

Distinct T cell correlates of COVID-19 risk before and after a second booster in previously SARS-CoV-2 infected vs. naïve COVAIL trial participants

Corresponding Author: Professor Peter Gilbert

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)

In this study, Bo Zhang and colleagues, investigated correlates of T cell protection by stratifying their COVAIL trial participants according to their infection status and prior booster vaccine regimen. Although the authors show important findings related to their research objective, they expanded their analysis with immunogenicity against 3 different vaccines. This makes the manuscript more of a mixture of two different objectives, although somewhat related, and does not contribute to the readability.

Part of the immunogenicity data have been published before (e.g. CD4/CD8 ratio), and the main results are not clearly supported by the figures (Fig1 and Fig2). The part of the manuscript that contributes to the main objective to assess the correlates, the authors write that they have an improved functionality assessment of the antigen-specific T cells compared to other studies, this is not sufficiently introduced and also not used to its full potential. For example, the authors say to fill the "gap" with T-cell "markers", while they actually look at antigen-specific T cells, using a subset of cytokines + CD154, where only a few represent the majority of the antigen-specific response. The authors do not make use of the other potential markers, or markers on those antigen-specific T cells. In addition, the shown correlations are observational and not mechanistic, providing no explanation about these differences. For example, there is no hypothesis on why polyfunctional CD4 T cell responses were found in higher numbers in the non-cases. Where these cells differ in terms of TCRs compared to the non-polyfunctional cells? Since stimulated with a pool of peptides, could the polyfunctionality truly arise from conserved epitopes? Or are these polyfunctional responses a result of a mutation-driven response, something that could be identified partly with TCR seq.

Major;

The title does not fully represent the results of the research findings as it assumes that these are correlates found at baseline (before actual infection and/or vaccination) but baseline is actually after the primary vaccination series. A suggestion for a more representative title could be; Distinct T cell correlates for the protection against COVID19 after booster for pre-exposed and naïve SARS-CoV-2 individuals.

Adhering to the maximum word count will enhance the readability and reviewing of the manuscript.

Introducing more structure of the introduction will be improved readability. The introduction in its current state reads as the results section of the COVAIL trials that were previously published (ref 10 and ref 11) and can be shortened, and perhaps work from others in the same field can be referenced to widen the angle.

Naïve individuals are self-test negative or negative for anti-N. How well does the anti-N predict pre-exposure, as it wanes over time? Instead of focusing on the distribution of the CD4 and CD8 T cell response, it might be more informative to elucidate more on the naïve vs. exposed individuals and their response to N and perform analysis on that. This would strengthen the claim that these individuals are more likely naïve and did not have an asymptomatic infection in the past which was missed. In line with this it is vital to understand and know how many are considered naïve based on self-tests and how many are actually anti-N negative? Naïve could be a group comprised of asymptomatics that have never had a reason to perform a self-test. This result should be included in the first paragraph of the paper.

Since the data rely on the antigen-specific T cell data which are generated via their ICS panel, it is essential to also show a gating strategy and include a representative example of the gating. It is also not clear which control has been used in these assays, it seems the authors have used medium control as negative control, where they should have used a DMSO control with equimolar DMSO% compared to the peptide mixes. Please clarify.

The results section is a mix of methods, figure legends without guiding the reader and actual results. Please structure.

Methods; the description of the ICS assay is not clear, no peptide pool supplier is listed, nor the used concentration, also no stimulation duration was listed. Was there no DMSO control as negative control used? (see comment above). Since the results of the manuscript are based on this assay, it is expected to show more details.

Results from Figure 1; the conclusion from the authors is that CD4 T cells are more dominant at D1 compared to CD8 T cells both in naïve and non-naïve. The figures are plotted in such a way that they are difficult to interpret. A suggestion would be to split the data on showing only IFN γ +, since the CD4 T cells from peripheral blood secrete IL2 whereas CD8 T cells do not. As the gating strategy is missing, it is difficult to assess whether this is not simply comparing different entities: CD4 T cells producing IFN γ and IL2) versus CD8 T cells solely producing IFN γ . Nevertheless, previous research indeed already shows that the CD4 T cell pool dominates over the CD8 T cell pool.

Also, most of the anti-N responses were non-responders both in the naïve pool as well as in the non-naïve pool, because the data are now split across all arms, the authors can also show a certain % of responder from the exposed (non-naïve) versus % responders from naïve. Then it becomes more evident that indeed the naïves are truly naïve.

The stages are not explained well, and therefore the results can be difficult to interpret. For example, lines 301-302 the authors write that the response is lower in stage 3 compared to stages 1 and 2. It would be more informative to refer to their primary one-dose vaccine regimen. Also, it is not easy for the reader to make the comparisons the authors show. For example, an overlaid dot plot can be generated for each "stage" whereby the dots are colored according to their D15 booster vaccine. Then, higher/lower can be compared much easier.

Lines 330-344; Suggestion to replace marker for upregulation of CD154 and production of xxxx (markers are not "produced", cytokines/ILs are). Also, this section would benefit from a clearer introduction. The reader cannot assess what the functionality score actually is, how many cytokines (or markers) are considered? The terminology remains unclear. As an example, what is the proportion of CD154 and the production of these cytokines for CD4 and CD8 responders?

Lines 346-383; This part describes that, when stage 1 and stage 2 results are combined, CD4 and CD8 responses are higher at D1 for those that have previously been exposed. Then, after a booster, CD4 levels equalize between those naïve and exposed while this is not the case for CD8 T cells. Suggestion to reduce this section if this is the conclusion the authors want to make. Also, we suggest to use a figure instead of a 2-page long table (lines 218 till 345).

Lines 385-394; this is known and can be removed. If the authors do want to integrate it, we suggest to refer to it in one sentence during the ICS results highlighting, Spike BA peptides correlated with the Index, and D1 correlated with D15 Frequencies, whereas Spike responses did not correlate with N.

It is not entirely made clear that lines 435 onwards show results of the risk of COVID19 after the booster regimen. It would be informative to add a simplified figure showing numbers of how many cases and non-cases there were after the booster. Also suggest to reduce the length of the text.

The word "marker" is used throughout the text, since the authors describe a 27 marker panel in the introduction, focusing on T cells, it is expected that certain markers on the antigen-specific T cells drive this "correlation", please rephrase since this is now misleading, or introduce the panel as a result showing marker expression of these antigen-specific T cells and compared across stages/ groups using for example UMAPs. Markers are mostly just IFN γ and/or IL2, and it is unclear whether the authors measured these on those antigen-specific T cells. The authors only state that a panel adapted from Cohen is used. Please add a table at line 862.

The authors find an inverse correlation for the exposed individuals for the HR of COVID19 within the first 90 days of booster. Have the authors looked into the antigen-specific T cells of those subjects and found an interesting phenotype or poly/oligo-functional properties of these cells from those subjects? Please clarify.

Lines 472-474; please clarify where can the results of those 5 functionality markers are described. S15 does not show this separately, but as a sum.

Line 511; the amount of markers, and which markers are changed constantly, without specifying the specific markers. The authors describe: defined by the expression of 9 markers, but the figure shows 5. Please clarify.

Lines 537-552; it is unclear whether the authors performed only clustering on the antigen-specific cells (by using which parameters/marker set?). It appears that this is not per se the case, if not, it is expected that the same populations with expressing cytokines are shown. Did the authors also compare the expression of PD1, CXCR3, CCR6 and other markers within those clusters between the groups? The latter would provide additional information and insights into the results of the inverse correlation. A suggestion is to perform a UMAP and measure the density level whether large shifts occurred.

Minor:

Please check spelling consistency (e.g. T-cell or T cell)

Please review the terminology regarding specific responses throughout the paper. For example, Fig1 and Fig 2 the authors say that these figures resemble "markers", while factually looking at the percentage of SARS-COV-2-specific CD4+ or CD8+ T-cells by looking at upregulation of IFN γ and/or IL2 after in vitro stimulation. Please adjust accordingly to improve clarity

(Lines-242-247 for example, but also line 292 “marker” levels)

Line 292, marker levels, is unclear what the authors mean. This could be any of the markers in the ICS panel, while the authors probably mean (only) IFN γ or IL2

What do the authors mean with N “index” ? Please clarify

Naïve and non-naïve are not the correct terminology when also “not” naïve is used in the same sentence (see line 95). Recommend to use naïve for those that are PCR negative and “exposed” for those that are PCR positive and thus non-naïve.

Please review the correct use of COVID-19 and SARS-CoV-2.

Line 120; administered on Day 1 is not informative, it would be more informative to indicate how many days after the 1st booster it is.

Line 122; the authors describe prototypes but the correct term is “variant or strains”, as only the Wuhan-Hu1 is a prototype, ancestral or wild-type and refers to the original Wuhan strain. In line 152 these are then called variants again. In Fig 1 prototype and wildtype/prototype are being used. Please review.

Line 128; please specify “illnesses”, based on the previous sentence it seemed the authors mean severe disease, but illnesses in this setting could also mean mild symptomatic (with fever, cough or respiratory distress?)

Line 131; Please indicate whether baseline is day 1, or before the primary vaccination regimen. As the authors used specific naming of day 1 in line 120, day 0 can be considered baseline, but is actually after 2 vaccine doses.

Line 131; suggestion to start a new paragraph as from there the focus changes.

Lines 153 and 156; suggestion to remove cross-reactive, if the T cells target the same epitope, although conserved, then from a T cell perspective this is not cross-reactive.

Line 190; suggestion to add; T cell markers expressed by SARS-CoV-2-specific T cells

Line 204; please add COVAIL to stages, without this it is not clear what stages mean, also, this can perhaps be considered as methods, since the authors do not explain what the stages mean.

Lines 218-223 can be shortened and perhaps included in lines 211-212 since almost all individuals included also had T cell data.

Lines 225-239 are all not results but part of a method section. Please adjust.

Lines 307-311, Please clarify the hypothesis on “cross-reactive” (which are actually conserved-region specific) T cells in line with these results.

Lines 320-328; These comparisons are not shown in a figure, and this text could be shortened with the use of a figure.

Lines 355-356; the authors mean the antigen-specific for primary. Please adjust, it now reads as non-specific responses.

Line 398; Please rephrase “marker”

Line 508; CD154 is not a cytokine, please adjust.

Line 858; CXCR5 and PD1 are not cytokines. Please adjust

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

The manuscript “T-cell immune correlates in COVAIL trial participants by prior SARS-CoV-2 infection status” by Zhang et al. investigated the correlation of T cell immune responses with risk or protection of COVID-19 in COVAIL trial participants following a second variant-adapted vaccine booster. In this study the CD4 and CD8 T cell responses against spike and N peptides in participants were assessed by using a comprehensive 27-colour intracellular cytokine staining assay at two time points, the same day prior to (D1) and 14 days (D15) after a second mono- or bivalent mRNA or protein-based booster vaccine. The correlation of T cell immune index at D1 or D15 with COVID-19 risk was further analyzed in a 6-month follow-up. The main finding of this study was that CD4 spike T cell IFN- γ and/or IL-2 spike BA.4/5 responses correlated with low COVID-19 risk in participants with prior infection (non-naïve) but not in naïve participants.

Strength: 1. The topic the study address is interesting and also important. So far, the correlation of T cell immunity established from previous infection as well as vaccination with protection of SARS-CoV-2 infection is less investigated compared to humoral immunity. This information is very important for public health and vaccination strategies.

2. The study is based on a randomized clinical trial with a great number of participants (n=932) and the authors utilized a comprehensive intercellular cytokine staining assay with 27-markers to analyze T cell immunity against spike and nucleocapsid. Furthermore, the authors performed comprehensive analysis of this large dataset and also a developed machine learning model Superlearner to assess T cell immune correlates with risk of COVID-19.

Major concerns

1. For non-naïve participants, non-spike specific T cells play an important role in the control of viral infections. Lang-Meli et al. 2022 (<https://doi.org/10.1038/s41564-022-01106-y>) showed that spike-derived CD8 T cell epitopes were not dominantly targeted in individuals with infection, non-structural proteins such as ORF1ab contained the main targeted antigens. This implied that non-spike specific T cells, especially CD8 T cells, play an important role in non-naïve individuals. Therefore, spike-specific CD8 T cell immunity might not be the strong immune correlates of protection in this group. And the design of this study might not be suitable to find the optimal immune correlates in non-naïve individuals.

On the other hand, Lang-Meli et al. 2022 also demonstrated that the breadth of spike-derived CD4 T cell epitopes targeted in convalescent individuals is broader than that in 2x vaccinated individuals. Moreover, booster vaccination increased the breadth of the spike-specific CD4 and CD8 T-cell responses in convalescent individuals but not in individuals with two vaccinations. These results are consistent with the main finding of this current manuscript that the CD4-specific T cells prior to the second booster vaccine correlate with low risk of COVID-19 infection in non-naïve group but not in the naïve group.

Since the COVAIL trial was designed to focus on analysis of T cell immune response against mainly Spike and N, the authors should describe these and emphasize that Spike and N-specific T cells alone could not represent a comprehensive immune landscape in non-naïve individuals.

2. The individuals in each stage of the COVAIL trial showed considerable heterogeneity the viral insert in the booster vaccine. This may negatively affect the power of correlation analysis when they were grouped.
3. The authors performed a variety of correlation assay and thereby provide a large dataset with analyses and figures. However, it is difficult to follow the authors' narrative through the manuscript. It seems that a clear and biologically logic line is missing. The authors may polish the manuscript and make it more readable and concise.
4. Given the complex immune landscape against SARS-CoV-2 at the time of conduction this trial, the study includes heterogeneous subgroups of participants with different infection histories, different vaccination schemes etc. This was not easy to avoid at that time, and the authors try to address this in their subgroupings and multivariate analyses. However, this should be considered a limitation of this study and be emphasized as such.

Minor concerns

1. 27-colour ICS method were utilized to analyze functional T cells expressing cytokines. As expected the percentages of cytokine-expressing T cells, especially CD4 T cells, in PBMC of participants are in the lower range with the majority between 0.05-1%. The reliability of specific responses below 0.1% is limited. A more detailed description of the ICS method will be helpful to show the reliability.
2. In the manuscript long sentences were often used which were not easy to follow. For example, line 120-124. It could help to shorten them.
3. The magnitude of the T-cell response prior to the booster vaccination might be a proxy marker of recent infection, which could explain the protective correlations of this biomarker against SARS-CoV-2 infection in the follow up period. Is there more detailed data about the timepoint of prior infections before the booster vaccination (i.e. timepoints of infections before the D1 timepoint)? This could be used as a separate parameter if available.

Abstract

- Line 93: greater -> stronger
- Line 97: Cox modeling of CD4+ T cells -> spike CD4+ T cells
- Line 101: titers-> neutralizing antibody titers
- Line 106: CoP -> correlates of protection (CoP)

Introduction

- Line 127: would asymptomatic participants with COVID-19 breakthrough infection be identified? What was the percentage of non-naïve individuals without positive anti-N response?
- Line 129: would SARS-CoV-2 variants in all infected participants with symptoms be characterized?

Results

- Line 209: please provide a clear schema of COVAIL trial (combining the Fig.S1 and Table 1 in SAP) similar as supplementary Figure S1 for from Zhang et al. 2025. Explain all the groups of vaccine boost arm as mentioned in section 2 of Statistical Analysis Plan (SAP) in the supplementary file.
- Line 213-214: non-case and case were mentioned for first time in this part, explain them as participants without infection and infections, respectively. The detail descriptions were in
- Line 551-552: EM -> Tem, CM -> Tcm.

Methods:

- L857-859: what is the reason for changing 3 antibodies? Please describe the ICS FACS protocol in more details, for example, the amount of PBMCs analyzed, the FACS gating strategy

Supplementary info.

1. On Index page, the Figure S5 title is missing
2. Figure S17 legend, (B,D, D-F), (B,D) are redundant.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

The aim of the study by Zhang et al. was to investigate immune correlates for COVID-19 risk or protection through SARS-CoV-2-specific T-cell responses after vaccination, including in relation to the prevailing variants of concern (VoCs) and variant-adapted boosters.

Overall, this is a relatively large study involving 970 recipients of a single-dose vaccination in phases 1, 2, and 3, with (polyfunctional) T cell determination performed only for a subset. Participants were randomly assigned to one of four phases (no T cell detection in phase 4) and received a second booster vaccination (day 1, D1), with immunological parameters measured again 14 days after the second booster vaccination (D15).

Analysis of CD4+ and CD8+ T-cell responses against spike (S) and nucleocapsid (N) peptides was performed using a 27-color intracellular cytokine staining (ICS) assay in vaccinated subjects from the COVAIL (COVID-19 Variant Immunologic Landscape) study over an extended period of time.

The study was not designed to evaluate the relative effectiveness of booster shots in preventing COVID-19, but immunological correlates for risk assessment could be found based on the COVID-19 symptoms reported by participants.

From the results, the authors conclude that spike CD4⁺ T cells are a non-mechanistic CoP that serves as a surrogate for one or more unmeasured protective immune mechanisms that are only effective after prior SARS-CoV-2 infection.

The results of the study showed that while neutralizing antibodies are the dominant protective factor, CD4⁺ T-cell responses provide important complementary protection, especially in individuals with lower antibody concentrations. This suggests that both humoral and cellular immunity should be considered in vaccine evaluation and risk stratification strategies.

Overall, this is an extremely comprehensive study involving a very large cohort, data analyses at two different time points after booster vaccinations with different booster vaccines, etc.

Nevertheless, I find the truly novel insights gained from this study to be limited. The fact that memory T-cell responses can be enhanced by vaccination or infection is not new, and neither is the data showing the synergistic protective effect of the combination of neutralizing antibodies and CD4⁺ T-cell responses.

Major Concerns:

Most results are well documented, but are described with long insertions, which impairs readability. A summary of the areas and highlighting the direction and strength of the observed results would provide greater clarity and improve readability.

The definitions of endpoints and cases/non-cases are precise and consistent with previous COVAIL publications and CDC criteria, but the repeated references to previous articles, etc., make the text cumbersome.

