Peer Review File

Manuscript Title: Influenza vaccination reveals sex dimorphic imprints of prior mild COVID-19

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

We initial read with enthusiasm the title and abstract of this paper and we commend the authors on the amount of work that was put into this study. Unfortunately, there are too many major issues with the analyses that dampen enthusiasm.

Conceptual
The authors do not adequately justify why they chose influenza vaccination as an immune challenge or address confounding that could have been caused by pre-existing immunity to influenza. As is well-documented in the literature, prior exposure to influenza through either infection or vaccination is nearly universal in middle-aged individuals, leading to heterogeneous levels of pre-existing immunity, which is a fundamental predictor of the response to vaccination. Despite some effort to statistically correct for vaccination history (which is never addressed in the main body of the text), the authors do not adequately address this. For example, the low percentage of 'responders' in figure 2B is likely due to high levels of pre-existing immunity precluding a measurable 4-fold change. While one justification for using influenza as an immune challenge could be that the response to influenza vaccination is very well-characterized in the literature, the authors transform the data in such a way that it cannot be directly compared to the literature. As a result, we think that the interpretations of transformed data by the authors are not completely accurate and become so difficult for the reader to assess that we found ourselves questioning the validity of many statements by the authors.

Text-specific
• Line 15: [5th-95th percentile:58 -235] days after diagnosis need clarification. It is a little confusing what the authors meant '5th - 95th percentile' and range of days.
• Line 20 -21 - Authors categorized CD71hi B cells as a subset including memory B cells and plasmablasts (as described in Supplementary Figure 2B). Authors need further clarification and/or explanation why CD71hi B cells (including influenza-specific subsets) increased before vaccination in COVID-19-recovered male subjects. Other very well-respected vaccine immunologists define IgD- and CD71- as memory B cells in the context of influenza vaccination (doi: 10.1038/ni.3533).
• Line 67: Authors need to identify the time of symptom assessment. Also, the number of total subjects and the percentage are important.
• Line 77: Was there any significant difference between male and female subjects.
• Line 82: the red cell distribution' should be 'the red blood cell distribution'.
• LIne 82 - 84: The normal range of RDW is from 11.8 to 15.6 percent, which tends to increase with age. The authors showed the elevation and decrease of RDW in COVID-19 subjects, but the changes

are within the normal range. In this case, the authors should have a healthy control group as a baseline.

• Line 84 -87: The decrease appears to be 1.5 -2 % which is within the normal range. Thus, the data do not support the statements made.
• Line 88 - 92: The decrease of plasmablast over time is a normal immunological reaction, which is not specific to COVID-19. The authors should have looked at plasmablast between Day 0 and Day 20, which is more sensitive to demonstrate the change in plasmablasts (Extended Figure 1 a -c).
• Line 113 -117: COVR-M has a higher innate immune activation and metabolic process than COVR-F. But T cell surface signature and T cell differentiation in COVR-M are lower than in COVR-F. These data suggest that the adaptive immune response might be better in COVR-F than in COVR-M, which is different from the conclusions drawn by the authors.
• Line 150 - 157: The author should elaborate on the meaning of 'true' and 'false' in Figure 2b. It is unclear whether the authors combine the titers of different strains. and how the normalization was processed. In healthy subjects, the difference between male and female subjects is present.
• Line 160-163: Additional clarification is needed for a better understanding of sex difference in Flu-specific memory B cell response. Influenza A -specific B cells (H3+H1-, H3+Flu-B-, H1+H3-) in HC show a higher fold change than those in COVR-F. Influenza B specific memory B cells showed the opposite trend. Also, all memory B cells in males show higher levels than those in females. The representation of the data is so transformed that it was particularly difficult to interpret.
• Line 310: Authors stated that the adaptive immune response is elevated in COVR-M. However, supportive data were not presented in the manuscript.
• Line 339: Authors stated the plasmablasts increased in COVR-M. However, Figure 2b shows the comparison of frequency (%). We suggest that the comparison should be conducted using ratio, not the frequency.

