID-19 after NCO calibration."*
2. In Line 199, we modified the sentence to: "*Both Hispanic patients (RR 0.99, 95% CI 0.91 to 1.08, P = 0.804) and NHB patients (RR 0.93, 95% CI 0.85 to 1.24, P = 0.147) showed a non-significant decrease in at least one condition compared to NHW patients*."