The use of overlapping peptide pools generally has the advantage of being “HLA-independent.” Nevertheless, the peptides must of course be presented in HLA molecules. How can the authors be sure that peptides with sequence differences between the individual variants are actually presented (and not just conserved regions) and that the T-cell responses are then actually specific to these differences? What is the proportion of “non-matching” peptides in the pools and for the comparison of variants in the boosters? The authors should at least discuss this.

The nature of overlapping pools of 15-mer peptides generally induces stronger CD4 than CD8 T cell responses. How can the authors ensure that CD8 responses are not overlooked solely because peptide presentation in HLA class I is not optimal?

Although this study provides valuable insights into the relationship between CD4⁺ T cell responses and vaccine efficacy, several methodological limitations and interpretive challenges must be considered. The findings are largely consistent with the current understanding of the immunological correlates of protection, but the evidence for T cell markers as independent correlates is weaker than suggested by the authors. The study would benefit from more rigorous statistical validation and clearer communication of the uncertainty in the results.

The key findings that higher S-specific CD4⁺ than CD8⁺ T-cell responses, virtually no N-specific responses in naive participants, and higher baseline responses in non-naive participants were to be expected and are not new, but well documented. But what about cross-reactivity to other coronaviruses, which has also been demonstrated for N-specific T cells in SAR-CoV-2-naive individuals? And again, the question arises: are CD8⁺ T cell responses being overlooked?

The conclusion that on day 15 (D15) the CD4⁺ S-specific T-cell responses between naive and non-naive cohorts are “balanced” after the booster vaccination is justified and expected for CD4⁺, but the persistent CD8⁺ differences and p-values are only briefly mentioned. A more explicit contrast would better support the mechanistic interpretation of the different priming responses of CD4⁺ and CD8⁺ compartments, if possible.

High correlations between spike peptide pools and time points and weak spike, N, and T cell nAb correlations are clearly described, but here too, independence of humoral and cellular responses was to be expected, as described for both naive and vaccinated subjects. The authors showed, that combination of high booster-elicited neutralizing antibodies and high pre-existing S-specific CD4⁺ T-cell responses from prior infection resulted in very low COVID-19 risk. It is not clear whether this is also due to a broader T-cell repertoire and the presence of T-cells against the N and M protein generated during infection. If necessary, a statement on the possible persistence of the Ak and different T-cells against different viral antigens would be helpful here.

The Cox and controlled risk analyses show convincing inverse CD4⁺ correlates in the non-naive cohort, but not in the naive cohort, but the repeated emphasis on “clearly indicated inverse correlations” could be tempered by acknowledging the moderate HRs, some p-values close to conventional thresholds, and limited case numbers, particularly in the subgroup and functionality score analyses.

The COMPASS-based findings on polyfunctionality and Leiden clustering are interesting, well presented, and consistent with the most important CD4⁺ T-cell signal (EM/CM CD4⁺ subsets that co-express CD154, IFN- γ , IL-2, and TNF- α and are enriched in non-naive cases), but the description is very technical and contains many “marker” lists. Here, it could be summarized more quickly and concisely that polyfunctional effector memory CD4⁺ T cells are associated with protection in non-naive participants—which is, after all, the core message.

The exploratory nature of the clustering analysis is only mentioned implicitly; it would strengthen the section to explicitly state that these are hypothesis-generating results to avoid overinterpretation of specific clusters.

The definitions of endpoints and cases/non-cases are precise and consistent with previous COVAIL publications and CDC criteria, but the repeated references to previous articles and SAP make the text cumbersome (a concise cross-reference would suffice).

Minor note:

The manuscript is very long and sometimes difficult to read, and it often takes a long time to find the key message of each chapter. Shortening the text and presenting the most important points more quickly would improve readability.

Abbreviations are often not introduced

In the chapter "Machine Learning," terms such as "baseline adjustment factors" and 'superlearner' are not defined, and to my knowledge, "machine superlearning" is not standardized terminology.

(Remarks on code availability)

In my opinion, the results of the work are reproducible. The code has been opened.

Reviewer #4

(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)

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

All my points have been addressed sufficiently.

(Remarks on code availability)

I have not reviewed the code

Reviewer #3

(Remarks to the Author)

The manuscript has become significantly clearer and more readable as a result of the abridgement and substantial revision.

I do not yet find the data and conclusions regarding peptide presentation and T-cell recognition of the variable regions convincing, especially since simple peptide presentation predictions for selected common HLAs could have provided some guidance here.

(Remarks on code availability)

Reviewer #4

(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 #5

(Remarks to the Author)

This revision responds to prior critiques although it still has poor readability and mostly limited novel findings. In my opinion the primary message of this manuscript is presented in Fig. 3 and 4. In the revision, the logic of the analysis is more clear although on the whole the manuscript is extremely difficult to comprehend. The main conclusions can be stressed better. Strengths include the extensive description, rigorous data and statistical analysis. The authors stress that manuscript has been shortened, but unfortunately, for non-highly biostatistics trained researchers the manuscript is cumbersome. One of the main conclusions that high spike and somewhat NC CD4 T cell responses in non-naïve individuals correlate with lower COVID-19 is not necessarily news worthy. There are multiple prior studies that have reached a similar conclusion

often with spike, NC or other SARS-CoV-2 regions. The authors and prior reviewers have highlighted some of the previous publications. While the number of individuals analyzed in this study is a major strength; the authors, however, rightly point out the heterogeneity in the population is a limitation. Besides the number of individuals, the author could better highlight the novel findings. The significance of these findings relative to prior knowledge is somewhat marginal or incremental. It is generally well established that neutralizing antibodies along with T cell responses are likely important in ameliorating disease.

Readability can be improved by focusing on the main points. One of the main ideas is differences in T cell responses among naïve versus non-naïve individuals. Figure 1 and Figure 2 do not convey this point adequately because it muddled with vaccine antigen and stage. There should be a figure showing direct comparisons. In Figure 1 and 2, it is unclear how “negative” can have higher values than “responders.” Highlighting this in the main text will improve readability. While it is important to consider the vaccine inserts, stages, days post vaccination, would this main point be conveyed by showing naïve versus non-naïve responses as a main figure. Different individuals can be symbol and color coded based on vaccine and insert etc. The current form is not very understandable.

Another major finding is that NC and spike CD4+ T cell responses are a correlate of protection against COVID-19 in the non-naïve population. Figure 3 and 4 could be simplified by reducing the number of panels. As mentioned by multiple prior reviewers, this finding is not necessarily novel. I find the discussion regarding this point informative. They may also consider that It is likely that enhanced CD4 responses don't protect against infection but rather ameliorate COVID19 disease. Thus individuals with high CD4+ T cell directed responses were likely not diagnosed with COVID-19 because it may have been mostly asymptomatic or not meriting any diagnostic evaluation.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #3

(Remarks to the Author)

As mentioned before, all my points have been addressed sufficiently.

Overall, the manuscript has become significantly more readable.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)

I have no further comments. The authors have addressed my prior concerns as well as possible.

(Remarks on code availability)

None

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Note: All lines numbers refer to the clean version of the Word document of our revision.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this study, Bo Zhang and colleagues, investigated correlates of T cell protection by stratifying their COVAIL trial participants according to their infection status and prior booster vaccine regimen. Although the authors show important findings related to their research objective, they expanded their analysis with immunogenicity against 3 different vaccines. This makes the manuscript more of a mixture of two different objectives, although somewhat related, and does not contribute to the readability.

Part of the immunogenicity data have been published before (e.g. CD4/CD8 ratio), and the main results are not clearly supported by the figures (Fig1 and Fig2). The part of the manuscript that contributes to the main objective to assess the correlates, the authors write that they have an improved functionality assessment of the antigen-specific T cells compared to other studies, this is not sufficiently introduced and also not used to its full potential. For example, the authors say to fill the “gap” with T-cell “markers”, while they actually look at antigen-specific T cells, using a subset of cytokines + CD154, where only a few represent the majority of the antigen-specific response. The authors do not make use of the other potential markers, or markers on those antigen-specific T cells. In addition, the shown correlations are observational and not mechanistic, providing no explanation about these differences. For example, there is no hypothesis on why polyfunctional CD4 T cell responses were found in higher numbers in the non-cases. Where these cells differ in terms of TCRs compared to the non-polyfunctional cells? Since stimulated with a pool of peptides, could the polyfunctionality truly arise from conserved epitopes? Or are these polyfunctional responses a result of a mutation-driven response, something that could be identified partly with TCR seq.

Major;

The title does not fully represent the results of the research findings as it assumes that these are correlates found at baseline (before actual infection and/or vaccination) but baseline is actually after the primary vaccination series. A suggestion for a more representative title could be; Distinct T cell correlates for the protection against COVID19 after booster for pre-exposed and naïve SARS-CoV-2 individuals.

Response: Given that differences were observed between naïve vs. non-naïve participants in D1 (pre-second booster) T-cell responses as well as in D15 (post-

second booster) T-cell responses, we have slightly revised the reviewer's suggestion to: "Distinct T-cell correlates of COVID-19 risk before and after a second booster in previously SARS-CoV-2 infected vs. naïve COVAIL trial participants."

Adhering to the maximum word count will enhance the readability and reviewing of the manuscript.

Response: We have shortened word count, and the length of the text (minus Methods) has been reduced from 9294 to 5563 words (within ~10% of the ideal word limit, 5000 words, as stated on the journal's webpage).

Introducing more structure of the introduction will be improved readability. The introduction in its current state reads as the results section of the COVAIL trials that were previously published (ref 10 and ref 11) and can be shortened, and perhaps work from others in the same field can be referenced to widen the angle.

Response: Thank you for this suggestion. We have revised the Introduction as noted, including shortening the description of the previous COVAIL results as well as adding work from others in the same field, including adding references of previous T-cell correlates studies ("Compared to the large body of work assessing antibody responses to SARS-CoV-2 as immune correlates of COVID-19,^{1,2} fewer studies have focused on T-cells³⁻⁶..." lines 103-104) and "Complementing these findings, a longitudinal study reported that a combination of antibody and T-cell response magnitude thresholds predicted breakthrough SARS-CoV-2 infection after a two-dose primary series (mRNA or viral vector) in non-naïve, but not naïve, individuals." (lines 124-127)

Naïve individuals are self-test negative or negative for anti-N. How well does the anti-N predict pre-exposure, as it wanes over time? Instead of focusing on the distribution of the CD4 and CD8 T cell response, it might be more informative to elucidate more on the naïve vs. exposed individuals and their response to N and perform analysis on that. This would strengthen the claim that these individuals are more likely naïve and did not have an asymptomatic infection in the past which was missed. In line with this it is vital to understand and know how many are considered naïve based on self-tests and how many are actually anti-N negative? Naïve could be a group comprised of asymptomatics that have never had a reason to perform a self-test. This result should be included in the first paragraph of the paper.

Response: We have clarified that nearly all non-naïve participants had a positive anti-Nucleocapsid test result at baseline:

“Of the 387 (68.0%) naïve participants in the PP-IAS, 183 were female (mean age 51.6 years) and 204 were male (mean age 52.2 years). The remaining 182 (32.0%) were non-naïve (73 female, mean age 48.1 years and 109 male, mean age 46.7 years), almost all of whom (172/182, 94.5%) had a positive anti-Nucleocapsid test result at baseline. The remaining 10 non-naïve participants (5.5%) tested anti-Nucleocapsid negative at baseline and were classified as non-naïve due to self-reporting previous SARS-CoV-2 infection.” (lines 146-151)

As to the reviewer’s suggestion to perform separate analyses of Nucleocapsid-specific responses in naïve and non-naïve participants, this was done for the immunogenicity analyses, but not the correlates analyses. The relevant immunogenicity results are shown, separately for naïve and non-naïve participants, in Table 1, Supplementary Table 4, and Supplementary Figures 3-6 (geometric mean responses and positive response proportions/positive vaccine response proportions, distributions of response magnitudes). We now include “As expected due to their lack of prior SARS-CoV-2 infection, D1 positive response proportions for N Index-specific (“Index” refers to Wuhan-Hu-1) CD4+ and CD8+ T cells in the naïve cohort were zero or near zero (CD4+: 0.0% in 7/14 one-dose booster arms and ranging from 2.5% to 9.2% in the other 7/14 arms; CD8+ 0.0% in 12/14 arms and 2.2% to 2.6% in the other 2/14 arms (Supplementary Figs. 3, 4). In contrast, in the non-naïve cohort, the D1 positive response proportions ranged from 8.1% to 42.5% (CD4+) and from 0.0% (6 arms) to 21.5% (CD8+) across one-dose booster arms (Supplementary Figs. 5, 6).” (lines 178-184).

Nucleocapsid-specific T cell responses were not studied as correlates in naïve participants as prespecified in the SAP; thus, we do not present Nucleocapsid-specific correlates results in naïve participants.

As to the persistence of anti-Nucleocapsid antibodies, we have added the following to the limitations section in the Discussion:

“Ninth, given that the majority of the non-naïve individuals were classified as non-naïve based on a positive anti-Nucleocapsid test result at baseline and that anti-Nucleocapsid seropositivity has been shown to decline more rapidly than anti-Spike seropositivity,⁸ it is possible that some individuals had experienced previous infection yet their anti-Nucleocapsid antibodies had waned such that they were not detected in the baseline serology test, leading to mis-classification as naïve. However, the low positive response rates and magnitudes to Nucleocapsid in the naïve cohort suggest that, if this occurred, it was only in a minority of participants and is unlikely to influence the conclusions.” (lines 546-553)

Since the data rely on the antigen-specific T cell data which are generated via their ICS panel, it is essential to also show a gating strategy and include a representative

example of the gating. It is also not clear which control has been used in these assays, it seems the authors have used medium control as negative control, where they should have used a DMSO control with equimolar DMSO% compared to the peptide mixes. Please clarify.

Response: The gating strategy is now detailed in Supplementary Figure 32, and Supplementary Figures 33 and 34 provide representative examples (from the same participant) of gating after stimulation with Spike BA.4/5 S1 and the negative control, respectively.

Additionally, we have added the following details on the negative control:

“Peptide diluent [dimethyl sulfoxide (DMSO) at a final concentration of 0.5%] served as the negative control. The peptide pools had varying concentrations of DMSO, depending on the number of peptides in the pool and/or the amount of DMSO required for reconstituting. The final concentration of DMSO in the negative control (0.5%) was chosen such that it would always be higher than the final concentration of DMSO in the peptide pools. This has been our standard practice and is how the assay was validated (for the peptide pools used in this assay, the final concentrations of DMSO were: COV2 CON S1: lot 20K072A, 0.12% DMSO; COV2 CON S2: lot 20K073A, 0.12% DMSO; Spike BA.4/5 S1: lot 22Q070A, 0.12% DMSO; Spike BA.4/5 S2: lot 22Q075A, 0.18% DMSO; Nucleocapsid Index: lot 22P078A, 0.06% DMSO). [The COV2 CON S1 peptide pool is the Index strain but the pool also contains additional peptides covering the D614G amino acid mutation.]” (lines 613-624).

The results section is a mix of methods, figure legends without guiding the reader and actual results. Please structure.

Response: In the process of revising and shortening, we have moved portions of Results text into a combination of Methods, Supplementary Text, and figure captions, or removed Results text where the essential information was already present in one of the above. Specific examples include removing the methods description of the controlled relative vaccine efficacy curves (already present in lines 754-763), removing the figure-caption like description of Fig. 6 (section “High booster-elicited neutralizing antibodies with high pre-existing Spike-specific CD4+ T-cell responses”), and moving portions of the superlearning results section to Methods (within lines 777-816).

Methods; the description of the ICS assay is not clear, no peptide pool supplier is listed, nor the used concentration, also no stimulation duration was listed. Was there no

DMSO control as negative control used? (see comment above). Since the results of the manuscript are based on this assay, it is expected to show more details.

Response: We have added comprehensive details on the ICS assay, including details on the peptide pool supplier (Bio-Synthesis Inc., Lewisville, TX), the used concentration (final concentration of 1 µg/mL per 15-mer peptide), and the stimulation duration (6 hours). The revised ICS Methods section now reads as follows (lines 591-673):

“SARS-CoV-2 Spike- and Nucleocapsid-specific CD4⁺ and CD8⁺ T-cell responses were evaluated using a validated (FDA Drug Master File 026862) 27-color ICS assay essentially as described,⁹ with the slight modification that CXCR5, PD-1, and IL-21 (for circulating T follicular helper cell detection) were used in replacement of CRTh2, CD32, and CD64. This switch was made due to increased interest in follicular helper T cells (Tfh) cells and their role in protective immunity. All antibodies were titrated upon receipt using human PBMC to verify performance and select the optimal titer. Titrations took place in the context of the intended use of the antibody in the assay. Antibodies detecting cytokines or other activation markers were titrated on PBMC that had been stimulated with Staphylococcus aureus enterotoxin B. Where a given lot of antibody was insufficient to assay all samples, the new lot of antibody was titrated upon receipt and bridged with the current lot to ensure equivalent performance.

For the ICS assay, after thaw and overnight rest at 37°C maintained with 5% CO₂, peripheral blood mononuclear cells (PBMCs) were plated at 1M-1.2M cells per well for each condition and stimulated ex vivo for 6 h with pooled peptides covering the SARS-CoV-2 D614G Spike subunit 1 (S1; comprising the N-terminal domain and the receptor-binding domain) and Spike subunit 2 (S2), SARS-CoV-2 Spike BA.4/5 S1 and S2, and the SARS-CoV-2 Nucleocapsid Index protein (final concentration of 1 µg/mL per 15-mer peptide, overlapping by 11 amino acids; Bio-Synthesis Inc., Lewisville, TX), in the presence of the secretory inhibitor brefeldin A (10 µg/mL; Sigma-Aldrich, St. Louis, MO) and the co-stimulatory antibodies CD28 and CD49d (each at 1 µg/mL; Becton Dickinson Biosciences, San Jose, CA). Fluorescently labeled CXCR5 phycoerythrin antibody (BioLegend, San Diego, CA) was also included in the stimulation cocktail. Stimulation with staphylococcal enterotoxin B (0.25 µg/ml; Sigma-Aldrich) served as a positive control. Peptide diluent [dimethyl sulfoxide (DMSO) at a final concentration of 0.5%] served as the negative control. The peptide pools had varying concentrations of DMSO, depending on the number of peptides in the pool and/or the amount of DMSO required for reconstituting. The final concentration of DMSO in the negative control (0.5%) was chosen such that it would always be higher than the final concentration of DMSO in the peptide pools. This has been our standard practice and is how the assay was validated (for the peptide pools used in this assay, the final concentrations of DMSO were: COV2

CON S1: lot 20K072A, 0.12% DMSO; COV2 CON S2: lot 20K073A, 0.12% DMSO; Spike BA.4/5 S1: lot 22Q070A, 0.12% DMSO; Spike BA.4/5 S2: lot 22Q075A, 0.18% DMSO; Nucleocapsid Index: lot 22P078A, 0.06% DMSO). [The COV2 CON S1 peptide pool is the Index strain but the pool also contains additional peptides covering the D614G amino acid mutation.]