Figure-specific
• All stacked bar graphs (1B, 2B, 4F) are difficult to interpret and do not support the conclusions made in the text.
• 1D/E: Unclear why correlation plots here. There is a clear dependent (nAb/RBC) and independent variable (TSD). Why not use regression analyses with a time by sex interaction term, which would allow direct testing of the sex different in trends? Also for RBC where are control patient data?
• 2B: typo in Male panel, says that OR =0.
• 2C: unclear if statistics account for that fact that the timepoints are not independent of each other. I do not think this panel was addressed in the results section.
• 3: unclear how samples at day 1 OR day 7 were used. Did the same individuals contribute samples at both timepoints? Or did participants contribute samples at either day 1 or day 7? Time-course analyses (C and E) suggest that these are the same people over time. Could small sample sizes in at day 7 explain the dips seen at this timepoint in C and E?
• 1Figure 2d: It appears that the authors subtract frequency on Day 0 from Day 7. Subtraction of % (frequency) should be replaced with actual cell count numbers. The strain of influenza should be provided. The authors stated 'Flu-specific plasmablast' in Figure 2d. Sex differences in flu-specific plasmablasts must be evaluated.
• Figure 3b-d: Figure 3b (transcriptomic analysis) shows little difference in interferon-gamma response between COVR-F and COVR-M. However, Figures 3c and d show that significantly higher

IFN-gamma response in COVR-M compared to that in COVR-F. Additional information would be helpful for a better understanding.
• Figure 3e: While interesting, absolute monocyte counts presented in this figure are confusing compared to Figure 1g data (frequency). In Figure 1g, the frequency of monocytes was similar between HC_M and COVR-M on Day 0 whereas a marked difference was presented in absolute monocyte counts between HC-M and COVR-M in Figure 3e. While the frequency and absolute count are different measures, additional description and explanation would be helpful.
• Figure 4d: It is unclear how many subjects were included in the gene heatmap. HC appears to have 3 subjects and COVR appears to have 8 subjects. However, there seems a notable variation within the group.
• Figure 4c: If a sex difference is present in this case, then separate figures for males and females should be included.
• The entirety of Figure 4 and the use of a new population of people from Italy without having the HC day 1 control group, which makes all of these comparisons confusing and, in some cases, incorrect.

Referee #2 (Remarks to the Author):

Sparks et al characterise the long-term impact of mild SARS-CoV-2 infection on immune status at baseline and following seasonal Flu vaccination. The study identifies shifts in immune signatures between healthy and following mild COVID19 that are different among sexes. In terms of effects of mild COVID19 on baseline immune parameters, sex differences were observed with increased monocytic, oxidative phosphorylation and fatty acid metabolism in males vs increased T cell processes in females. Following influenza vaccination, a proportion of males had higher Ab titers at day 28 post vaccination than females, together with higher frequency of flu-specific B memory populations and plasmablast frequency. Vaccination induced shifts from baselines in a sex-dependent manner, with males showing significant up-regulation of IFN signalling and IFNg transcriptomic responses and IFNg levels at day 1 following vaccination when compared to females. Other features that were differentially observed between the sexes were increased immune activation in males and increased neutrophil, NK and T cell modules in females at day 1 post vaccination, and increased B/plasmablast and neutrophil signatures in males at day 7 post vaccination. Finally, the authors identify a depressed immune state in COVR that is partially resolved upon vaccination, with genes relating to oxidative phosphorylation, monocyte and cell cycle processes partially returning to baseline in females, and processes relating to plasmablasts returning to baseline levels in males after 28 days post vaccination. Vaccination induces transcriptional changes in classical monocytes at day 1, similar to baseline. Interestingly, a larger fraction of females had genes whose expression reset by day 28 when compared to males suggesting females are able to return to closer to a baseline more effectively than males. This manuscript offers several interesting new perspectives on immune recovery after acute infection in general and COVID-19 in particular and thus opens up several new avenues for future studies.

