Review of: "Immune responses to SARS-CoV-2 infection and vaccine in a big Italian COVID-19 Hospital: an 18-month follow-up"

Ruy Ribeiro[1]

1 Los Alamos National Laboratory

Potential competing interests: No potential competing interests to declare.

The study by Sansone et al analyzes the decay of anti-S antibodies is a large cohort of vaccinated hospital workers followed over up to 9 months pre-vaccine and another 9 months post-vaccine. The study is interesting because its size allows to test statistically several hypotheses about the decay of antibody. Moreover, the longitudinal nature of the study allows better understanding of the decay of antibodies over time. It is also interesting that a sizeable number of people in the study had pre-vaccine infection adding another covariate to the study.

Overall, the study is well-written and what was done is clearly described. In my view there are a couple of main issues that should be further elucidated to improve the quality of the manuscript and the robustness of the findings.

In general, the analyses of antibody titers shown should be done after transforming the antibody levels with the logarithm. This transformation is appropriate to normalize the distribution of the antibody levels. This may be especially important for the decay analyses, because it has been shown that the decay of antibodies is exponential (e.g., Cohen et al Cell Reports Medicine 2, 100354), so the logarithm would transform into a linear decay. Note that figure 2 shows that the decay is not linear and it may be compatible with an exponential decay (how would this figure look if the y-axis is logarithmic?). This transformation would also likely make the parametric tests more appropriate. (Not using the logarithm is main reason for my 2-star rating, which would increase to 4 otherwise.)

Figure 3 and following, the analyses done are not clear. Did you analyze separately those people with pre-vaccine infection adjusting for age and sex, and then analyzed the people with no pre-vaccine infection adjusting from age and sex? (And corresponding changes for the other figures?) If so, rather than calculating separate regression, here it would make sense to simply use the fully adjusted linear mixed effects model and then show the curves from that model – i.e., only the figure would be stratified by pre-vaccine infection, etc, but not the analyses. In fact, then all the figures (fig 3 to 8) would come from the same mixed effect regression, including interactions as appropriate. Please clarify this issue.

Perhaps it would make sense to include in your mixed effect model the baseline level of antibody. It would be interesting to see if the decay rate depends on the initial levels, which maybe would help explain some of the results in figures 3 to 8.

The anti-N results are especially puzzling. Do you have any information on the severity of infection, and could older people have had more symptomatic infection that would explain higher anti-N levels? Also could seroreversion be more

common in younger people because they had lower levels to start with? (Again here, it may be possible to analyze seroreversion controlling for baseline levels.)

Near the end of the discussion, you state “Such figure and the constant paucisymptomatic clinical course of all such cases allowed us to estimate a protection from SARS-CoV-2 infection above 95% and a 100% protection from hospitalization during the first nine months following the 1st jab administration”, but it is not clear how you calculate vaccine efficacy, because you don’t have a control group of unvaccinated people. Are you comparing with pre-vaccine? That would not be a proper vaccine efficacy, because of the evaluation of the control and vaccinated group at two different times, when the epidemic is in a different state.

There are a few other smaller issues that you could clarify:

It is not very clear why only the 7 individuals without measurable Ab were analyzed for T cell responses. A short clarification would be helpful.

Abstract: it is not clear what the “higher anti-S titers” shown are relative to.

Section 2 summary (and also first line of Discussion): Stating “this is the first study” is very risky, especially in COVID-19 since it is virtually impossible to know all that has been published to date.

Introduction “compulsory for HCW” where? Add Italy

Introduction: “Before and besides the vaccine,”… this is not clear. Do you mean “Before and after the vaccine…”?

Methods: “whenever available, the results of serological screening performed during 2020 (T-2 and T-1) were cumulated” This is unclear as T1 and T2 were in 2021 (according with figure 1). Also, it would make sense to integrate/consider/bring together (“cumulated” is not an appropriate word) the results for the earlier screening in 2020, but it is not clear how you would add the results from T1 and T2 (in 2021). Maybe T-2 and T-1, mean “T minus 2” and “T minus 1”, which would make sense, but I think these are not defined in the figure or anywhere else. Please clarify.

Results: perhaps better than “on May 6th, 2021, 8648 workers…” would be “by May 6th, 2021, 8648 workers…”

Results say 6862 but figure 1 says 6826.

Table 1: what does “according to standardized residuals” mean in this context? Isn’t significance evaluated by p-value?

Page 9: “Significantly higher anti-S titers were observed (…) (median antibody levels of 5,000 U/ml”. Since 5000 is the upper limit of quantification, it is important to clarify here that the titers were >5000 or to say that this is the upper limit.

Table 2 reports the medians and first and third quartile, but ANOVA is a parametric test for means, it would perhaps be better to show means and standard deviation. (See above about using logarithm of antibody levels and also given the large n, parametric tests are appropriate.)

Page 12, line after figure 4, what does “always on sample” mean in this sentence?

Page 13, four lines after figure 5, “The subjects who got a pre-vaccine infection” is not appropriate, it would be better “The subjects who had a pre-vaccine infection”.

Page 15, two lines after figure 7, “who got a pre-vaccine infection later in time. ” (… who had a pre-vaccine…). What does later in time mean here? When and what likely variant?

Page 18, “Such findings, also considering the level of in vitro neutralizing activity of anti-S (as low as 15 U/mL)”, it is not clear what you intend to say by referring to the neutralizing activity here. Please clarify.

What were the sensitivity and specificity of the assays?