Cells were incubated for 6 hours at 37°C maintained with 5% CO₂. After the incubation, 20 µL 20 mM EDTA 1X PBS was added to each well and the plates were transferred to 4°C for overnight storage. The next day, cells were washed and stained with viability dye (Thermo Fisher, Waltham, MA) in 1X PBS for 20 min in the dark. The cells were washed again and incubated with the surface staining antibodies [Supplementary Table 6, panel adapted from Cohen et al.¹⁰] in FACS Wash (2% FBS in 1X PBS) for 20 min in the dark. Subsequently, cells were fixed (BD FACS Lyse; BD Biosciences) for 10 min, centrifuged and permeabilized (BD FACS Perm II; BD Biosciences) for 10 min followed by two washes. Cells were then incubated with intracellular antibodies in FACS Wash in the dark for 30 min, followed by two washes and a final resuspension in FACS Wash prior to overnight storage at 4°C. Cells were acquired on a custom BD FACSymphony A5 (BD Biosciences). Data were acquired using BD FACSymphony A5 instruments equipped with 28 fluorescence detectors and 5 lasers (355 nm UV laser, 405 nm violet laser, 488 nm blue laser, 561 nm green laser, and 640 nm red laser). All instruments were previously characterized and standardized based on a slightly modified approach to the procedure described by Perfetto et al.¹¹ Instruments are also standardized to each other, allowing for interchangeable use. Prior to each acquisition, instruments were quality controlled to their optimal settings using hard dyed beads.

Collection of the cells on the cytometer resulted in approximately 200K-500K total events collected. Actual numbers of CD4⁺ and CD8⁺ T cells analyzed were participant- and stimulation-dependent, with some stimulations causing more toxicity. Data were analyzed using FlowJo version 9.9.4 (FlowJo LLC). Data were excluded if the number of CD4⁺ T cells acquired on the cytometer for a particular stimulation were <10,000, or CD8⁺ T cells <5,000. These criteria were pre-established in the assay Standard Operating Procedure. Selected data were excluded if the data were determined to be unreliable by the laboratory. Reliability was based on review of the data as analyzed in the FlowJo analysis software. It is very rare for data to be determined to be unreliable and this would typically be due to unusual staining profiles.

All ICS assays were conducted in a blinded manner.

Gating strategy

As visualized in Supplementary Fig. 32, cells were initially gated by Time followed by additional gates to remove aggregates/non-specific staining and dead cells.

Further gates identified lymphocytes, singlets, CD3+ T cells, and CD4+ and CD8+ T-cell subsets. Additional upstream gates also identified monocytes, B cells, natural killer (NK), and natural killer T (NKT) cells. For each of CD4+ and CD8+, the expression of functional and phenotypic markers was determined using marginal gates. Where applicable, placement of gates was determined using Fluorescence Minus One (FMO) testing, which defined the lowest possible bound of the gate. For cytokine gate placement, the lower bounds of the gate were placed higher than the FMO to improve specificity without negatively impacting sensitivity. Select cytokine and phenotyping gates were also applied to the monocytes, B cells, NK, and NKT cells.

Spike-specific and Nucleocapsid-specific CD8 T cell responses were defined as live, CD14-, FSClow SSClow, singlet, CD19-, CD16-, CD56-, CD3+, and CD8+ T cells that expressed the designated cytokines/markers. Cytokine/marker positive Spike- and Nucleocapsid-specific CD4+ memory T cell populations were further characterized as naïve (CD45RA+CCR7+), effector memory (CD45RA-CCR7-), and central memory (CD45RA-CCR7+). Cytokine/marker positive Spike- and Nucleocapsid-specific CD8+ memory T cell populations were characterized as naïve (CD45RA+CCR7+), effector memory (CD45RA-CCR7-), and TEMRA (CD45RA+CCR7-).

Representative flow plots (from the same participant) after stimulation with Spike BA.4/5 S1 and after negative control stimulation are shown in Supplementary Figs. 33 and 34, respectively.”

Results from Figure 1; the conclusion from the authors is that CD4 T cells are more dominant at D1 compared to CD8 T cells both in naïve and non-naïve. The figures are plotted in such a way that they are difficult to interpret. A suggestion would be to split the data on showing only IFN γ +, since the CD4 T cells from peripheral blood secrete IL2 whereas CD8 T cells do not. As the gating strategy is missing, it is difficult to assess whether this is not simply comparing different entities: CD4 T cells producing IFN γ and IL2) versus CD8 T cells solely producing IFN γ . Nevertheless, previous research indeed already shows that the CD4 T cell pool dominates over the CD8 T cell pool.

Also, most of the anti-N responses were non-responders both in the naïve pool as well as in the non-naïve pool, because the data are now split across all arms, the authors can also show a certain % of responder from the exposed (non-naïve) versus % responders from naïve. Then it becomes more evident that indeed the naïves are truly naïve.

Response: Thank you for this comment. First, we now provide our gating strategy as Supplementary Figure 32, which helps clarify the separation of CD4+ vs. CD8+ T cells.

We also clarify in the text:

“Spike-specific and Nucleocapsid-specific CD8 T cell responses were defined as live, CD14-, FSClow SSClow, singlet, CD19-, CD16-, CD56-, CD3+, and CD8+ T cells that expressed the designated cytokines/markers. Spike-specific and Nucleocapsid-specific CD4 T cell responses were defined as live, CD14-, FSClow SSClow, singlet, CD19-, CD16-, CD56-, CD3+, and CD4+ T cells that expressed the designated cytokines/markers. Cytokine/marker positive Spike- and/or Nucleocapsid-specific CD4+ memory T cell populations were further characterized as naïve (CD45RA+CCR7+), effector memory (CD45RA-CCR7-), central memory (CD45RA-CCR7+), and terminally differentiated effector memory T cells (TEMRA; CD45RA+CCR7-).” (lines 664-671)

Our choice of plotting IFN- γ and/or IL-2 responses to Spike BA.4/5 in Figures 1 and 2 was made to encompass the majority of responding T cells, where most Spike-specific CD4+ T cells are expected to co-express both IFN- γ and IL-2. CD8+ T cells are more likely to express IFN- γ without IL-2 and this is also captured well by the cytokine endpoint parameter of IFN- γ and/or IL-2.

As to the reviewers’ comment that “most anti-N responses were non-responders both in the naïve pool and in the non-naïve pool”, Figure 1 does not show any Nucleocapsid-specific T-cell responses. These are shown in Supplementary Figures 3 through 6 (Supplementary Fig. 3: Naïve, CD4+ cells; Supplementary Fig. 4: Naïve, CD8+ cells; Supplementary Fig. 5: Non-naïve, CD4+ cells; Supplementary Fig. 6: Non-naïve, CD8+ cells). In fact, the frequencies of Nucleocapsid-specific positive response (pre-second booster, D1) were low in the naïve cohort, as expected, but high in the non-naïve cohort. We have revised the results text on Supplementary Figures 3 to 6 to focus more on this result and to address the reviewer’s concern that the naïve cohort is truly naïve:

“As expected due to their lack of prior SARS-CoV-2 infection, D1 positive response proportions for N Index-specific (“Index” refers to Wuhan-Hu-1) CD4+ and CD8+ T cells in the naïve cohort were zero or near zero (CD4+: 0.0% in 7/14 one-dose booster arms and ranging from 2.5% to 9.2% in the other 7/14 arms; CD8+ 0.0% in 12/14 arms and 2.2% to 2.6% in the other 2/14 arms (Supplementary Figs. 3, 4). In contrast, in the non-naïve cohort, the D1 positive response proportions ranged from 8.1% to 42.5% (CD4+) and from 0.0% (6 arms) to 21.5% (CD8+) across one-dose booster arms (Supplementary Figs. 5, 6).” (lines 178-184)

The stages are not explained well, and therefore the results can be difficult to interpret. For example, lines 301-302 the authors write that the response is lower in stage 3 compared to stages 1 and 2. It would be more informative to refer to their primary one-dose vaccine regimen. Also, it is not easy for the reader to make the comparisons the authors show. For example, an overlaid dot plot can be generated for each “stage” whereby the dots are colored according to their D15 booster vaccine. Then, higher/lower can be compared much easier.

Response: We have added a new supplementary figure, Supplementary Figure 1, that visualizes the COVAIL stages as well as provides detail on the specific study vaccine administered for each COVAIL study arm in panel A. Panel B of Supplementary Figure 1 shows the COVAIL study timeline, including the study inclusion criterion of receipt of a complete primary and booster COVID-19 vaccine series prior to enrolment in the COVAIL study as well as the day of receipt of COVAIL study vaccine (second boost) on D1.

As to the reviewer’s suggestion to refer to participants’ primary one-dose regimen, participants received different primary one-dose regimens as we clarify in the caption of Supplementary Fig. 1: “Before enrollment, all participants received a complete primary and booster COVID-19 vaccine series (eligible regimens listed in Table S2 of Branche et al.¹²).” We do not have data on which primary one-dose regimen the participants received other than it was one of the eligible regimens. Note we have added as a limitation to the Discussion: “Second, the study included heterogeneous subgroups of participants with different infection histories (variant, timing), different primary series and boosters received, etc., and the extent to which this heterogeneity may have confounded the results is unknown.” (lines 523-526)

As to the reviewer’s suggestion to overlay the plot where the dot is colored according to the D15 booster vaccine, Figs. 1 and 2 already display participants separately according to D1 second booster vaccine (i.e., COVAIL study vaccine, the relevant vaccine in this analysis). No booster vaccine was received at D15, which is clarified in the new Supplementary Fig. 1B.

For additional clarity we have made the following revisions (underlined) to the captions of Figures 1 and 2:

“Figure 1. Distributions of T-cell responses at D1 and at D15, shown separately for each one-dose vaccine second booster arm, in a representative random cohort (non-cases plus a subset of cases) of the naïve cohort. Study vaccine (second booster) was administered on D1 and corresponded to an mRNA-1273 vaccine (stage 1, Moderna), a BNT162b2 vaccine (stage 2, Pfizer–BioNTech, or a pre-S DTM

AS03 protein vaccine (stage 3, Sanofi), with the specific characteristics of each study arm on the x-axis (wild type vs. variant, monovalent vs. bivalent)."

"Figure 2. Distributions of T-cell responses at D1 and at D15, shown separately for each one-dose vaccine second booster arm, in a representative random cohort (non-cases plus a subset of cases) of the non-naïve cohort. Study vaccine (second booster) was administered on D1 and corresponded to an mRNA-1273 vaccine (stage 1, Moderna), a BNT162b2 vaccine (stage 2, Pfizer–BioNTech, or a pre-S DTM AS03 protein vaccine (stage 3, Sanofi), with the specific characteristics of each study arm on the x-axis (wild type vs. variant, monovalent vs. bivalent)."

Lines 330-344; Suggestion to replace marker for upregulation of CD154 and production of xxxx (markers are not "produced", cytokines/ILs are). Also, this section would benefit from a clearer introduction. The reader cannot assess what the functionality score actually is, how many cytokines (or markers) are considered? The terminology remains unclear. As an example, what is the proportion of CD154 and the production of these cytokines for CD4 and CD8 responders?

Response: We have clarified at the first introduction of "functionality scores":
"...functionality scores, which represent the overall proportion of different antigen-specific T cell subsets defined by expression/production of eight different cytokines/markers, were higher..." (lines 210-212) The specific eight cytokines/markers, provided in Methods, are: "Functionality scores were calculated as non-negligible combinations of a set of eight cytokines/markers: CD154, IFN- γ , IL-2, IL-4, IL-5 and/or IL-13, IL-17a, IL-21, and TNF- α ..." (lines 692-695)

We have also revised the legend of Figure 5 pertaining to the reviewer's comment:

"All cytokines/markers included in the COMPASS analyses (CD154, IFN- γ , IL-2, IL-4, IL-5 and/or IL-13, IL-17a, IL-21, and TNF- α) were assessed, but only cytokine/marker combinations for which at least 25% of observations had a probability of a positive response of 0.8 or greater [as estimated by COMPASS¹³] are plotted here. The specific cytokine(s)/marker(s) defining each subset are designated in the lower panel of each plot, where a filled square designates that the designated cytokine(s)/marker(s) were expressed in response to stimulation with Spike BA.4/5 peptide pools.

We also revised in the text to clarify that the subset analysis (Fig. 5) was performed to try to understand which subsets were driving the polyfunctional score inverse correlate results:

"Polyfunctional effector memory CD4⁺ T cells are associated with protection in non-naïve participants

As the functionality score represents the overall proportion of antigen-specific T cell polyfunctional subsets, we sought further insight by considering individually the antigen-specific T cell polyfunctional subsets that contribute to the functionality score.” (lines 336-338)

Lines-346-383; This part describes that, when stage 1 and stage 2 results are combined, CD4 and CD8 responses are higher at D1 for those that have previously been exposed. Then, after a booster, CD4 levels equalize between those naïve and exposed while this is not the case for CD8 T cells. Suggestion to reduce this section if this is the conclusion the authors want to make. Also, we suggest to use a figure instead of a 2-page long table (lines-218 till 345).

Response: Thank you for the suggestion. We have reduced and simplified this section, focusing on the conclusion outlined by the reviewer.

On the suggestion to use a figure instead of a table, we agree the table was long when double-spaced. We have single-spaced the table and revised its formatting to aid legibility. We prefer the table format for completeness, because it allows the reader to look up specific GM values, as well as proportions with positive response (D1) or positive vaccine response (D15). (We have made the same revisions to Supplementary Table 4 as well).

Lines 385-394; this is known and can be removed. If the authors do want to integrate it, we suggest to refer to it in one sentence during the ICS results highlighting, Spike BA peptides correlated with the Index, and D1 correlated with D15 Frequencies, whereas Spike responses did not correlate with N.

Response: We have reduced this portion by 40% (116 to 71 words, now only two sentences). Given that the correlation analyses were performed in the Per-Protocol Immunogenicity Analysis Set (PP-IAS), we prefer not to integrate these into the immunogenicity results, given that those analyses were performed in the representative random cohort of the non-cases in the PP-IAS plus a subset of cases in the PP-IAS.

It is not entirely made clear that lines 435 onwards show results of the risk of COVID19 after the booster regimen. It would be informative to add a simplified figure showing numbers of how many cases and non-cases there were after the booster. Also suggest to reduce the length of the text.

Response: We now lead this section with the sentence “Univariable correlate of risk analyses in the non-naïve cohort showed that Spike-specific and Nucleocapsid-specific CD4+ T-cell responses correlated inversely with COVID-19

risk over ~3 months follow-up, as well as over ~6 months follow-up” (lines 285-288).

We also now include information on cases and non-cases (underlined) in the Fig. 3 caption:

“Data points are from eligible COVID-19 cases (naïve: n=122, non-naïve: n=22) and non-cases (naïve: n=209, non-naïve: n=117) in the Per-Protocol Immunogenicity Analysis Set.”

In addition, the Supplementary Fig. 2 flowchart provides the numbers of cases and non-cases after the second booster that were included in the immune correlates analyses. To clarify, we have repeated the non-case and case definitions in the caption: “Non-cases: No evidence of SARS-CoV-2 infection after D1 and through to 188 days post D15; booster-proximal cases: COVID-19 endpoint between 7 and 91 days post-D15; booster-distal cases: COVID-19 endpoint between 92 and 188 days post-D15”, as well as clarifying that the correlates analyses were conducted in the Per-Protocol Immunogenicity Analysis Set: “through to membership in the Per-Protocol Immunogenicity Analysis Set, in which the correlates analyses were conducted.”

We have also reduced this section from 583 to 407 words.

The word “marker” is used throughout the text, since the authors describe a 27 marker panel in the introduction, focusing on T cells, it is expected that certain markers on the antigen-specific T cells drive this “correlation”, please rephrase since this is now misleading, or introduce the panel as a result showing marker expression of these antigen-specific T cells and compared across stages/ groups using for example UMAPs. Markers are mostly just IFN γ and/or IL2, and it is unclear whether the authors measured these on those antigen-specific T cells. The authors only state that a panel adapted from Cohen is used. Please add a table at line 862.

Response: Thank you for the suggestion. We have searched the main text and supplement and largely replaced “marker” with “response” (or removed, as for functionality score), although in some occasions we keep “marker” in the context of specific cytokine(s)/marker(s) expressed.

The requested antibody staining table is now included as Supplementary Table 6.

The authors find an inverse correlation for the exposed individuals for the HR of COVID19 within the first 90 days of booster. Have the authors looked into the antigen-specific T cells of those subjects and found an interesting phenotype or poly/oligo-

functional properties of these cells from those subjects? Please clarify.

Response: Thank you for the question. We believe the reviewer is referring to the inverse correlations of pre-second booster (D1), as well as post-second booster (D15), Spike-specific and Nucleocapsid-specific T-cell responses with COVID-19 risk when assessed over the boost-proximal period (COVID-19 endpoint between 7 and 91 days post-D15) in the non-naïve cohort.