General comments and suggestions
- Subjects with variable ages were included in the study. Because sex differences were observed for

many features, it would be interesting to indicate how sex differences occur when comparing subjects pre and post andropause/ menopause; if data is not available, perhaps separating by age groups.

- More insights on how sex differences in immune and metabolic findings relate to one another?

- Fig 2: More discussion is warranted on why male influenza antibody titers are higher in COVR vs HC in males while the opposite happens in females? Any difference in prior immunization history between groups?

- Were there reported days following SARSCoV2-specific vaccination and/ or differences between the sexes and groups of COVR vs HC?

- Some of the individuals were samples quite recently after a prior COVID-infection. Is there any difference in relation to time from prior COVID in these subjects?

- The observation of CD71 hi B-cells after COVID-19 is curious and I wonder whether any particular plasma protein correlated with the abundance of these cells?

Minor points:
- Clarify age range and distribution among groups
- Clarify in text whether average time from diagnosis was accounted for in analysis.
- Specify range of self-reported average length of illness 20.8 days and whether counted for in analysis
- Clarify Figure 1b legend in terms of Y N bars relating to residual symptoms
- Clarify in text that Fig 1d only show significant anticorrelation between time of diagnosis and neutralizing antibody titers in males: Lines 78, 93, 94
- Clarify normal range for RDW in Fig 1e and comparison with uninfected males and females
- Line 109 conclusion based on males, but Figure 1g suggests no changes in monocytes in males between healthy vs COVR, whereas females have decreased % monocytes.
- Information/ differences in vaccinations between subjects who got covid/ did not?
- Verify Fig 2b male OR=0?
- Fig 2e indicates higher FC COVR vs HC in males but not females; are there differences in baseline values (rather than FC) between males and females?
- Figure 3 indicates that vaccine induced shits from baseline states are quite between sexes; hard to interpret

- Figure 4 indicates partial reset of a previously depressed immune state in COVR subjects following influenza vaccination. More discussion on why would influenza vaccine reset post covid immune states back to healthy controls? And why sex effects are impacting this reset?

- Overall observation of a depressed immune state following COVID. Are these effects speculated to result from SARS-CoV-2-specific effects or due to overall effects of a respiratory infection? As for the HC and COVR, did any of the subjects have any symptomatic respiratory infection, other than COVID19, that could be accounted for in some of the analyses?

Referee #4 (Remarks to the Author):

This is an ambitious attempt to address a series of important questions: first, whether mild to moderate COVID-19 leaves a long-term imprint in the immune system that can be detected by -omics approaches of the (potentially) new steady state or after a normalised, antigenically different challenge of the immune system, namely influenza vaccination. Second, whether any of the observed imprints are sex-specific. Overall, this paper is well written and holds together (at least in most places) the high complexity of analysis. Also, the graphic displays are helpful to guide the reader through the complex data, and overall, the discussion is fair. My main concerns with this manuscript are that the observed phenotypes are very subtle, the n numbers are not very high, the phenomena described are slightly unlinked, and due to the nature of the study, correlative data is combined to develop a narrative that remains however speculative. This in combination with the fact that only blood data is available for analysis, while both in SARS-CoV-2 infection and in vaccination, local processes are crucial, it remains difficult to judge the biological meaning of the phenomena described here.

Major concerns:

1.Inter-individual variation in response to influenza vaccines is strongly determined by the individual pre-vaccination exposure history. A vaccine that all study participants were naïve to would have been a better choice and would have facilitated some of the analysis.

2. It is a pity that true longitudinal data, i.e. different time points since diagnosis from the same donor, is not part of this study – again this would have strengthened the analysis.

3. An important point that is at least missing in the main text is the question of participant infections between COVID and vaccination – all assumptions here are based on the absence of such confounding factors – are we sure that participants did not have anything, not even a cold in that period?

