Review of: "COVID-19 Vaccine Effectiveness Against Long-COVID-19 Condition in Pakistan"

Omid Keshavarzian[1]

1 Transplant Research Center, Shiraz University of Medical Sciences, Iran, Islamic Republic of

Potential competing interests: No potential competing interests to declare.

This study has been performed on an important topic. However, in my opinion, there is a way to improve it. Please consider the following comments.

All the best,

1. Title can be revised to “COVID-19 Vaccine Effectiveness Against Long-COVID-19 diagnosis: A Retrospective Cohort from Pakistan”.
2. In the abstract:
    a. What is/are the long COVID outcomes? Please mention.
    b. The objective should be corrected according to the work you did.
    c. Please mention point estimates with precisions.
    d. Revise the methods and results according to the following comments.
3. In the introduction:
    a. Please update this sentence until Jan 2024: “According to the World Health Organization (WHO), there were 599 million confirmed cases of COVID-19 as of August 31, 2022[1].”.
    b. Please update the second paragraph to the latest definition of Long COVID [https://doi.org/10.1056/NEJMsb2408466].
    c. Use a uniform term instead of “post-acute sequelae of COVID-19” or “post-COVID-19 syndrome”.
    d. The last sentence of this section (… diagnosed with long-COVID who underwent COVID-19 vaccination.) needs corrections. You’ve also included unvaccinated subjects.
4. In the inclusion criteria section, you have mentioned that patients were included if they were alive one year before the study endpoint (i.e., February 1, 2023), and had at least one documented healthcare visit for both before and after January 1, 2023. I cannot understand these criteria according to the study purposes. Please make these clearer.
5. For the Long-COVID Identification section, considering the latest definition of “Long COVID”, at least an aggregate (continuous or not) of 12 weeks of having symptom(s) after recovery:
    a. What was the rationale to choose a minimum of a 20-week interval between COVID-19 diagnosis to 1st February 2024? Did you, for example, wrongly exclude a subject who recovered after a week from COVID-19, developed

Long COVID a week later that lasted for 12 weeks continuously? I think you should revise it.

b. “Patients who died within twelve weeks of this index date were excluded.” What was your rationale?

c. “Long-COVID cases were identified if patients reported one or more COVID-associated symptoms between 12 and 20 weeks after the initial diagnosis.”. Please add a reference for this definition, although it is not the latest definition suggested so far. Additionally, did you mean 12-20 weeks after the initial diagnosis of COVID-19 or recovery from the COVID-19? Please critically address this.

6. For the Patient Demographics and Health Status section, were vaccination data, medical conditions, and symptoms (used for Long COVID diagnosis) assessed through self-declarations? If so, this is a considerable source of bias.
7. Who diagnosed Long COVID? Was it assessed from the health records, retrospectively, or was a physician recorded the diagnosis, independently?
8. Since the primary aim of your study is COVID-19 VE against the occurrence of Long COVID, it would be majorly valuable to enhance vaccination data.

    a. Although you have defined a variable as “vaccination status” – with no, partial, and complete responses – it would be better to revise it to: no, first dose, second dose/complete, third dose/booster.

    b. Also, you’ve stated that “data on vaccinations, including the timing of the first dose, was also considered.”. How did you handle subjects who were vaccinated after COVID-19 infection?

9. Although you’ve reported VE against the "Symptom Count" using GLM, I think it would be very interesting to assess VE against: each Long COVID symptom (provide tables as supplement), No/Yes Long COVID, severity of Long COVID, and aggregated duration of Long COVID.
10. Please define the study type. It appears to be a retrospective cohort (real-world data) study. Therefore, in addition to the fact that you are assessing VE, I strongly encourage the authors to estimate relative risks (RRs) instead of ORs. Then, VE can be simply calculated as a percentage (VE [%] = 1-RR). It should also be employed to the 95% CIs. For this purpose, you can apply Log–binomial regression to your data.
11. How did you discriminate a Long COVID symptom from a pre-existing symptom or a symptom originating from other sources?
12. Several other covariables can be incorporated in the multivariable models, for example, WHO severity scale, definition of predominant variant, ICU admission, etc.
13. In the manuscript, “long-COVID-19 outcomes” is used several times. Please clarify what outcome(s) is exactly mentioned.
14. A remarkable portion of the results section is dedicated to the descriptive reports. It would be better to report the VE-related results more.
15. The Long COVID profile of the cohort is missed in the results section. You should report, for example, frequencies of symptoms, type of syndrome, duration, recovery-to-Long COVID time interval, severity, etc.
16. “Pfizer/BioNTech was the most common vaccine type (17.5%), followed by AstraZeneca (23.9%).” This sentence is wrong.

17. The third paragraph of the results section, as well as the corresponding table 2, cannot be easily understood. In the first model, you've used a logistic regression model with a Newton Conjugate Gradient solution to identify factors influencing the persistence or onset of long-COVID symptoms. I suppose that the binary outcome in the model is the vaccination status in Long COVID patients (Yes/No vaccination within 12 weeks after diagnosis) and the main explanatory factor is the vaccination timing. In this respect:

    a. How can the vaccination status and vaccination timing be used as the major blocks of the model? What is your rationale? Please clearly define the outcome (dependent variable). Persistent/not persistent? Days from recovery to Long COVID diagnosis? Which?
    b. You've mentioned that "Vaccination timing was analyzed as a ratio representing the timing of the first vaccination dose relative to the infection.". Based on the study aims for VE against Long COVID, vaccination precedes the event (COVID-19 infection or Long COVID). You should critically revise it.
    c. As the model is fitted using the other independent variables, it would be better to report the statistics for them.
    d. "… Within 0-4 weeks post-diagnosis, the odds ratio was 2.2X (95% CI: 0.390-2.445, $p < 0.05$), indicating an increased likelihood of vaccination …". Nothing can be understood.
    e. Moreover, as previously mentioned, it would be better to report RR and VE%. For more information, you can use this publication [https://doi.org/10.1503/cmaj.101715].
    f. Totally, this section is meaningless.

18. Additionally, for the last paragraph of the results section, there are some misleading points:

    a. This is a GLM model with "Symptom Count" as the response variable, which is predicted with various independent factors, including vaccine-related variables. In table 3 title, what is the meaning of "… who were vaccinated before diagnosis, within 20 weeks after diagnosis, or remained unvaccinated 20 weeks after diagnosis."?
    b. How about the reason for the significant increase for both male and female genders?
    c. How can you define the findings showing increased Symptom Count for all of the vaccination statuses or types?
    d. I think it would be better to use a binary outcome to assess the VE.

19. Please uniformly mention the study aim in different parts of the article. For example, you've stated "Our study aimed to investigate the effectiveness of COVID-19 vaccination in altering the presence of long COVID-19 conditions among vaccinated individuals." in the discussion section, different from the abstract, introduction, and results.
20. The fourth paragraph of the discussion should be clarified.
21. Please remember to revise the discussion according to the possibly different results after the corresponding comments.