Figure 5c and 5d and Supplementary Figure 27c and 27d present results on the antigen-specific T cells in these individuals, i.e. the non-naïve cohort. Specifically, Figure 5c and 5d show polyfunctionality results for Spike-specific CD4⁺ T cells in booster-proximal cases vs. non-cases in the non-naïve cohort. We found that polyfunctionality profiles appeared different between cases and non-cases:

“In the non-naïve cohort of the one-dose mRNA group, the proportions of multiple CD4⁺ T-cell sets defined by Spike BA.4/5-specific cytokine/marker combinations, including many with polyfunctionality degree three or higher, were higher in non-cases vs. booster-proximal cases at D1 and, to a potentially lesser extent, at D15 (Fig. 5c, 5d). In contrast, the same CD4⁺ T-cell subsets showed generally similar proportions between non-cases and booster-proximal cases in the naïve cohort (Fig. 5a, 5b).” (lines 338-344)

Supplementary Figure 27c and 27d present the exploratory clustering results, specifically the proportions of CD4⁺ T-cell clusters defined by cytokines/markers in each of the 7 Leiden clusters passing the MIMOSA-based positivity threshold in the non-naïve cohort:

“In general, clusters with the largest differential proportion between non-cases vs. booster-proximal cases in the non-naïve cohort tended to be classified as effector memory and to include IFN- γ , IL-2, CD154, and TNF- α (Supplementary Fig. 27c, 27d), suggesting that polyfunctional effector memory CD4⁺ T cells are associated with protection in non-naïve participants. Given the exploratory nature of this analysis, these results should be considered hypothesis-generating.

In the naïve cohort, cluster proportions were similar between booster-proximal cases and non-cases (Supplementary Fig. 27a, 27b).” (lines 347-354)

Given that previous studies suggest that T-cell response quality (vs. magnitude) is more important for determining infection outcomes (or vaccine responses) for various pathogens including TB¹⁴ and HIV¹³, the difference in polyfunctionality profiles between COVID-19 cases and non-cases in the present study is not surprising. We hypothesize that the greater polyfunctionality may be beneficial due to enhanced cytokine synergy and/or effector function.

Lines 472-474; please clarify where can the results of those 5 functionality markers are described. S15 does not show this separately, but as a sum.

Response: We have removed the use of the word “marker”, which we realize inadvertently introduced potential confusion, and now simply refer to “functionality score”. The functionality score is a summary measure of an individual’s quality of antigen-specific, polyfunctional T-cell responses, with each response defined by the simultaneous production of different markers (alone or in combination); thus it is the summary measure that is presented in Supplementary Figure 16 (Supplementary Figure 15 in the original submission).

The “5 functionality score markers” was intended to refer to the five functionality scores presented in Supplementary Figure 16:

- D1 CD4+ Spike BA.4/5 functionality score
- D1 CD8+ Spike BA.4/5 functionality score
- D1 CD4+ Nucleocapsid Index functionality score (in non-naïve participants only)
- D15 CD4+ Spike BA.4/5 functionality score
- D15 CD8+ Spike BA.4/5 functionality score

We have revised the text (underlined) to clarify:

“D1 and D15 Spike BA.4/5 functionality scores, which represent the overall proportion of different antigen-specific T cell subsets defined by expression/production of eight different cytokines/markers...” (lines 210-212)

The univariable correlate of risk analysis results for functionality scores can be found in Supplementary Fig. 16:

“A similar trend was seen for functionality scores as well as the other T-cell responses assessed (Supplementary Fig. 16).” (lines 297-299)

Line 511; the amount of markers, and which markers are changed constantly, without specifying the specific markers. The authors describe: defined by the expression of 9 markers, but the figure shows 5. Please clarify.

Response: Thank you for this opportunity to clarify. We have revised (underlined) the Figure 5 caption to state: “All cytokines/markers included in the COMPASS analyses (CD154, IFN- γ , IL-2, IL-4, IL-5 and/or IL-13, IL-17a, IL-21, and TNF- α) were assessed, but only cytokine/marker combinations for which at least 25% of observations had a probability of a positive response of 0.8 or greater [as estimated by COMPASS¹³] are plotted here. The specific cytokine(s)/marker(s)

defining each subset are designated in the lower panel of each plot, where a filled square designates that the designated cytokine(s)/marker(s) were expressed in response to stimulation with Spike BA.4/5 peptide pools.”

Lines 537-552; it is unclear whether the authors performed only clustering on the antigen-specific cells (by using which parameters/marker set?). It appears that this is not per se the case, if not, it is expected that the same populations with expressing cytokines are shown. Did the authors also compare the expression of PD1, CXCR3, CCR6 and other markers within those clusters between the groups? The latter would provide additional information and insights into the results of the inverse correlation. A suggestion is to perform a UMAP and measure the density level whether large shifts occurred.

Response: We have clarified that in the clustering, 1) CD4+ T cells expressing at least one of the eight cytokines and/or functional activation markers were extracted; 2) Next, from these cells, transformed fluorescence intensities were obtained and used for input in clustering (using all markers except the ones used in the FlowJo gating hierarchy); 3) After obtaining the Leiden clusters, the downstream statistical analyses are performed.

We have added additional details to the Methods section on the clustering, including the antigens used in the clustering and the expression of the different markers/cytokines:

“Leiden clustering¹⁵ was also performed as an exploratory analysis. CD4+ T cells expressing at least one of the eight cytokines and/or functional activation markers (IFN- γ , IL-2, TNF, CD154, IL-4, IL-17A, IL-21, and IL-5/13) were extracted and analyzed as described.¹⁶ Briefly, fluorescence intensity (FI) values for all markers in the ICS panel were transformed using an arcsinh transformation (cofactor 500) with an additional asymmetrical adjustment to deemphasize cells with highly negative FI in the arcsinh transformed distribution. Transformed data were then used for clustering, performed using the Leiden algorithm¹⁷ (using the Leiden R package) with $res = 1$, $k = 30$ and $n_iterations = 100$. All markers, except the ones being used in the *FlowJo* gating hierarchy to identify the CD4+ and CD8+ T cells (/Time/K1/K2/K3/K4/K5/K6/K7/K8/Lv/14-/L/S/19-/3+/3+excl 16br/56-16-/4+: UViD, CD14, CD19, CD3, CD56, CD16, CD4 and CD8), were used as input for clustering. Clustering was performed using all cells from all the samples combined [naïve and non-naïve; all one-dose second booster arms; six antigens (Nucleocapsid Index, BA.4/5 S1, BA.4/5 S2, COV2 CON S1, COV2 CON S2, negative control); and four visit numbers (D1, D15, D91, and D181)]. After filtering, a total of 2,501,319 CD4+ T cells were analyzed, yielding 21 distinct clusters. Before plotting cluster distributions by case/non-case status, positivity calls were made using the MIMOSA method⁴² by comparing the frequency of cells in each antigen-stimulated condition to the

negative control stimulation; 7 of the 21 clusters passed this threshold for at least one of the five test antigens. For each peptide pool, multiplicity adjustment was made. Individual peptide pool magnitudes < 0.001 were set to 0.001.

COMPASS and Leiden clustering were performed blinded to the meta data.” (lines 696-716)

We have also added more details to the caption of Supplementary Figure 26: “Leiden clustering of CD4+ T cells. The dot plot shows, for each of the 21 distinct Leiden clusters that were found (listed in Supplementary Table 5), percents of each of four different T cell subsets (top panel, gray: EM = effector memory (T_{em}), CM = central memory (T_{cm}), TEMRA = Terminally Differentiated Effector Memory cells re-expressing CD45RA) and percents of cells positive for expression of each of the cytokine(s)/marker(s) shown on the left. Clusters 1, 2, 4, 5, 6, 13, and 16 passed the MIMOSA-based positivity threshold for at least one of the five test antigens. Cluster 1 was clearly 4-function polyfunctional with expression of CD154 (>90%), IFN- γ (>90%), IL-2 (>90%) and TNF (>85%). Two main clusters (2 and 4) were 3-function polyfunctional. Data are from one independent experiment, run using a validated assay. After filtering, a total of 2,501,319 CD4+ T cells were included in this analysis.”

We also now include a UMAP figure of the Leiden clusters, which is Supplementary Figure 25 in the revision.

Minor:

Please check spelling consistency (e.g. T-cell or T cell)

Response: We have harmonized the manuscript to use “T-cell” when used as a compound modifier, and “T cell” when used as a noun.

Please review the terminology regarding specific responses throughout the paper. For example. Fig1 and Fig 2 the authors say that these figures resemble “markers”, while factually looking at the percentage of SARS-COV-2-specific CD4+ or CD8+ T-cells by looking at upregulation of IFN γ and/or IL2 after in vitro stimulation. Please adjust accordingly to improve clarity (Lines-242-247 for example, but also line 292 “marker” levels)

Response: Thank you for the suggestion. We have searched the main text and supplement and largely replaced “marker” with “response” (or removed, as for functionality score), although in some occasions we keep “marker” in the context of specific cytokine(s)/marker(s) expressed.

Line 292, marker levels, is unclear what the authors mean. This could be any of the markers in the ICS panel, while the authors probably mean (only) IFN γ or IL2

Response: Please see the response above. Note we also now clarify “response magnitudes (i.e., percentages of antigen-specific, cytokine/marker-positive T cells)” early in the manuscript (line 169).

What do the authors mean with N “index” ? Please clarify

Response: The use of “Index” rather than “Wuhan” as a label throughout the paper is based on WHO advice discouraging the association of the virus with specific locations like Wuhan (WHO, IFRC, and UNICEF: “Social Stigma associated with COVID-19”, 24 February 2020). We now clarify once in the text that “Index” refers to “Wuhan”, and avoid using the name of a specific geographical location after this clarification.

“(“Index” refers to Wuhan-Hu-1).” (line 179)

The SAP also clarifies that the Wuhan Nucleocapsid sequence is used (see section 5, “T cell markers included in analyses”).

Naïve and non-naïve are not the correct terminology when also “not” naïve is used in the same sentence (see line 95). Recommend to use naïve for those that are PCR negative and “exposed” for those that are PCR positive and thus non-naïve.

Response: We rectified this by adding “for” to this sentence: “for previously-infected (non-naïve) [hazard ratio (HR)=0.62 (95% CI 0.39, 0.99) per 3-fold increase], but not for SARS-CoV-2 naïve [HR=1.03 (0.81, 1.32)], participants” (lines 88-89)

We prefer to keep with the existing definitions of “naïve” and “non-naïve” given that they are consistent with our previous neutralizing antibody correlates analysis of COVAIL (Zhang et al.):

“As in Zhang et al.,¹⁸ participants were classified as SARS-CoV-2 non-naïve (hereafter, “non-naïve”) if they self-reported a previous SARS-CoV-2 infection or had detectable anti-N antibodies [defined as Elecsys Anti-SARS-CoV-2 assay (Roche) cutoff index ≥ 1.0] at D1. Otherwise, participants were classified as naïve.” (lines 577-580, in Methods)

Please review the correct use of COVID-19 and SARS-CoV-2.

Response: We have reviewed this.

Line 120; administered on Day 1 is not informative, it would be more informative to indicate how many days after the 1st booster it is.

Response: Day 1 marks enrolment in the COVAIL study and the day of receipt of the second booster (i.e. the COVAIL study vaccine), and the T-cell responses assessed as correlates in the present study were measured on Day 1 and 14 days later, on Day 15. Inclusion criteria for COVAIL included receipt of a primary vaccination series and one ancestral (Wild Type) booster dose; the first booster was not part of the COVAIL study. This is now clarified in the new Supplementary Figure 1, which shows that receipt of the complete primary and booster COVID-19 vaccine series must have occurred at least 16 weeks prior to Day 1.

Line 122; the authors describe prototypes but the correct term is “variant or strains”, as only the Wuhan-Hu1 is a prototype, ancestral or wild-type and refers to the original Wuhan strain. In line 152 these are then called variants again. In Fig 1 prototype and wildtype/prototype are being used. Please review.

Response: We have removed the use of “Prototype” throughout the manuscript and have harmonized this to “wild type”, including in all figures (Fig. 1, Fig. 2, Supplementary Figs. 3-10, Supplementary Fig. 28).

In the caption of Supplementary Figure 1, we now specify that “wild type” refers to the WA-1 ancestral strain.

Line 128; please specify “illnesses”, based on the previous sentence it seemed the authors mean severe disease, but illnesses in this setting could also mean mild symptomatic (with fever, cough or respiratory distress?)

Response: We have made the following revisions:

“ The COVID-19 endpoint was symptomatic SARS-CoV-2 infection, as examined previously.¹⁸” (line 152)

To provide more details on the symptoms of the COVID-19 cases, we have added Supplementary Table 1, which details the distributions of symptoms among the n=135 naïve cases and the n=26 non-naïve cases in the Per-Protocol Immunogenicity Analysis Set. It shows that, on average, naïve cases reported 6.5 symptoms per case, and non-naïve cases reported 5.7 symptoms per case. The most common three symptoms were congestion (85% of naïve cases and 65% of non-naïve cases), cough (80% of naïve cases and 77% of non-naïve cases), and runny nose (79% of naïve cases) or sore throat (73% of non-naïve cases).

Line 131; Please indicate whether baseline is day 1, or before the primary vaccination regimen. As the authors used specific naming of day 1 in line 120, day 0 can be considered baseline, but is actually after 2 vaccine doses.

Response: We have revised to: “Longitudinal sera and peripheral blood mononuclear cell (PBMC) samples were collected at D1 and at 14 days post-

second boost (D15)” (lines 117-118). The sampling scheme is also visualized in the new Supplementary Fig. 1, referenced in line 116.

Line 131; suggestion to start a new paragraph as from there the focus changes.

Response: We took this suggestion.

Lines 153 and 156; suggestion to remove cross-reactive, if the T cells target the same epitope, although conserved, then from a T cell perspective this is not cross-reactive.

Response: As part of shortening and restructuring the Introduction, we have removed this sentence.

Line 190; suggestion to add; T cell markers expressed by SARS-CoV-2-specific T cells

Response: We see there is some confusion between the use of “marker” in the sense of surface markers expressed by T cells used to define T cell subsets vs. the statistical sense of “marker” as a measured immune response that is assessed, for example, as a correlate of risk. Section 5 of the SAP defines “T cell markers included in analyses” as percentages of CD4⁺/CD8⁺ T cells expressing certain cytokine combinations, and functionality scores derived from COMPASS. However, given that we now have revised “cytokine” to “cytokine/marker” (in response to the comment on lines 330-344), we have elected to use “response” instead of “marker” in the statistical sense throughout the manuscript.

The revised (underlined) sentence now reads:

“Here, we analyze Spike-specific and Nucleocapsid-specific T-cell immune responses, measured using a 27-color intracellular cytokine staining (ICS) assay developed and validated for SARS-CoV-2 specific CD4⁺ T-cell and CD8⁺ T-cell responses, as correlates of risk and of protection against COVID-19 in COVAIL second boost recipients.” (lines 128-131)

Line 204; please add COVAIL to stages, without this it is not clear what stages mean, also, this can perhaps be considered as methods, since the authors do not explain what the stages mean.

Response: We have added “COVAIL” (line 138).

Regarding explanation of stages in Methods, we feel that the stages need to be introduced before the results, as this knowledge is necessary for results interpretation. We have clarified (underlined) the explanation of the COVAIL study

stages in the Introduction: “Participants were enrolled sequentially in four different stages; within each stage, participants were randomized to receive one of the following study vaccines (i.e. second booster) on Day 1 (D1): stage 1, mRNA-1273 (Moderna); stage 2, BNT162b2 (Pfizer-BioNTech); stage 3, prefusion Spike transmembrane-deleted (pre-S DTM) protein vaccine (Sanofi). Study vaccines included monovalent or bivalent variant vaccines targeting wild type, Beta (B.1.351), Delta (B.1.617.2), Omicron BA.1 (B.1.1.529.1), or Omicron BA.4/BA.5 (B.1.1.529.4/B.1.1.529.5) Spike protein (Supplementary Fig. 1).” (lines 110-116)

Note we also now include Supplementary Fig. 1, which provides a visualization of the COVAIL study stages and specific study arm vaccines.

Lines 218-223 can be shortened and perhaps included in lines 211-212 since almost all individuals included also had T cell data.

Response: We have shortened this section, and have combined the previous three Results sections on analysis cohorts; available T-cell response data; and COVID-19 endpoint definition, case, and non-case definition into a single Results section “Analysis cohorts, available data, and COVID-19 endpoint” (lines 137-163), endeavoring to keep in this section a shortened summary of what is needed for interpreting the results that follow.

Lines 225-239 are all not results but part of a method section. Please adjust.

Response: In response to reviewer #2’s comment on defining “case” and “non-case” at their first mention (their comment on Line 213-214), we have elected to keep the definitions of case and non-case early in the text rather than moving these definitions to Methods (which is at the end of the manuscript). We also keep a brief definition of the COVID-19 endpoint in response to this reviewer’s comment to clarify “illness” (comment on Line 128; please specify “illnesses”). However, we take the point this portion of Results could be condensed. Please see the response to the point above, where we have revised three Results sections into one and substantially shortened.

Lines 307-311, Please clarify the hypothesis on “cross-reactive” (which are actually conserved-region specific) T cells in line with these results.

Response: In the process of shortening the manuscript, we have removed this sentence.

Lines 320-328; These comparisons are not shown in a figure, and this text could be shortened with the use of a figure.

Response: The referenced text (lines 320-328 in the original manuscript) described the distributions of D15 Spike-specific IFN- γ and/or IL-2 producing CD4+ and CD8+ T cells and D15 positive vaccine response proportions in the naïve and in the non-naïve cohorts. These data are shown in Figures 1 and 2. The numbers quoted in the text in this section correspond to the GMs and positive vaccine response proportions that are overlaid on top of the box plots in Figures 1 and 2, and the relevant figure panels are cited in the text (Fig. 1a, Fig. 2a, Fig. 1b, Fig. 2b).

We have shortened the reporting of these values in the text by making the text more concise, resulting in a reduction of approximately 20%.

Lines 355-356; the authors mean the antigen-specific for primary. Please adjust, it now reads as non-specific responses.