4. Following on from point 3: COVID infection of the participants was not the first, but the nth infections of their life. I understand from the comparison made with flu infection data that similar imprints were found after other infections (and this is what I would expect). So the concept of a sort of persisting immune-paralysis that is partially or fully relieved is difficult to reconcile with a long sequence of infections – which one depresses, which one relieves the immune depression? Or would it be specific to COVID as compared to other infections to induce immunodepression, and of vaccination as compared to infection to be able to reinstate to original state? Apart from flu data, the authors cite literature on vaccinations to have immune-depressive effects, so the answer to both questions above seem to be no. I think it is important not only to discuss this, but to re-consider the basic assumptions of the study in the light of sequential immune-exposure that started long before COVID for these individuals.

5. Line 60 why include 2 asymptomatic subjects? Do you know when they were infected, probably not? How do you establish TSD? How do the analyses look without them?

6. Along the same lines, it is not clear to me why in different analyses, different n numbers appear, when figure 1c suggests that the key analyses were done on all 71 study participants (e.g. 2b, c, also later figures).

7. It is a pity that no PBMCs were frozen to allow testing of some of the findings more directly – for instance, the authors link higher serum IFNg on d1 post vaccination to the increased presence of CD8 effector cells in M convalescents per-vaccination – it could be tested whether these cells are prone to IFNg production if such samples were available. It could also be tested whether increased NK cells on d1 post vaccination in F would not produce higher IFNg. Such minimal functional experiments would strengthen some of the links that the authors plausibly draw but which remain speculative.

8. In figure 1 and related data, stable and evolving changes in immune set points are analysed. Subtle differences in monocytes and T cells are reported, partly reminiscent of the reported more dramatic changes during acute and severe COVID-19. Many of the cellular, protein and Cytek analyses do not reach a FDR< 0.05 significance value.

In the related data in fig. ED1A, of five pathways shown, the only one with a significant (FDR <0.05 – not a very stringent cut-off, we hardly ever use anything less stringent than < 0.01) enrichment factor in F does not correlate significantly with TSD, and vice versa for the four other ones. I appreciate the authors' argument that the (significant) loss over time in F but not in M leads to (significant) sex differences overall in a subset of these categories, but overall, my impression is that the long-term imprints detected here are quite subtle. The fact that we don't really know what a change in e.g. DC representation in the blood means, given the tissue- or lymph node restricted role of these cells, makes it difficult to develop a clear narrative of the changes we observe and their biological meaning.

Line 129 I agree that M/F combined analysis yields less significant differences, but could this be due to increased n number for analysis? Concretely, the stats shown if EDFig1F refer to sex differences, while the figure is used for an argument about sex-independent differences (line 130), while the Cytek analysis EDFig1d does not show sex-independent differences for pDCs.

9. In figure 2, the Ab and B cell response to influenza vaccine is analysed, and recovered M appear to have a slightly stronger Ab response that correlates with higher frequencies of plasmablasts and higher baseline levels of a specific memory B cell subset, which in turn correlates with an mTORC-driven gene signature. If I understand correctly, the metabolic signature is derived from bulk analysis of PBMCs, so this may or may not reflect metabolic changes in this B cell subset. I do not see that the B-cell specific metabolic argument is picked up upon once the scRNAseq in fig.4 is done, and EDFig.5B does not seem to show detectable differences for B cells in males.

10. In fig. 3, further analysis of the anti-vaccine response is performed, and strong early IFN responses were found in recovered M. The authors link IFNg reponses to higher base-line CD8 effector cells levels in M, and I wonder why increased NK-frequencies in F cannot provide a similar level IFNg. As mentioned above, testing function of cells from frozen aliquots would come in handy here.

11. Finally fig.4 explores a vaccination-driven immunological “reset”. Apart from the conceptual issues (explained in point 4) to consider, his figure is not easy to understand. In particular, for Fig. 4B the legend talks about four lines, including the HC base line, but this fourth line is missing – is the HC base line the comparator and therefore would be the zero value on the x axis? Based on this initial issue, the whole argument of vaccination releasing immune paralysis post COVID is difficult to follow. The x-axis annotation (abs(effect size)) is not helpful to understand this, and the precise selection of genes that are channelled into this analysis is difficult to follow. It is also not helpful for the comparison that the scale of the x axis for M and F is not the same. The authors conclude from the data analysis in this and associated figures that indeed, at d1 post vaccination of convalescents, there is a release of depressed genes, in particular in monocytes, which persists up to day 28 more in F than in M. Early monocyte activation post vaccination is perhaps unsurprising, so to talk about a relief of an immune-depressed state may be pushing it a little – but it does show that we ae not in front of an immune-paralysis that cannot be overcome.

Minor concerns:

1. Typo EDFig.2c: COVR-F instead of COR-F

2. File names (at least of the files extracted from the zip) do not indicate which suppl. Tables are in the excel sheet, often it is not indicated in the excel sheet itself – this makes it difficult to follow the data referenced to tables. Having all tables, well annotated, in on excel sheet may be preferable.

Author Rebuttals to Initial Comments:

Response to Reviewers

Common response to reviewers

We thank the reviewers for the constructive comments that have helped us significantly improve our manuscript. We are pleased and excited to learn that all three reviewers are enthusiastic about the topic and questions that our study explored. For example:

"*this manuscript offers several interesting new perspectives on immune recovery after acute infection in general and COVID-19 in particular and thus opens up several new avenues for further studies.*" (**Reviewer #2**)

"*this is an ambitious attempt to address a series of important questions: first, whether mild to moderate COVID-19 leaves a long-term imprint in the immune system that can be detected by -omics approaches of the (potentially) new steady state or after a normalised, antigenically different challenge of the immune system, namely influenza vaccination. Second, whether any of the observed imprints are sex-specific.*" (**Reviewer #3**).

Reviewer #1 is also enthusiastic about the questions and premise of the study and commended us on the "*amount of work that was put into this study*".

Overview of our revised manuscript

A substantial number of concerns from the reviewers stemmed from the complexity of the data and dense presentation in the original manuscript. Thus, in this revision we have simplified the data presentation to better highlight key findings and clarify potential confusions. Importantly, we have generated new multimodal single cell data (using CITE-seq measuring 138 surface proteins, transcriptome, and V(D)J simultaneously in single cells) now covering 40 subjects total (12 COVID-19-recovered males, 8 healthy males, 12 COVID-19-recovered females and 8 healthy females) including days 0, 1 and 28. Our single cell data significantly strengthen our conclusions and help to address the reviewers' questions. To the best of our knowledge, this also represents the largest multimodal single cell dataset on vaccination responses to date and could thus also serve as a data resource for the community.

Some of the results in the previous manuscript that might have appeared less concrete and connected—for example, because the bulk blood transcriptomic data is confounded by contributions from changes in cell frequency and cell state—are now clarified with the single cell analyses, which reveal the cell type-specific transcriptomic signatures associated with prior mild COVID-19. This data also helps solidify and streamline the presentation of our findings: we now present the key single cell-based results in the main figures and place the bulk transcriptomic and related supporting data in the Extended Data Figures and Supplementary Figures/Tables. Finally, as suggested by the reviewers, we added discussion, including additional background and context for our study design and findings, such as the rationale for using the influenza vaccine as an antigen-agnostic challenge of the human immune system.

Our current manuscript has a revised flow as summarized below:

1. In Fig. 1 (and Extended Data Fig. 1) we present the study design and highlight key examples of stable cellular differences associated with prior mild COVID-19. Stability is defined as differences between COVID-19 recoverees and their matched healthy controls that do *not* correlate with the time since COVID-19 diagnosis (TSD). These examples include transcriptional states of monocytes assessed using

the new single cell CITE-seq data. Throughout the manuscript, full analysis results (e.g., differences associated with prior mild COVID-19) are included in the Extended Data Figures and Supplementary Tables. The associated software code and raw data will be published together with our manuscript.