Response: Please see our response to the comment above “Line 190; suggestion to add; T cell markers expressed by SARS-CoV-2-specific T cells”. This is another example where “marker” is used in the statistical sense, not the T-cell phenotypic sense.

The statistical primary and secondary markers are defined in section 5 of the SAP. However, to avoid confusion between these two different senses of the word “marker”, we use “response” in the manuscript now. In the caption of Table 1 we now clarify: “Rows for primary responses (also referred to as primary markers, section 5 in the SAP) are shaded.”

Line 398; Please rephrase “marker”

Response: Please see the response above.

Line 508; CD154 is not a cytokine, please adjust.

Response: We revised to “cytokines/markers”.

Line 858; CXCR5 and PD1 are not cytokines. Please adjust

Response: Revised to “cytokines/markers”.

Reviewer #2 (Remarks to the Author):

The manuscript “T-cell immune correlates in COVAIL trial participants by prior SARS-CoV-2 infection status” by Zhang et al. investigated the correlation of T cell immune

responses with risk or protection of COVID-19 in COVAIL trial participants following a second variant-adapted vaccine booster. In this study the CD4 and CD8 T cell responses against spike and N peptides in participants were assessed by using a comprehensive 27-colour intracellular cytokine staining assay at two time points, the same day prior to (D1) and 14 days (D15) after a second mono- or bivalent mRNA or protein-based booster vaccine. The correlation of T cell immune index at D1 or D15 with COVID-19 risk was further on analyzed in a 6-month follow-up. The main finding of this study was that CD4 spike T cell IFN- γ and/or IL-2 spike BA.4/5 responses correlated with low COVID-19 risk in participants with prior infection (non-naïve) but not in naïve participants.

Strength: 1. The topic the study address is interesting and also important. So far, the correlation of T cell immunity established from previous infection as well as vaccination with protection of SARS-CoV-2 infection is less investigated compared to humoral immunity. This information is very important for public health and vaccination strategies.

2. The study is based on a randomized clinical trial with a great number of participants (n=932) and the authors utilized a comprehensive intercellular cytokine staining assay with 27-markers to analyze T cell immunity against spike and nucleocapsid.

Furthermore, the authors performed comprehensive analysis of this large dataset and also a developed machine learning model Superlearner to assess T cell immune correlates with risk of COVID-19.

Major concerns

1. For non-naïve participants, non-spike specific T cells play an important role in the control of viral infections. Lang-Meli et al. 2022 (<https://doi.org/10.1038/s41564-022-01106-y>) showed that spike-derived CD8 T cell epitopes were not dominantly targeted in individuals with infection, non-structural proteins such as ORF1ab contained the main targeted antigens. This implied that non-spike specific T cells, especially CD8 T cells, play an important role in non-naïve individuals. Therefore, spike-specific CD8 T cell immunity might not be the strong immune correlates of protection in this group. And the design of this study might not be suitable to find the optimal immune correlates in non-naïve individuals.

On the other hand, Lang-Meli et al. 2022 also demonstrated that the breadth of spike-derived CD4 T cell epitopes targeted in convalescent individuals is broader than that in 2x vaccinated individuals. Moreover, booster vaccination increased the breadth of the spike-specific CD4 and CD8 T-cell responses in convalescent individuals but not in individuals with two vaccinations. These results are consistent with the main finding of this current manuscript that the CD4-specific T cells prior to the second booster vaccine correlate with low risk of COVID-19 infection in non-naïve group but not in the naïve group.

Since the COVAIL trial was designed to focus on analysis of T cell immune response against mainly Spike and N, the authors should describe these and emphasize that Spike and N-specific T cells alone could not represent a comprehensive immune landscape in non-naïve individuals.

Response: Thank you for this comment. We have made the following revisions (underlined) to our Discussion:

“Fifth, assessing additional antigens or overlapping peptide pools beyond Spike and Nucleocapsid would be required to capture a more comprehensive T-cell response for analysis. This limitation is particularly applicable to the analyses in non-naïve individuals, given that non-structural proteins such as ORF1ab (not structural proteins such as Spike) were previously shown to include the main CD8+ T-cell targeted antigens in previously-infected individuals.¹⁹” (lines 529-534)

2. The individuals in each stage of the COVAIL trial showed considerable heterogeneity the viral insert in the booster vaccine. This may negatively affect the power of correlation analysis when they were grouped.

Response: We balanced the need for a sufficient number of COVID-19 endpoints to give good precision in the correlates analyses with the advantage of reducing heterogeneity as much as possible in the study vaccine insert by focusing the correlates analyses on the one-dose mRNA group: “For consistency with our previous neutralizing antibody correlates analyses of COVAIL,¹⁸ the correlates analyses focus on the combined group of all one-dose mRNA vaccine boosters across stages 1 and 2 (hereafter, the “one-dose mRNA group”), due to their similar vaccine platforms and the fact that they had the majority of COVID-19 endpoints (in the Per-Protocol Immunogenicity Analysis Set: 122 naïve and 22 non-naïve cases in stages 1 and 2 vs. 13 naïve and 4 non-naïve cases in stage 3, Supplementary Fig. 2).” (Supplementary Methods).

However, we did previously report a protective advantage of one-dose mRNA second booster recipients in stage 2 (Pfizer-BioNTech) who received an Omicron-containing study vaccine vs. the wild type study vaccine.^{18,20} Thus, if sufficient breakthrough COVID-19 endpoints had occurred in each individual vaccine arm to perform immune correlates analyses separately by arm, we may expect to see different inverse correlates of risk across the different arms, or if not across arms, at least across logical groups such as mRNA wild type, mRNA Omicron-containing, mRNA bivalent, and mRNA monovalent. The SAP outlines these last four groups as groups of interest in which the correlates analyses may be conducted, given time (Section 2: “Vaccine groups of interest for studying correlates of risk and controlled risk correlates of protection”).

Given the need to shorten the manuscript and the many reviewer comments to make the results more concise and compact, we elected not to introduce additional correlates analyses by these viral insert groups, and have added this as a limitation in the Discussion:

“Second, the study included heterogeneous subgroups of participants with different infection histories (variant, timing), different primary series and boosters received, etc., and the extent to which this heterogeneity may have confounded the results is unknown.” (lines 523-526)

3. The authors performed a variety of correlation assay and thereby provide a large dataset with analyses and figures. However, it is difficult to follow the authors' narrative through the manuscript. It seems that a clear and biologically logic line is missing. The authors may polish the manuscript and make it more readable and concise.

Response: We have substantially shortened the manuscript, making all results sections more concise and polished. In this process we also restructured each results section so that the major result leads the section. We have also restructured the Discussion so that it begins with a summary of the five major conclusions of this manuscript.

4. Given the complex immune landscape against SARS-CoV-2 at the time of conduction this trial, the study includes heterogeneous subgroups of participants with different infection histories, different vaccination schemes etc. This was not easy to avoid at that time, and the authors try to address this in their subgroupings and multivariate analyses. However, this should be considered a limitation of this study and be emphasized as such.

Response: We have added this to our “Limitations” paragraph in the discussion:

“Second, the study included heterogeneous subgroups of participants with different infection histories (variant, timing), different primary series and boosters received, etc., and the extent to which this heterogeneity may have confounded the results is unknown.” (lines 523-526)

Minor concerns

1. 27-colour ICS method were utilized to analyze functional T cells expressing cytokines. As expected the percentages of cytokine-expressing T cells, especially CD4 T cells, in PBMC of participants are in the lower range with the majority between 0.05-1%.

The reliability of specific responses below 0.1% is limited. A more detailed description of the ICS method will be helpful to show the reliability.

Response: We have added substantially more detail to the ICS methods section (lines 591-686 of the revision).

On positivity calls, we include the following details:

“At D1, a responder was defined by a positivity call made by comparing the percentage of T cells secreting the given set of cytokines in the relevant sample under stimulation compared to that in the paired control (sample from the same person-time-point in an unstimulated condition). Those with Fisher’s exact test $P < 0.0001$ were defined as responders. At D15, a positive vaccine-response for percentage of CD4+ T cells expressing a given set of cytokines (“booster response over baseline”) was defined by the worst-case p-value comparing the percentage of positive cells post-vaccination versus baseline being significant when controlling the false discovery rate at 5%, considering all assay misclassification rates that are consistent with the paired control samples at baseline and post-vaccination.²¹ A positive vaccine-response against Spike BA.4/5 peptides was defined as a positive vaccine-response against Spike BA.4/5 subunit S1 or subunit S2 peptides.” (lines 675-686)

2. In the manuscript long sentences were often used which were not easy to follow. For example, line 120-124. It could help to shorten them.

Response: We have shortened this sentence and also reviewed the rest of the manuscript to shorten particularly long sentences where possible.

3. The magnitude of the T-cell response prior to the booster vaccination might be a proxy marker of recent infection, which could explain the protective correlations of this biomarker against SARS-CoV-2 infection in the follow up period. Is there more detailed data about the timepoint of prior infections before the booster vaccination (i.e. timepoints of infections before the D1 timepoint)? This could be used as a separate parameter if available.

Response: Thank you for the suggestion. We do have some data on pre-study infection time among non-naïve participants, with the caveats that this data is incomplete (only available for ~75% of all non-naïve participants) and is self-reported data. As suggested, we explored whether the magnitude of the Nucleocapsid-specific T-cell response could potentially be used as a proxy marker for prior infection recency by plotting pre-study infection time versus D1 CD4+ Nucleocapsid-specific IFN-g and/or IL-2 response magnitude, among non-naïve participants (Figure R1, below). We saw no evidence to support that Nucleocapsid-

specific T-cell response could be used as a proxy for prior infection recency (correlation -0.1, $p=0.91$). We also saw no difference in pre-study infection time between cases and non-cases ($p=0.15$ when comparing all cases vs. non-cases and $p=0.17$ when comparing booster-proximal cases vs. non-cases).

Given these findings and the limitations of these data as discussed above, as well as other reviewer comments to shorten and streamline the revision, we have opted not to include this plot in the revision.

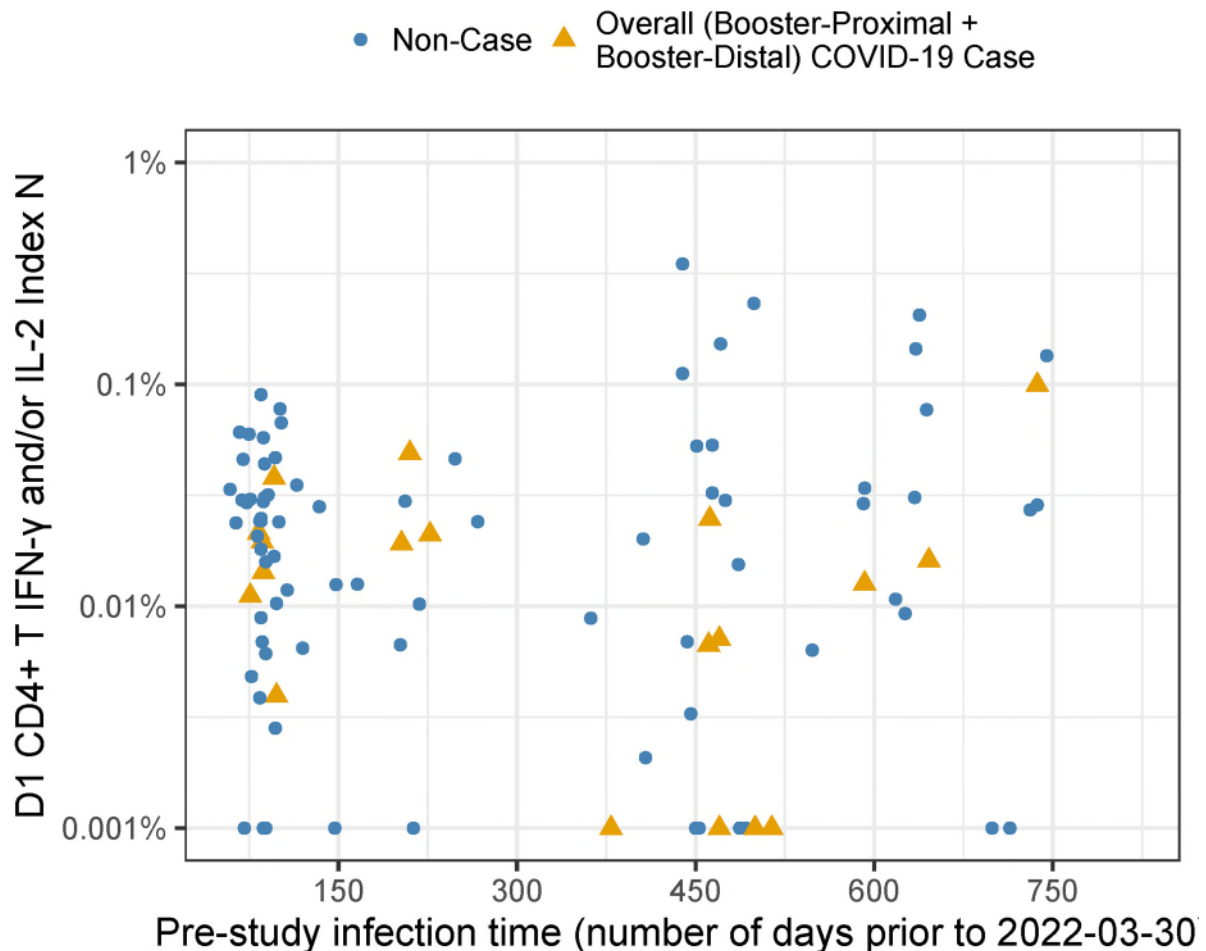


Figure R1. Scatterplot of pre-study infection time (number of days prior to March 30, 2022) versus D1 Nucleocapsid-specific CD4+ T-cell IFN- γ and/or IL-2 response magnitude. Data points are from one-dose mRNA recipients in the non-naïve cohort. Pre-study infection times were self-reported and were available for ~75% of all one-dose mRNA recipients in the non-naïve cohort of the Per-Protocol Immunogenicity Analysis Set. Blue dots represent non-cases and orange triangles represent COVID-19 cases. Correlation for the scatterplot was -0.1, with $p=0.91$.

Abstract

- Line 93: greater -> stronger

Response: In the process of shortening the Abstract to reach the journal's word limit, this phrase was removed.

- Line 97: Cox modeling of CD4+ T cells -> spike CD4+ T cells

Response: We added "Spike-specific CD4+ T cells".

- Line 101: titers-> neutralizing antibody titers

Response: Revised.

- Line 106: CoP -> correlates of protection (CoP)

Response: We wrote this out as suggested.

Introduction

- Line 127: would asymptomatic participants with COVID-19 breakthrough infection be identified? What was the percentage of non-naïve individuals without positive anti-N response?

Response:

1. In the present work, in accordance with Zhang et al.¹⁸ we considered only symptomatic SARS-CoV-2 infections. We now state this more clearly in the manuscript and include a supplementary table detailing the distributions of symptoms among the 161 cases in the Per-Protocol Immunogenicity Analysis Set, separately for naïve and non-naïve participants:

"The COVID-19 endpoint was symptomatic SARS-CoV-2 infection, as examined previously.¹⁸" (line 152)

"In each stage, COVID-19 cases were caused predominantly by BA.5 (lineage distributions provided in the Supplementary Text and symptom distributions in Supplementary Table 1)." (lines 159-160)

2. On the percentage of non-naïve individuals without positive anti-N response, only a small proportion (~5%) of participants were classified as non-naïve based on self-report, not a positive anti-Nucleocapsid test result at baseline. We have added the following to the revision:

“Of the 387 (68.0%) naïve participants in the PP-IAS, 183 were female (mean age 51.6 years) and 204 were male (mean age 52.2 years). The remaining 182 (32.0%) were non-naïve (73 female, mean age 48.1 years and 109 male, mean age 46.7 years), almost all of whom (172/182, 94.5%) had a positive anti-Nucleocapsid test result at baseline. The remaining 10 non-naïve participants (5.5%) tested anti-Nucleocapsid negative at baseline and were classified as non-naïve due to self-reporting previous SARS-CoV-2 infection.” (lines 146-151)

- Line 129: would SARS-CoV-2 variants in all infected participants with symptoms be characterized?

Response: We have added the following to the revision:

“In each stage, COVID-19 cases were caused predominantly by BA.5 (lineage distributions provided in the Supplementary Text...” (lines 159-160)

In the Supplementary Text (p. 64), we now detail:

“Lineages of COVID-19 cases

Distributions of lineages among the n=161 COVID-19 cases in the Per-Protocol Immunogenicity Analysis Set, listed by stage, were as follows:

Stage 1: BA.2, n=32; BA.4, n=11; BA.5, n=56; XZ, n=1.

Stage 2: BA.2, n=5; BA.4, n=4; BA.5, n=35.

Stage 3: BA.4, n=1; BA.5, n=16.

The majority of the lineages above (121/161) were observed via viral sequencing of the nasal swab sample collected in persons testing positive. The rest were imputed by using the same method as described in Zhang et al.¹⁸”

Results

- Line 209: please provide a clear schema of COVAIL trial (combining the Fig.S1 and Table 1 in SAP) similar as supplementary Figure S1 for from Zhang et al. 2025. Explain all the groups of vaccine boost arm as mentioned in section 2 of Statistical Analysis Plan (SAP) in the supplementary file.

Response: We have added the figure suggested by the reviewer as Supplementary Figure 1. The study arm groups considered in the analyses presented in the manuscript are denoted by shading and outlining.

- Line 213-214: non-case and case were mentioned for first time in this part, explain them as participants without infection and infections, respectively. The detail

descriptions were in

Response: We have moved the definitions of “case” and “non-case” to accompany their first mention in the manuscript:

“...Cases were endpoints in PP-IAS participants stratified in the three time periods above. Non-cases were PP-IAS participants with no evidence of SARS-CoV-2 infection after D1 and through to 188 days post-D15.” (lines 154-156)

- Line 551-552: EM -> Tem, CM -> Tcm.

Response: In the process of shortening the manuscript, we no longer use these abbreviations in the text. We have, however, added these abbreviations in the caption of Supplementary Figs. 26 and 27.

Methods:

- L857-859: what is the reason for changing 3 antibodies? Please describe the ICS FACS protocol in more details, for example, the amount of PBMCs analyzed, the FACS gating strategy

Response: In the previously validated COVID v1 panel, CRTh2 was included as a surface marker for Th2 cells and CD32 and CD64, as Fcγ receptor markers, were included related to the potential for infection or vaccination to induce antibody-mediated enhancement of disease and/or infection. It was found that these markers did not add significant value to the information provided by the panel. Since CRTh2 likely identifies additional Th2 cells and not just those induced by infection or vaccination, it was not critical to the primary purpose of the assay. CD32 and CD64 are likely more applicable to non-T cells and are therefore not of primary interest to this assay. Due to increased interest in follicular helper T cells (Tfh) cells and the role they play in protective immunity, it was decided to replace CRTh2, CD32, and CD64 with the Tfh identifying markers CXCR5, PD1, and IL-21. Tfh cells constitutively express CXCR5 and express PD1 at high levels. IL-21 is an important cytokine produced by Tfh cells since production of IL-21 by Tfh cells drives the differentiation of B cells and affinity maturation of antibodies. In the modification of the original version of the panel, the new markers were inserted in the positions vacated by the removal of CRTh2, CD32, and CD64. CXCR5 is included with the stimulation cocktail rather than included with either the surface or intracellular antibody staining cocktails. CXCR5 is downregulated with stimulation and loss of IL-21 co-expression with CXCR5 is also observed. By staining for this marker during stimulation, we maintain the detection of the IL-21 producing Tfh cells.

Substantially more detail has been added to the ICS assay methods section (lines 591-673), including the addition of Supplementary Fig. 32, which provides the gating strategy. In lines 642-644 we now state “Collection of the cells on the

cytometer resulted in approximately 200K-500K total events collected. Actual numbers of CD4+ and CD8+ T cells analyzed were participant- and stimulation-dependent, with some stimulations causing more toxicity.” The captions of Supplementary Figs. 25 and 26 include the numbers of CD4+ T cells analyzed.

Supplementary info.

1. On Index page, the Figure S5 title is missing

Response: Thank you, fixed.

2. Figure S17 legend, (B,D, D-F), (B,D) are redundant.

Response: We carefully reviewed the legends of the supplementary controlled risk figures and corrected any existing typos (we found just one).

Reviewer #3 (Remarks to the Author):

The aim of the study by Zhang et al. was to investigate immune correlates for COVID-19 risk or protection through SARS-CoV-2-specific T-cell responses after vaccination, including in relation to the prevailing variants of concern (VoCs) and variant-adapted boosters.

Overall, this is a relatively large study involving 970 recipients of a single-dose vaccination in phases 1, 2, and 3, with (polyfunctional) T cell determination performed only for a subset. Participants were randomly assigned to one of four phases (no T cell detection in phase 4) and received a second booster vaccination (day 1, D1), with immunological parameters measured again 14 days after the second booster vaccination (D15).

Analysis of CD4+ and CD8+ T-cell responses against spike (S) and nucleocapsid (N) peptides was performed using a 27-color intracellular cytokine staining (ICS) assay in vaccinated subjects from the COVAIL (COVID-19 Variant Immunologic Landscape) study over an extended period of time.

The study was not designed to evaluate the relative effectiveness of booster shots in preventing COVID-19, but immunological correlates for risk assessment could be found based on the COVID-19 symptoms reported by participants. From the results, the authors conclude that spike CD4+ T cells are a non-mechanistic CoP that serves as a surrogate for one or more unmeasured protective immune mechanisms that are only effective after prior SARS-CoV-2 infection.

The results of the study showed that while neutralizing antibodies are the dominant protective factor, CD4+ T-cell responses provide important complementary protection, especially in individuals with lower antibody concentrations. This suggests that both humoral and cellular immunity should be considered in vaccine evaluation and risk stratification strategies.

Overall, this is an extremely comprehensive study involving a very large cohort, data analyses at two different time points after booster vaccinations with different booster vaccines, etc.

Nevertheless, I find the truly novel insights gained from this study to be limited. The fact that memory T-cell responses can be enhanced by vaccination or infection is not new, and neither is the data showing the synergistic protective effect of the combination of neutralizing antibodies and CD4+ T-cell responses.

Major Concerns:

Most results are well documented, but are described with long insertions, which impairs readability. A summary of the areas and highlighting the direction and strength of the observed results would provide greater clarity and improve readability.

Response: We have substantially shortened our manuscript, and the length of the text (minus Methods) has been reduced from 9294 to 5563 words. During the revision process, we strove to increase readability by moving several portions of text to Methods or to figure captions. On summarizing the findings of this work, we have restructured our Discussion so that it now begins with an overview of the five main conclusions we draw from this work. We then discuss each of these five conclusions in turn.

The definitions of endpoints and cases/non-cases are precise and consistent with previous COVAIL publications and CDC criteria, but the repeated references to previous articles, etc., make the text cumbersome.

Response: We have substantially shortened the text in this area and avoided repeated references to previous articles.

The use of overlapping peptide pools generally has the advantage of being “HLA-independent.” Nevertheless, the peptides must of course be presented in HLA molecules. How can the authors be sure that peptides with sequence differences between the individual variants are actually presented (and not just conserved regions) and that the T-cell responses are then actually specific to these differences? What is the proportion of “non-matching” peptides in the pools and for the comparison of variants in the boosters? The authors should at least discuss this.

Response: Thank you for the question. First, we note that the use of 15-mers with 11 amino acid overlap is a well-accepted procedure for generating peptides in ICS studies aimed at assessing cellular responses.²² Moreover, we expect that there is a diversity of epitopes within the peptide pools that would likely be presented across various HLA types.

As far as how we can ensure that peptides with sequence differences between individual variants are presented and that the T-cell responses are specific to these differences, we do not have data on this and cannot answer this question. While we did address the objective of characterizing immunogenicity for each of the three vaccine manufacturers across Spike vaccine inserts, similar to Branche et al.,¹² no prespecified hypothesis tests were planned on this and the comparisons were descriptive and not statistical. Our comparisons of T-cell responses across vaccine arms were limited to within-stage comparisons (necessary due to randomization only being conducted within-stage) and consisted of the following:

“D1 Spike-specific T-cell response magnitudes and proportions with positive response were largely similar across booster arms within each stage, as expected due to the within-stage randomization (Fig. 1A, Fig. 2A).” (lines 175-177)

“We focus the post-second booster comparisons on within-stage comparisons, given that randomization was conducted within stages. Within each stage, post-booster antigen-specific T cell responses were generally comparable at D15, regardless of prior infection status (Figs. 1 and 2). However, naïve recipients of a wild type mRNA vaccine had lower D15 Spike-specific CD4+ T-cell responses compared to naïve recipients of an Omicron-containing mRNA vaccine (Spike BA.4/5-specific IFN-γ and/or IL-2 GMs in stage 1: wild type 0.17%, Omicron 0.23% to 0.27%; in stage 2: wild type 0.12%, Omicron BA.1 0.16% to 0.25% (Fig. 1A). This finding is consistent with the previous neutralizing antibody results¹² and was not seen in the non-naïve cohort (Fig. 2a) or in CD8+ T cells (Fig. 1b, Fig. 2b).” (lines 190-198)

We have added to the Discussion:

“Moreover, from the overlapping peptide pools strategy we cannot be certain that peptides with sequence differences across the study vaccine variants are actually presented (not just conserved regions) in HLA molecules.” (lines 538-540)

The nature of overlapping pools of 15-mer peptides generally induces stronger CD4 than CD8 T cell responses. How can the authors ensure that CD8 responses are not overlooked solely because peptide presentation in HLA class I is not optimal?

Response: Thank you for raising this point. We agree that we cannot be certain that some CD8+ T-cell responses are not missed due to the use of 15-mers and the potentially sub-optimal peptide presentation in HLA class I. However, even if some CD8+ T-cell responses are missed, this would not be anticipated to have any effect on the results of our immune correlates analyses, given that the same CD8+ T-cell responses would be missed in both cases and non-cases.

Although this study provides valuable insights into the relationship between CD4+ T cell responses and vaccine efficacy, several methodological limitations and interpretive challenges must be considered. The findings are largely consistent with the current understanding of the immunological correlates of protection, but the evidence for T cell markers as independent correlates is weaker than suggested by the authors. The study would benefit from more rigorous statistical validation and clearer communication of the uncertainty in the results.

Response: Thank you for these good points. Regarding clearer communication of uncertainty, we have made the following revisions:

First, related to the reviewer's comment that is four comments down from this one, we have removed all instances of the word "clearly". For instance, in the "D1 and D15 Spike-specific and D1 Nucleocapsid-specific CD4+ T-cell responses are inverse correlates of COVID-19 risk among the non-naïve cohort" section of Results, we now state "Univariable correlate of risk analyses in the non-naïve cohort of the one-dose mRNA group showed that Spike-specific and Nucleocapsid-specific CD4+ T-cell responses correlated inversely with COVID-19 risk over ~3 months follow-up, as well as over ~6 months follow-up." (lines 285-288). Second, for all HRs (or HR ranges) provided in the same Results section, we made sure to provide the corresponding p-value (or p-value ranges) in the text, so that the reader can easily locate this information. We also added "...albeit with some of these p-values close to conventional thresholds of significance" (line 295) and "A similar trend was seen for functionality score...." (line 298).

We additionally now state "Similar results, albeit with lower precision, were also obtained in stage 1 and stage 2 separately (Supplementary Fig. 17)." (lines 310-311) In our Limitations paragraph in the Discussion, we state "Fourth, there were not enough breakthrough COVID-19 endpoints in Sanofi recombinant protein booster recipients to be able to assess, with acceptable precision, T-cell correlates for this vaccine platform." (lines 527-529)

The key findings that higher S-specific CD4⁺ than CD8⁺ T-cell responses, virtually no N-specific responses in naïve participants, and higher baseline responses in non-naïve

participants were to be expected and are not new, but well documented. But what about cross-reactivity to other coronaviruses, which has also been demonstrated for N-specific T cells in SAR-CoV-2-naïve individuals? And again, the question arises: are CD8+ T cell responses being overlooked?

Response: Thank you for bringing to our attention the issue of cross-reactivity to other coronaviruses, which we agree is important to consider. However, we see little evidence of cross-reactive Nucleocapsid-specific CD4+ or CD8+ T cells in the SARS-CoV-2 naïve individuals in this study. For example, in the naïve cohort, the proportion of pre-second booster (D1) CD4+ T cells with a positive response to IFN- γ and/or IL-2 in response to Nucleocapsid was 2.0% (0.7%, 5.2%); for CD8+ T cells, this was 0.5% (0.1%, 3.4%) (Table 1).

We have added to the revision: “While cross-reactive CD8+ T cells targeting a region of N that is highly conserved between SARS-CoV-2 and other human coronaviruses have been reported,²³⁻²⁵ there was little evidence of such cross-reactivity, with a D1 N-specific CD8+ T-cell positive response proportion of 0.5% (0.1%, 3.4%) in naïve non-cases (Table 1).” (lines 234-237)

In addition, to address the first point about lack of novelty/expected results, we have substantially shortened the text reporting the immunogenicity results (from 1934 words to 1023 words).

The conclusion that on day 15 (D15) the CD4⁺ S-specific T-cell responses between naïve and non-naïve cohorts are “balanced” after the booster vaccination is justified and expected for CD4⁺, but the persistent CD8⁺ differences and p-values are only briefly mentioned. A more explicit contrast would better support the mechanistic interpretation of the different priming responses of CD4⁺ and CD8⁺ compartments, if possible.

Response: We have revised the manuscript to give approximately equal weight to the D15 CD4+ and CD8+ immunogenicity findings. Revisions include:

-Including the CD8+ finding in the Results section header: “In the one-dose mRNA group, post-second booster equalization of CD4+ T-cell levels, but not CD8+ T-cell levels, between non-naïve and naïve participants” (lines 224-225)

-Quoting in the text the D15 GMs and positive vaccine-response proportions for CD8+ T cells, for parallel structure with the CD4+ T cell reporting in this section:

“After the second booster, CD4+ IFN- γ and/or IL-2 Spike BA.4/5 response levels became comparable across both cohorts, with D15 GMs of 0.214% (0.190%, 0.240%) in naïve vs. 0.216% (0.185%, 0.253%) in non-naïve non-cases (p=0.978) (Table 1). However, CD8+ IFN- γ and/or IL-2 Spike BA.4/5 D15 GMs remained

significantly higher in the non-naïve cohort [naïve 0.106% (0.082%, 0.138%) vs. non-naïve 0.165% (0.114%, 0.239%) non-cases, $p=0.002$]. These results suggest that the second booster equalized CD4+, but not CD8+, T-cell response levels between the naïve and non-naïve cohorts.” (lines 238-244)

-We also now draw out the difference in vaccine-response proportions, which, although no statistical comparison testing was done, were descriptively higher in the naïve vs. the non-naïve cohort for CD4+ but not CD8+ T cells (i.e. drawing an additional contrast between CD4+ and CD8+ T cells pertaining to D15 immunogenicity results):

“This difference in booster equalization was mirrored by a descriptive difference in positive vaccine-response proportions. D15 CD4+ positive vaccine-response proportions were higher in the naïve compared to the non-naïve cohort: 57.1% (50.2%, 63.8%) in naïve vs. 42.8% (34.1%, 51.9%) in non-naïve non-cases, while CD8+ positive vaccine-response proportions were similar in the two cohorts: 47.9% (41.1%, 54.8%) and 52.9% (43.7%, 61.9%) in naïve and non-naïve non-cases, respectively.” (lines 244-249)

In the Discussion, we now include the CD8+ D15 immunogenicity results:

“(1) receipt of a second booster equalized antigen-specific CD4+ T-cell levels (but not CD8+ T-cell levels) between the naïve and non-naïve cohorts” (lines 438-440).

Regarding a mechanistic hypothesis, we have added to our Discussion: “The first conclusion is that after the second booster (i.e. on D15), circulating Spike-specific CD4+ responses were similar in naïve and non-naïve participants, whereas CD8+ responses remained higher in non-naïve participants. This pattern is consistent with a more persistent influence of prior infection on circulating Spike-specific CD8+ than CD4+ T-cell responses. Because our PBMC-based ICS assay measured circulating peptide-responsive cells but did not resolve epitope-level breadth, TCR repertoire/clonality, direct cytolytic capacity, or tissue-resident memory, these findings should be interpreted as hypothesis-generating rather than mechanistic.” (lines 449-456)

High correlations between spike peptide pools and time points and weak spike, N, and T cell nAb correlations are clearly described, but here too, independence of humoral and cellular responses was to be expected, as described for both naïve and vaccinated subjects. The authors showed, that combination of high booster-elicited neutralizing antibodies and high pre-existing S-specific CD4+ T-cell responses from prior infection resulted in very low COVID-19 risk. It is not clear whether this is also due to a broader T-cell repertoire and the presence of T-cells against the N and M protein generated during

infection. If necessary, a statement on the possible persistence of the Ak and different T-cells against different viral antigens would be helpful here.

Response: We agree that the assessment of T-cell responses specific to Spike and Nucleocapsid alone is a limitation of this work. We did not assess any Membrane-specific T-cell responses and thus cannot address any hypothesis pertaining to the role of Membrane-specific T cells generated by prior infection.

We have added the following to our Discussion:

“Fifth, assessing additional antigens or overlapping peptide pools beyond Spike and Nucleocapsid would be required to capture a more comprehensive T-cell response for analysis. This limitation is particularly applicable to the analyses in non-naïve individuals, given that non-structural proteins such as ORF1ab harbor the main CD8+ T-cell targeted antigens in previously-infected individuals.¹⁹ It also limits interpretation of the result of high booster-elicited neutralizing antibodies and high pre-existing Spike-specific CD4+ T-cell responses from prior infection resulting in very low COVID-19 risk, given that this very low risk may be due to a broader T-cell repertoire and/or the presence of T cells specific to other SARS-CoV-2 proteins (e.g. Membrane).” (lines 529-538)

The Cox and controlled risk analyses show convincing inverse CD4⁺ correlates in the non-naïve cohort, but not in the naïve cohort, but the repeated emphasis on “clearly indicated inverse correlations” could be tempered by acknowledging the moderate HRs, some p-values close to conventional thresholds, and limited case numbers, particularly in the subgroup and functionality score analyses.

Response: Please see our response to your earlier comment on better communication regarding uncertainty:

We have removed all instances of the word “clearly” and now state “Univariable correlate of risk analyses in the non-naïve cohort of the one-dose mRNA group showed that Spike-specific and Nucleocapsid-specific CD4+ T-cell responses correlated inversely with COVID-19 risk over ~3 months follow-up, as well as over ~6 months follow-up.” (lines 285-288). Regarding p-values, for all HRs (or HR ranges) provided in the section “Spike-specific and Nucleocapsid-specific CD4+ T-cell responses are inverse correlates of COVID-19 risk among the non-naïve cohort”, we made sure to provide the corresponding p-value (or p-value ranges) in the text, so that the reader can easily locate this information. We also added “...albeit with some of these p-values close to conventional thresholds of significance” (line 295) and “A similar trend was seen for functionality score....” (line 298).

Regarding case numbers, we now state “Similar results, albeit with lower precision, were also obtained in stage 1 and stage 2 separately (Supplementary Fig. 17).” (lines 310-311)

In our Limitations paragraph in the Discussion, we state “Fourth, there were not enough breakthrough COVID-19 endpoints in Sanofi recombinant protein booster recipients to be able to assess, with acceptable precision, T-cell correlates for this vaccine platform.” (lines 527-529)

The COMPASS-based findings on polyfunctionality and Leiden clustering are interesting, well presented, and consistent with the most important CD4⁺ T-cell signal (EM/CM CD4⁺ subsets that co-express CD154, IFN- γ , IL-2, and TNF- α and are enriched in non-naive cases), but the description is very technical and contains many “marker” lists. Here, it could be summarized more quickly and concisely that polyfunctional effector memory CD4⁺ T cells are associated with protection in non-naive participants—which is, after all, the core message.