2. In Fig. 2 (and Extended Data Fig. 3) we present key sex-dimorphic differences in response to the influenza vaccine associated with prior mild COVID-19. These include the elevation of circulating IFNγ and its downstream transcriptional responses across different immune cell types on day 1 following vaccination, as well as increased humoral responses involving plasmablasts and antibodies on days 7 and 28, respectively, in COVID-19-recovered males. We also confirmed previously published findings (Extended Data Fig. 3c) that the humoral response was correlated with the IFNγ response on day 1, supporting the hypothesis that the elevated IFNγ response in COVID-19-recovered males could have contributed to their more robust humoral responses.

3. In Fig. 3 we identify GPR56+ effector memory CD8+ T-cells with a surface-protein phenotype that resembles "virtual memory" cells as a source for the elevated IFNγ response on day 1 following vaccination in COVID-19-recovered males. CITE-seq data reveal that the frequency of these cells was higher in COVID-19-recovered males at baseline (prior to vaccination) and the IFNG transcript in these cells increased significantly on day 1 compared to day 0 in COVID-19-recovered males only. Together, our data establish a link between baseline, pre-vaccination immune states in COVID-19-recovered males and their heightened innate responses following influenza vaccination.

4. Fig. 4 is similar in conclusion to the previous version of our manuscript but now includes additional CITE-seq data (more subjects and inclusion of day 28 data in addition to days 0 and 1) and a more simplified presentation focusing on the single cell analysis results only (we no longer use the bulk transcriptomic data). We show that the average expression of genes in the "depressed" monocyte transcriptional signature (involving immune receptors such as TLRs identified earlier and shown in current Fig. 1f and g) increased on day 1 following vaccination in COVID-19-recovered subjects and persisted to day 28 in both classical and non-classical monocytes. The effects are generally stronger in females with a significantly larger fraction of the genes demonstrating reversal. Thus, transcriptional changes in monocytes elicited early (day 1) by influenza vaccination partially reverse the immune status of COVID-19-recovered individuals towards that of matching healthy individuals who never had COVID-19 at the time of our study.

We believe our revision has adequately addressed the reviewers' concerns. Together, our manuscript provides exciting new data on using seasonal influenza vaccination and multiomic single cell approaches to reveal the impact of mild, non-hospitalized COVID-19 (with more than 500 million global cases thus far). Furthermore, our work provides conceptual advances regarding how mild viral infections can shape human immune status and function after acute illness and recovery to establish new antigen-agnostic set point that can impact future immune responses; in turn, how heterologous vaccination can be used to probe such imprints and help reverse the immune system back towards the state before SARS-CoV-2 infection.

Below we first respond to common questions raised by the reviewers followed by a point-by-point response to the reviewers' comments.

Common response #1: Why use the influenza vaccine in our study design?

Given the public health importance of seasonal influenza vaccination, and the possibility of concurrent circulation of influenza and COVID-19 in Fall/Winter 2020-21, we were interested in understanding whether prior mild COVID-19 may negatively affect the immune response to the seasonal influenza vaccine. We thus felt that important information, from both basic science and public health perspectives, could be gained from using the seasonal influenza vaccine as the immune challenge. Additionally, as we mentioned in the manuscript, the dynamics of the immune response to the seasonal influenza vaccine has been well characterized by previous studies using blood transcriptomics and immune cell phenotyping; this information allowed us to sample study subjects at the most informative time points post-vaccination.

Using this vaccine was also logistically convenient as we were able to catch the first wave of COVID-19 recoverees (infected early 2020) for influenza vaccination during the Fall of 2020. Importantly, our study/cohort is unique due to the timing of vaccination: to the best of our knowledge, the influenza vaccine was the first immune challenge that the COVID-19-recovered individuals likely encountered after SARS-CoV-2 infection. At the time of the study, COVID-19 vaccines were not yet available, and the circulation of other viruses (including influenza) was largely non-existent, partly due to social distancing and pervasive use of face masks in the Washington DC area. Because vaccines and natural viral infections may induce immunological changes that could have hindered our ability to discern immune imprints and influenza vaccine responses associated with prior COVID-19, it would be impossible to pursue a similar study later during the pandemic (i.e., after exposure to the COVID-19 vaccine, other SARS-CoV-2 strains, or other pathogens and/or vaccines). Thus, our study is uniquely positioned to map out the immune imprints and altered responses to an antigenically distinct immune challenge following prior mild COVID-19.