Response: We revised the section header to communicate the core message (“Polyfunctional effector memory CD4⁺ T cells are associated with protection in non-naive participants”). We also removed the list of specific cytokine/marker combinations defining subsets that tended to be higher in non-cases vs. booster-proximal cases (“including CD154 and TNF- α ; CD154, IFN- γ , and TNF- α ; CD154, IL-2, and TNF; and CD154, IFN- γ , IL-2, and TNF- α ” was removed). We also shortened the Leiden clustering text and added “...suggesting that polyfunctional effector memory CD4⁺ T cells are associated with protection in non-naive participants.” (lines 349-351)

The exploratory nature of the clustering analysis is only mentioned implicitly; it would strengthen the section to explicitly state that these are hypothesis-generating results to avoid overinterpretation of specific clusters.

Response: We appreciate the comment and have added “Given the exploratory nature of this analysis, these results should be considered hypothesis-generating.” (lines 351-352)

The definitions of endpoints and cases/non-cases are precise and consistent with previous COVAIL publications and CDC criteria, but the repeated references to previous articles and SAP make the text cumbersome (a concise cross-reference would suffice).

Response: We have substantially shortened the text in this area and avoided repeated references to previous articles.

Minor note:

The manuscript is very long and sometimes difficult to read, and it often takes a long time to find the key message of each chapter. Shortening the text and presenting the most important points more quickly would improve readability.

Response: As discussed above, we have substantially shortened the text and aimed to present the most important take-home conclusion of each Results section first.

Abbreviations are often not introduced

Response:

In the chapter “Machine Learning,” terms such as “baseline adjustment factors” and ‘superlearner’ are not defined, and to my knowledge, “machine superlearning” is not standardized terminology.

Response: We now provide more details, including three references, and a brief explanation of superlearner in Methods:

“Superlearner²⁸⁻³⁰ is a weighted ensemble machine learning approach. The model is based on a superlearner ensemble architecture for binary outcome risk prediction, as prespecified in the Statistical Analysis Plan. The superlearner estimates the best-performing classification model (in our case, classification into individual COVID-19 case versus non-case status) as a weighted combination of candidate learning algorithms selected to minimize cross-validated risk. The analysis was implemented separately by cohort (naïve “vaccine immunity” or non-naïve “hybrid immunity”), outcome definition, and prespecified sets of antibody and T-cell markers at D1 and at D15. The approach does not rely on a single fixed parametric model, but rather adaptively combines multiple learning algorithms to optimize predictive performance.” (line 778-786)

Note we do not use “machine superlearning” anywhere in our manuscript, although we do use “machine learning” and “superlearner” (to refer to the specific machine learning approach). This terminology is consistent with previous work from our group, as well as from other groups.

Example papers from our group using “machine learning” and “superlearner” (sometimes also written as Super Learner or super learner) include:

Gilbert and Luedtke, Statistical Learning Methods to Determine Immune Correlates of Herpes Zoster in Vaccine Efficacy Trials, *The Journal of Infectious*

Diseases, Volume 218, Issue suppl_2, 1 November 2018, Pages S99–S101,

<https://doi.org/10.1093/infdis/jiy421>

Magaret et al. (2019) Prediction of VRC01 neutralization sensitivity by HIV-1 gp160 sequence features. *PLoS Comput Biol* 15(4): e1006952.

<https://doi.org/10.1371/journal.pcbi.1006952>

Benkeser et al., Comparing antibody assays as correlates of protection against COVID-19 in the COVE mRNA-1273 vaccine efficacy trial. *Sci. Transl. Med.* 15,eade9078(2023).DOI:[10.1126/scitranslmed.ade9078](https://doi.org/10.1126/scitranslmed.ade9078)

From other groups:

McCoy, D., Mgbara, W., Horvitz, N. et al. Ensemble machine learning of factors influencing COVID-19 across US counties. *Sci Rep* 11, 11777 (2021).

<https://doi.org/10.1038/s41598-021-90827-x>

Ma, C., Du, J., Lan, L. et al. Immune correlates analysis of mRNA-1345 RSV vaccine efficacy clinical trial. *Nat Commun* 16, 6118 (2025). <https://doi.org/10.1038/s41467-025-61153-x>

Pourakbari, B., Mamishi, S., Valian, S.K. et al. Predicting COVID-19 severity in pediatric patients using machine learning: a comparative analysis of algorithms and ensemble methods. *Sci Rep* 15, 29118 (2025). <https://doi.org/10.1038/s41598-025-15366-1>

Regarding baseline adjustment factors, we have revised Methods so that it now provides more details:

“based on baseline adjustment factors (detailed below)...” (lines 788-789)

“All models included baseline adjustment factors. For the naïve cohort, two different sets of baseline adjustment factors were considered in input variable sets, one of which included baseline risk score, standardized force of infection (FOI) score, vaccine insert, and stage (vaccine manufacturer) information, and the other omitted the insert and stage information. For the non-naïve cohort, which had only a total of 17 cases for analysis, only one set of baseline adjustment factors was considered (baseline risk score only) in all models.” (lines 792-798)

Reviewer #3 (Remarks on code availability):

In my opinion, the results of the work are reproducible. The code has been opened.

Response: Thank you for your time to investigate and confirm the reproducibility of these findings.

Note that we have also added significantly more detail to our Code Availability statement, to further guide the reader on how to run the code used to generate each specific item of this manuscript.

Reviewer #4 (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.

Response: Thank you for your time and consideration of our manuscript.

References

- 1 Gilbert, P. B. *et al.* A Covid-19 Milestone Attained — A Correlate of Protection for Vaccines. *New England Journal of Medicine* **387**, 2203-2206 (2022).
- 2 Khoury, D. S. *et al.* Correlates of Protection, Thresholds of Protection, and Immunobridging among Persons with SARS-CoV-2 Infection. *Emerg Infect Dis* **29**, 381-388, doi:10.3201/eid2902.221422 (2023).
- 3 Hertoghs, N. *et al.* Vaccine-induced T cell responses correlate with reduced risk of severe COVID-19 in a placebo-controlled efficacy trial. *EBioMedicine* **117**, 105809, doi:10.1016/j.ebiom.2025.105809 (2025).
- 4 Kemlin, D. *et al.* Humoral and cellular immune correlates of protection against COVID-19 in kidney transplant recipients. *Am J Transplant* **23**, 649-658, doi:10.1016/j.ajt.2023.02.015 (2023).
- 5 Torres, I. *et al.* SARS-CoV-2 Omicron BA.1 variant breakthrough infections in nursing home residents after an homologous third dose of the Comirnaty(R) COVID-19 vaccine: Looking for correlates of protection. *J Med Virol* **94**, 4216-4223, doi:10.1002/jmv.27867 (2022).
- 6 Brasu, N. *et al.* Memory CD8(+) T cell diversity and B cell responses correlate with protection against SARS-CoV-2 following mRNA vaccination. *Nat Immunol* **23**, 1445-1456, doi:10.1038/s41590-022-01313-z (2022).
- 7 Baum, H. E. *et al.* Hybrid B- and T-Cell Immunity Associates With Protection Against Breakthrough Infection After Severe Acute Respiratory Syndrome Coronavirus 2 Vaccination in Avon Longitudinal Study of Parents and Children (ALSPAC) Participants. *J Infect Dis* **232**, e327-e340, doi:10.1093/infdis/jiaf246 (2025).
- 8 Van Elslande, J. *et al.* Lower persistence of anti-nucleocapsid compared to anti-spike antibodies up to one year after SARS-CoV-2 infection. *Diagn Microbiol Infect Dis* **103**, 115659, doi:10.1016/j.diagmicrobio.2022.115659 (2022).
- 9 Horton, H. *et al.* Optimization and validation of an 8-color intracellular cytokine staining (ICS) assay to quantify antigen-specific T cells induced by vaccination. *J Immunol Methods* **323**, 39-54, doi:10.1016/j.jim.2007.03.002 (2007).

- 10 Cohen, K. W. *et al.* Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep Med* **2**, 100354, doi:10.1016/j.xcrm.2021.100354 (2021).
- 11 Perfetto, S. P., Ambrozak, D., Nguyen, R., Chattopadhyay, P. K. & Roederer, M. Quality assurance for polychromatic flow cytometry using a suite of calibration beads. *Nat Protoc* **7**, 2067-2079, doi:10.1038/nprot.2012.126 (2012).
- 12 Branche, A. R. *et al.* Comparison of bivalent and monovalent SARS-CoV-2 variant vaccines: the phase 2 randomized open-label COVAIL trial. *Nat Med* **29**, 2334-2346, doi:10.1038/s41591-023-02503-4 (2023).
- 13 Lin, L. *et al.* COMPASS identifies T-cell subsets correlated with clinical outcomes. *Nat Biotechnol* **33**, 610-616, doi:10.1038/nbt.3187 (2015).
- 14 Hawkridge, T. *et al.* Safety and immunogenicity of a new tuberculosis vaccine, MVA85A, in healthy adults in South Africa. *J Infect Dis* **198**, 544-552, doi:10.1086/590185 (2008).
- 15 Levine, J. H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184-197, doi:10.1016/j.cell.2015.05.047 (2015).
- 16 Voillet, V. *et al.* Unbiased cell clustering analysis of vaccine-induced T cell responses in the Imbokodo HIV-1 vaccine trial. *EBioMedicine* **122**, 106017, doi:10.1016/j.ebiom.2025.106017 (2025).
- 17 Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* **9**, 5233, doi:10.1038/s41598-019-41695-z (2019).
- 18 Zhang, B. *et al.* Neutralizing antibody immune correlates in COVAIL trial recipients of an mRNA second COVID-19 vaccine boost. *Nat Commun* **16**, 759, doi:10.1038/s41467-025-55931-w (2025).
- 19 Lang-Meli, J. *et al.* SARS-CoV-2-specific T-cell epitope repertoire in convalescent and mRNA-vaccinated individuals. *Nat Microbiol* **7**, 675-679, doi:10.1038/s41564-022-01106-y (2022).
- 20 Diemert, D. J. *et al.* Effect of Omicron BA.1-based compared to prototype booster mRNA vaccination on incidence of COVID-19 in the COVAIL trial. *Vaccine* **64**, 127718, doi:10.1016/j.vaccine.2025.127718 (2025).
- 21 Chen, Z., Heng, S., Tapley, A., De Rosa, S. & Zhang, B. Determining vaccine responders in the presence of baseline immunity using single-cell assays and paired control samples. [Preprint] arXiv:2507.06451 [stat.ME]. doi: 10.48550/arXiv.2507.06451. Submitted on 8 Jul 2025. Access date 10 Jul 2025.
- 22 Xu, Y., Wasnik, S., Baylink, D. J., Berumen, E. C. & Tang, X. Overlapping Peptide Library to Map Qa-1 Epitopes in a Protein. *J Vis Exp*, doi:10.3791/56401 (2017).
- 23 Lineburg, K. E. *et al.* CD8(+) T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope cross-react with selective seasonal coronaviruses. *Immunity* **54**, 1055-1065 e1055, doi:10.1016/j.immuni.2021.04.006 (2021).
- 24 Nguyen, T. H. O. *et al.* CD8(+) T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope display high naive precursor frequency and TCR promiscuity. *Immunity* **54**, 1066-1082 e1065, doi:10.1016/j.immuni.2021.04.009 (2021).

- 25 Murray, S. M. *et al.* The impact of pre-existing cross-reactive immunity on SARS-CoV-2 infection and vaccine responses. *Nat Rev Immunol* **23**, 304-316, doi:10.1038/s41577-022-00809-x (2023).
- 26 Polley, E. C., LeDell, E., Kennedy, C. J. & van der Laan, M. J. SuperLearner: Super Learner Prediction <https://CRAN.R-project.org/package=SuperLearner>, R package version 2.0-28. (2022).
- 27 Polley, E. C. & van der Laan, M. J. Super Learner in Prediction. U.C. Berkeley Division of Biostatistics Working Paper Series. Paper 226. <http://biostats.bepress.com/ucbbiostat/paper266/>. (2010).
- 28 van der Laan, M. J., Polley, E. C. & Hubbard, A. E. Super learner. *Stat Appl Genet Mol Biol* **6**, Article25, doi:10.2202/1544-6115.1309 (2007).

REVIEWER COMMENTS

Reviewer #2 (Remarks to the Author):

All my points have been addressed sufficiently.

Response: Thank you for the positive feedback.

Reviewer #2 (Remarks on code availability):

I have not reviewed the code

Reviewer #3 (Remarks to the Author):

The manuscript has become significantly clearer and more readable as a result of the abridgement and substantial revision.

I do not yet find the data and conclusions regarding peptide presentation and T-cell recognition of the variable regions convincing, especially since simple peptide presentation predictions for selected common HLAs could have provided some guidance here.

Response:

We fully agree that, if participant HLA genotype data were available, such information should be used in concert with epitope mapping to identify which observed T-cell responses may be attributed to relevant peptide-HLA complexes (vs. general immune activation signals). For the T-cell responses that correlated inversely with COVID-19 risk, such information could be used to generate hypotheses on mechanism, and could identify correlates that merit further testing to establish whether they are true mechanistic correlates of protection (mCoPs).¹⁻³

However, in the COVAIL study, participant HLA genotyping was not performed, nor were TCR β -chain repertoires sequenced to enable probabilistic inference of participant HLA genotype. Epitope mapping studies also fall outside the scope of the current study. Moreover, if the value of HLA genotype data and epitope mapping data is to advance consideration of specific T-cell responses that correlated inversely with risk as mCoPs, such additional data would not be sufficient to establish the correlates as mCoPs. For this, an animal model for passive transfer and/or selective depletion would be needed, which is also outside the scope of the current study. This limitation also applies even if peptide presentation predictions for selected common HLAs were performed, as

suggested by the reviewer. Therefore, with the data we have, it is not feasible, nor in the scope of the current study, to generate hypotheses on mechanism involving specific peptide presentation and T-cell recognition.

Considering this, it is worth emphasizing that non-mechanistic CoPs can still be practically useful [Plotkin and Gilbert's Principle #10, "There Are Convenient Nonmechanistic Correlates"³]. Real-world examples of non-mechanistic CoPs include antibodies against varicella virus glycoproteins and serum IgA antibodies against rotavirus.³

We made minor revisions (underlined) of the text on non-mechanistic CoPs in our discussion (lines 474-477).

"Potentially, CD4+ T-cell responses may not function as direct (mechanistic) mediators of protection but instead as non-mechanistic² correlates, i.e. as proxy biomarkers for an unmeasured protective mechanism (e.g., HLA-specific peptide presentation, tissue-resident immunity, and/or innate priming) present only in previously-infected individuals"

Additionally, lines 482-486 speak to this point:

"Given that Nucleocapsid is a core protein, it is unlikely for these D1 T-cell responses to directly reduce infection risk; more likely they are a non-mechanistic correlate reflecting the 'dose' of the 'live unattenuated vaccine' that is natural infection, where a greater dose implies more protection due to greater levels of other immune functions that increase with dose".

Reviewer #4 (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.

Response:

Thank you for your co-review.

Reviewer #5 (Remarks to the Author):

This revision responds to prior critiques although it still has poor readability and mostly limited novel findings. In my opinion the primary message of this manuscript is presented in Fig. 3 and 4. In the revision, the logic of the analysis is more clear although on the whole the manuscript is extremely difficult to comprehend. The main conclusions can be stressed better.

Strengths include the extensive description, rigorous data and statistical analysis. The authors stress that manuscript has been shortened, but unfortunately, for non-highly biostatistics trained researchers the manuscript is cumbersome.

One of the main conclusions that high spike and somewhat NC CD4 T cell responses in non-naïve individuals correlate with lower COVID-19 is not necessarily news worthy. There are multiple prior studies that have reached a similar conclusion often with spike, NC or other SARS-CoV-2 regions. The authors and prior reviewers have highlighted some of the previous publications. While the number of individuals analyzed in this study is a major strength; the authors, however, rightly point out the heterogeneity in the population is a limitation. Besides the number of individuals, the author could better highlight the novel findings. The significance of these findings relative to prior knowledge is somewhat marginal or incremental. It is generally well established that neutralizing antibodies along with T cell responses are likely important in ameliorating disease.

Response:

Thank you for the opportunity to revisit the literature and update our manuscript to clarify the differences between our study and previous T-cell immune correlates studies of COVID-19. While we agree that it is generally well established that neutralizing antibodies along with T-cell responses are likely important in reducing COVID-risk, our study has important strengths compared to previous studies. We have revised our Discussion (underlined) as follows (lines 562-585):

“Most prior studies assessing T-cell responses as correlates of COVID-19 have focused exclusively on Spike-specific responses (e.g., refs.⁴⁻⁷). Nucleocapsid-specific responses have also been assessed as correlates,^{8,9} but we did not identify such a study conducted after the introduction of variant-adapted boosters (recruitment from May to December 2020⁸ or from October to December 2020⁹). Many previous T-cell correlates studies that assessed T-cell responses focused on ELISpot, FLISpot, or T-cell receptor sequencing data (e.g., refs.^{4,5,9,10}), none of which can discriminate CD4+ from CD8+ T-cell responses; analyzed only one or two cytokines/markers (e.g., refs.^{5,6,9}); or, if polyfunctional T-cell subsets were examined by ICS, did not include assessment of Nucleocapsid-specific responses.⁷ Moreover, the statistical methods used in the present study expand on those used in prior studies, e.g., to our knowledge, controlled risk curves of COVID-19 risk level by T-cell response level have not been previously reported. Therefore, strengths of this study that differentiate it from previous COVID-19 T-cell immune correlates studies are: (1) its relevance to the variant-adapted booster era; (2) the use of a validated 27-color ICS assay, which enabled interrogation of CD4+ and CD8+ responses, as well as examination of polyfunctional subsets based on expression of CD154 and multiple cytokines; (3) the focus on T-cell responses to Spike and Nucleocapsid peptides; and (4) the diverse set of complementary statistical approaches taken. A second strength is the assessment of T-cell responses in univariable analyses, as well as T-cell responses and neutralizing antibody responses jointly in multivariable analyses, as correlates of COVID-19 separately in naïve and previously-infected individuals. As we elaborated on above,

performing correlates analyses in the two distinct cohorts yielded potential insight into the nature of the D1 and D15 Spike-specific CD4+ T-cell correlates (e.g. possibly non-mechanistic). Moreover, while prior multivariable correlate of risk analyses have been reported, they restricted to individuals who were SARS-CoV-2 seronegative at baseline⁴ or did not have prior history of SARS-CoV-2 infection.⁵

If we are missing any key references that should also be discussed here, we are happy to incorporate suggestions from the reviewer.

Readability can be improved by focusing on the main points. One of the main ideas is differences in T cell responses among naïve versus non-naïve individuals. Figure 1 and Figure 2 do not convey this point adequately because it muddled with vaccine antigen and stage. There should be a figure showing direct comparisons.

Response:

We have added two new supplementary figures, Supplementary Fig. 7 and Supplementary Fig. 12, with box plots showing Nucleocapsid-specific (Supplementary Fig. 7) and Spike-specific (Supplementary Fig. 12) D1 and D15 CD4+ and CD8+ T-cell D1 positive response rates, D15 positive vaccine-response proportions, D1 and D15 geometric means (GMs), and D1 and D15 response distributions, pooled across vaccine inserts within each stage. The figure also provides p-values from a Wilcoxon rank sum test (for D1 comparisons) or a treatment-arm-stratified Wilcoxon rank sum test (for D15 comparisons) for whether the distributions differ significantly between the naïve and non-naïve cohorts, within each stage.

We have added the following text (lines 217-256), which references these two new figures, to the revision:

“Higher Nucleocapsid-specific responses at baseline in the non-naïve vs. naïve cohort For Nucleocapsid Index, D1 CD4+ and CD8+ T-cell IFN-γ and/or IL-2 positive response proportions and response levels were higher in the non-naïve vs. the naïve cohort within each stage. The D1 CD4+ positive response proportions (non-naïve vs. naïve) were 25.4% vs. 1.5% in stage 1, 29.0% vs. 3.0% in stage 2, and 27.8% vs. 1.7% in stage 3; similarly, all $p < 0.001$ for testing whether response levels significantly differed in the naïve vs. non-naïve cohort (Supplementary Fig. 7a).

While cross-reactive CD8+ T cells targeting a highly conserved region of Nucleocapsid have been reported,¹¹⁻¹³ there was little evidence of such cross-reactivity in this cohort, with D1 Nucleocapsid-specific CD8+ T-cell positive response proportions of 1%, 0%, and 0% in the naïve cohorts in stages 1, 2, and 3, respectively (Supplementary Fig. 7b). D1 CD8+ T-cell IFN-γ and/or IL-2 Nucleocapsid Index did not pass the prespecified positive response rate eligibility criterion for correlates assessment and was not included in subsequent analyses. Similar to the CD4+ results, D1 CD8+ T-cell Nucleocapsid-specific response levels were higher in the non-naïve vs. naïve cohort within each stage (all

p<0.001, Supplementary Fig. 7b). The Nucleocapsid-specific results are additionally shown stratified by study arm in Supplementary Figs. 8-11.

For participants with higher Spike-specific CD4+ T-cell responses in the non-naïve vs. naïve cohort at baseline, levels became comparable after the second boost

In contrast to the Nucleocapsid results, D1 Spike BA.4/5-specific CD4+ T-cell response levels were only higher in the non-naïve (vs. naïve) cohort in stage 1. The GMs (95% CIs) were 0.13% (0.10%, 0.17%) vs. 0.10% (0.09%, 0.12%) in stage 1 (p=0.035), 0.13% (0.10%, 0.17%) vs. 0.13% (0.11%, 0.15%) in stage 2 (p=0.53), and 0.14% (0.11%, 0.17%) vs. 0.12% (0.08%, 0.17%) in stage 3 (p=0.74) (Supplementary Fig. 12a). A similar result was seen for D1 CD8+ Spike BA.4/5-specific responses, with GMs (95% CIs) of 0.13% (0.08%, 0.21%) vs. 0.04% (0.03%, 0.06%) in stage 1 (p<0.001), 0.07% (0.04%, 0.10%) vs. 0.06% (0.04%, 0.09%) in stage 2 (p=0.88), and 0.08% (0.04%, 0.13%) vs. 0.06% (0.04%, 0.09%) (p=0.50) in stage 3 (Supplementary Fig. 12b). After the second booster, CD4+ Spike BA.4/5-specific response levels were comparable between the non-naïve and naïve cohorts within all three stages, with GMs (95% CIs) of 0.28% (0.23%, 0.34%) vs. 0.23% (0.21%, 0.26%) in stage 1 (p=0.06), 0.16% (0.13%, 0.20%) vs. 0.18% (0.15%, 0.22%) in stage 2 (p=0.19), and 0.21% (0.16%, 0.26%) vs. 0.20% (0.14%, 0.29%) in stage 3 (p=0.73) (Supplementary Fig. 12a). However, CD8+ Spike BA.4/5-specific GMs remained higher at D15 in the non-naïve cohort in stage 1 [non-naïve 0.25% (0.15%, 0.42%) vs. naïve 0.10% (0.08%, 0.13%), p<0.001], and remained comparable in stages 2 and 3, with non-naïve vs. naïve GMs (95% CIs) of 0.11% (0.07%, 0.18%) vs. 0.12% (0.08%, 0.18%) in stage 2 (p=0.89) and 0.08% (0.04%, 0.14%) vs. 0.06% (0.04%, 0.09%) in stage 3 (p=0.54) (Supplementary Fig. 12b). Thus, in participants in whom Spike-specific CD4+ T-cell responses were higher in the non-naïve vs. naïve cohort at baseline, these responses became comparable after the second boost. This finding was not seen for CD8+ T-cell responses.”

We have added a new section to our Statistical Analysis Plan (Section 10) that includes these additional post hoc analyses.

We have opted to keep Figs. 1 and 2, which present the analogous data but additionally stratified by study arm, as main figures given that they directly address one of the study’s prespecified objectives (see Section 6 in the Statistical Analysis Plan, “The primary goal of plots in Collection I is to characterize and compare immunogenicity of the vaccine booster arms for various T cell markers listed in Section 5”). We state this objective as the concluding sentence in our Introduction (lines 133-134): “Moreover, we characterize T-cell immunogenicity for each of the three vaccine manufacturers across Spike vaccine inserts”.

In Figure 1 and 2, it is unclear how “negative” can have higher values than “responders.” Highlighting this in the main text will improve readability.

Response:

Thank you for the opportunity to clarify this point. A “responder” at D1 is defined by comparing the D1 T cell response to the paired D1 negative control sample. So, for a specified immune marker, e.g., antigen-specific CD4+ T cells expressing a particular cytokine combination, a responder at D1 is a participant whose D1 antigen-specific CD4+ T cells showed any evidence of expressing the given cytokine combination based on comparison to the paired D1 negative control sample. Therefore, a responder at D1 means a SARS-CoV-2 specific T-cell response. On the other hand, at D15, a responder is defined as a “vaccine responder”. This is a participant whose D15 immune response (e.g., antigen-specific CD4+ or CD8+ expressing a particular cytokine combination) is significantly higher than their corresponding response at D1. Hence, a participant could have a high D15 response level and still be a non-responder at D15 if the participant’s D15 response is not meaningfully higher (according to the statistical definition) than the D1 response level.

We describe this in more detail, and include a reference to the relevant methods paper, in the Methods section (lines 700-711):

“D1 and D15 positivity calls for T-cell responders

At D1, a responder was defined by a positivity call made by comparing the percentage of T cells secreting the given set of cytokines in the relevant sample under stimulation compared to that in the paired control (sample from the same person-time-point in an unstimulated condition). Those with Fisher’s exact test $P < 0.0001$ were defined as responders. At D15, a positive vaccine-response for percentage of CD4+ T cells expressing a given set of cytokines (“booster response over baseline”) was defined by the worst-case p-value comparing the percentage of positive cells post-vaccination versus baseline being significant when controlling the false discovery rate at 5%, considering all assay misclassification rates that are consistent with the paired control samples at baseline and post-vaccination.¹⁴ A positive vaccine-response against Spike BA.4/5 peptides was defined as a positive vaccine-response against Spike BA.4/5 subunit S1 or subunit S2 peptides.”

This information is also repeated in the captions of Figs. 1 and 2, and in footnote #1 of Table 1.

As suggested by the reviewer, we have additionally added the following clarification in the main text, at the first instance of reporting a D15 vaccine response (lines 193-197):

“(The definition of a “positive vaccine response” at D15 differs from the definition of a “positive response” at D1, where the former is based on a participant’s D15 T-cell response being statistically significantly higher than their corresponding D1 T-cell response, whereas the latter is based on comparing the participant’s D1 T-cell response to their paired D1 negative control sample, see Methods).”

While it is important to consider the vaccine inserts, stages, days post vaccination, would this main point be conveyed by showing naïve versus non-naïve responses as a main figure. Different individuals can be symbol and color coded based on vaccine and insert etc. The current form is not very understandable.

Response:

Please see our response to your comment above, “Figure 1 and Figure 2 do not convey this point adequately because it muddled with vaccine antigen and stage. There should be a figure showing direct comparisons.” In summary, we have added two new additional supplementary figures (Supplementary Fig. 7 and Supplementary Fig. 12) that simplify the comparison (between the naïve and the non-naïve cohorts) of D1 and D15 CD4+ and CD8+ responses to Spike and to Nucleocapsid by pooling across study arms, within each stage.

Another major finding is that NC and spike CD4+ T cell responses are a correlate of protection against COVID-19 in the non-naïve population. Figure 3 and 4 could be simplified by reducing the number of panels.

Response:

We have reduced the number of panels in Fig. 4 from twelve to seven by moving the two booster-proximal controlled risk curves and the three functionality score controlled risk curves to two new supplementary figures (Supplementary Figs. 22 and 27 in the revision).

Given that Fig. 3 has only two panels, we have opted not to make any changes to it.

As mentioned by multiple prior reviewers, this finding is not necessarily novel. I find the discussion regarding this point informative. They may also consider that It is likely that enhanced CD4 responses don't protect against infection but rather ameliorate COVID19 disease. Thus individuals with high CD4+ T cell directed responses were likely not diagnosed with COVID-19 because it may have been mostly asymptomatic or not meriting any diagnostic evaluation.

Response:

Thank you for this question on asymptomatic SARS-CoV-2 infection. There is ample evidence supporting a role for SARS-CoV-2 specific CD8+ T cell responses in preventing severe disease,¹⁵⁻¹⁸ and we agree it is also plausible that high SARS-CoV-2 specific CD4+ T cell responses in individuals who have received a second booster may associate with reduced severity or even asymptomatic SARS-CoV-2 infection. This scenario could occur, for example, if activation of memory CD4+ T cells upon SARS-CoV-2 exposure leads to faster production of higher-affinity antibodies, limiting early viral replication and inflammation and increasing the likelihood of asymptomatic or mild infection.¹⁹

However, the endpoint considered in our study was symptomatic SARS-CoV-2 infection (“The COVID-19 endpoint was symptomatic SARS-CoV-2 infection, as examined previously²⁰”, line 152). In the COVAIL study, asymptomatic infection endpoints between 7 and 91 days (booster-proximal period) and between 7 and 188 days (overall period) post D15 were determined based on anti-Nucleocapsid serotesting at study visits D29, D91, and D181. In total, 17 and 29 asymptomatic infection endpoints were recorded during the booster-proximal and overall periods of COVAIL among the 569 participants in the Per-Protocol Immunogenicity Analysis Set, compared to 95 and 144 COVID-19 cases during the same period. Thus, the majority of infections in the COVAIL study — and all infections included in the correlates analyses — were symptomatic, with a small fraction that were asymptomatic and not included in the correlates analyses.

We agree that in theory it would be interesting to evaluate the relationship between cellular immune responses and asymptomatic infections. Correlates analyses for asymptomatic cases are hindered by (1) the small number of cases with PBMC samples; and (2) lack of precise time-to-event information about the asymptomatic cases. Specifically, the asymptomatic cases were interval-censored, meaning we only know they occurred between D29 and D91 (if the participant tested anti-N negative at D29 and then anti-N positive at D91) or between D91 and D181 (if the participant tested anti-N negative at D29 and D91 and then anti-N positive at D181). Moreover, a correlation between an immune marker and asymptomatic infection risk is harder to interpret than a correlation between an immune marker and COVID-19 risk. This is because the immune marker may increase asymptomatic infection risk via the immune marker’s effect on symptomatic cases. For instance, this could be the case if the immune marker prevented symptomatic cases, and thus increased the number of participants in the at-risk set for potential later asymptomatic infections. For these reasons, the scope of this article did not include the immune correlates analysis for asymptomatic infections.

References

- 1 Plotkin, S. A. Correlates of protection induced by vaccination. *Clin Vaccine Immunol* 17, 1055-1065, doi:10.1128/CVI.00131-10 (2010).
- 2 Plotkin, S. A. & Gilbert, P. B. Nomenclature for immune correlates of protection after vaccination. *Clin Infect Dis* 54, 1615-1617, doi:10.1093/cid/cis238 (2012).
- 3 Plotkin, S. A. & Gilbert, P. B. Correlates of Protection. In: *Vaccines*, Eighth Edition, Editors Walter Orenstein, Paul Offit, Kathryn Edwards, Stanley Plotkin. Pages 35-40. Elsevier Inc., New York. (2022).
- 4 Hertoghs, N. et al. Vaccine-induced T cell responses correlate with reduced risk of severe COVID-19 in a placebo-controlled efficacy trial. *EBioMedicine* 117, 105809, doi:10.1016/j.ebiom.2025.105809 (2025).
- 5 Kemlin, D. et al. Humoral and cellular immune correlates of protection against COVID-19 in kidney transplant recipients. *Am J Transplant* 23, 649-658, doi:10.1016/j.ajt.2023.02.015 (2023).

- 6 Torres, I. et al. SARS-CoV-2 Omicron BA.1 variant breakthrough infections in nursing home residents after an homologous third dose of the Comirnaty(R) COVID-19 vaccine: Looking for correlates of protection. *J Med Virol* 94, 4216-4223, doi:10.1002/jmv.27867 (2022).
- 7 Brasu, N. et al. Memory CD8(+) T cell diversity and B cell responses correlate with protection against SARS-CoV-2 following mRNA vaccination. *Nat Immunol* 23, 1445-1456, doi:10.1038/s41590-022-01313-z (2022).
- 8 Eser, T. M. et al. Nucleocapsid-specific T cell responses associate with control of SARS-CoV-2 in the upper airways before seroconversion. *Nat Commun* 14, 2952, doi:10.1038/s41467-023-38020-8 (2023).
- 9 Molodtsov, I. A. et al. Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)-Specific T Cells and Antibodies in Coronavirus Disease 2019 (COVID-19) Protection: A Prospective Study. *Clin Infect Dis* 75, e1-e9, doi:10.1093/cid/ciac278 (2022).
- 10 Fenn, J. et al. Early de novo T cell expansion following SARS-CoV-2 infection predicts favourable clinical and virological outcomes. *eBioMedicine* 117, doi:10.1016/j.ebiom.2025.105795 (2025).
- 11 Lineburg, K. E. et al. CD8(+) T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope cross-react with selective seasonal coronaviruses. *Immunity* 54, 1055-1065 e1055, doi:10.1016/j.immuni.2021.04.006 (2021).
- 12 Nguyen, T. H. O. et al. CD8(+) T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope display high naive precursor frequency and TCR promiscuity. *Immunity* 54, 1066-1082 e1065, doi:10.1016/j.immuni.2021.04.009 (2021).
- 13 Murray, S. M. et al. The impact of pre-existing cross-reactive immunity on SARS-CoV-2 infection and vaccine responses. *Nat Rev Immunol* 23, 304-316, doi:10.1038/s41577-022-00809-x (2023).
- 14 Chen, Z., Heng, S., Tapley, A., De Rosa, S. & Zhang, B. Determining vaccine responders in the presence of baseline immunity using single-cell assays and paired control samples. *Biostatistics* 26, doi:10.1093/biostatistics/kxaf045 (2025).
- 15 Goldblatt, D., Alter, G., Crotty, S. & Plotkin, S. A. Correlates of protection against SARS-CoV-2 infection and COVID-19 disease. *Immunol Rev* 310, 6-26, doi:10.1111/imr.13091 (2022).
- 16 Peng, Y. et al. An immunodominant NP(105-113)-B*07:02 cytotoxic T cell response controls viral replication and is associated with less severe COVID-19 disease. *Nat Immunol* 23, 50-61, doi:10.1038/s41590-021-01084-z (2022).
- 17 Yang, Y. et al. The characterization of CD8(+) T-cell responses in COVID-19. *Emerg Microbes Infect* 13, 2287118, doi:10.1080/22221751.2023.2287118 (2024).
- 18 Bertoletti, A., Le Bert, N. & Tan, A. T. SARS-CoV-2-specific T cells in the changing landscape of the COVID-19 pandemic. *Immunity* 55, 1764-1778, doi:10.1016/j.immuni.2022.08.008 (2022).

- 19 Le Bert, N. & Samandari, T. Silent battles: immune responses in asymptomatic SARS-CoV-2 infection. *Cell Mol Immunol* 21, 159-170, doi:10.1038/s41423-024-01127-z (2024).
- 20 Zhang, B. et al. Neutralizing antibody immune correlates in COVAIL trial recipients of an mRNA second COVID-19 vaccine boost. *Nat Commun* 16, 759, doi:10.1038/s41467-025-55931-w (2025).

REVIEWERS' COMMENTS

Reviewer #3 (Remarks to the Author):

As mentioned before, all my points have been addressed sufficiently.

Overall, the manuscript has become significantly more readable.

Response: Thank you for the positive comments.

Reviewer #5 (Remarks to the Author):

I have no further comments. The authors have addressed my prior concerns as well as possible.

Response: Thank you for the positive comments.

Reviewer #5 (Remarks on code availability):

None