While challenging individuals with naïve antigens (e.g., yellow fever or rabies vaccines) can provide informative and complementary data (e.g., how the post-COVID-19 immune system may respond to a previously unseen immune challenge), there were logistical challenges and lack of clinical indication for using those vaccines in this population. The immune response kinetics are also less well established compared to the influenza vaccine (e.g., the innate and plasmablast responses peak on days ~3-5 and ~10-14 after live-attenuated yellow fever vaccination, respectively).

We have added discussion to highlight these considerations (lines 497-507); see also lines 188-196.

Common response #2: How does influenza vaccination/infection history and prior influenza immunity impact the vaccine responses reported in this study?

We agree that prior exposure to the influenza virus (either through infection or vaccination) can impact the response to seasonal influenza vaccination. However, using the seasonal influenza vaccine as an immune challenge has clear merits (please see response to the question above).

In addition, while prior exposure history and related surrogates such as pre-existing antibody titers are important determinants of influenza vaccination responses, they can only partially explain inter-individual response variations; baseline levels of circulating influenza-specific memory B- and T-cells also tend not to correlate with antibody responses. For example, we and others have shown that baseline, pre-vaccination antigen-agnostic immune statuses can predict antibody responses to influenza vaccination beyond the effects of age, gender, and baseline antibody titers[1,2]. Based on those earlier observations and the concept of "trained immunity"[3], we designed our study to assess whether a prior viral infection (in this case, by SARS-CoV-2) can change the baseline immune statuses to impact the response to an antigenically distinct immune challenge.

While ascertaining and thus matching the influenza (vaccine or infection) exposure history of the COVID-19-recovered (COVR) and healthy groups is difficult, these groups are age- and sex/gender-matched and were drawn from the same geographic area. Thus, on average both the age-dependent (e.g., strain of the very first encounter) and recent exposure histories are likely quite comparable between the groups and therefore exposure history alone is unlikely the sole factor driving the differences we observed among the groups. Indeed, the baseline antibody titer levels are comparable and statistically indistinguishable between the COVR and healthy groups for both sexes (Extended Data Fig. 3e, f). This suggests that the plasmablast and antibody response differences among the groups cannot be attributed to differences in their baseline antibody titers, which is known to correlate negatively with the fold-changes in titers following influenza vaccination.

Finally, age, self-reported influenza vaccination history (of the past 10 years) and pre-vaccination (baseline) influenza antibody titers are accounted for in our results as they have been included as co-variates in our statistical model for assessing immune response differences among the groups. We model post-vaccination titer as the outcome variable with influenza vaccination history and pre-vaccination titer included as covariates. This approach has also been used for influenza vaccine evaluation by the Food and Drug Administration (e.g., see https://www.fda.gov/media/135687/download page 27).

We have added discussion to highlight the above considerations (lines 476-485); see also lines 232-238.

Common response #3: What was the health status of study subjects? Did they receive any COVID-19 vaccines or have any other infections prior to influenza vaccination?

Please see lines 86-90 in the revised manuscript. Both the COVR and healthy control (HC) groups did not receive any COVID-19 vaccine (none were available back in the Fall of 2020), did not receive any recent vaccines before the influenza vaccination we administered, and did not have autoimmunity, obesity, or active infection. Prior to each study visit, we collected information about any new medical illnesses, medications, vaccinations or infections since the last study visit/timepoint. Review of this information showed zero reported viral infections or vaccinations between study enrollment and day 28 of the study.

We agree that information about participant infections between COVID-19 and influenza vaccination could be useful information to add to the analysis. Although we did not collect this information at the time of study enrollment, the closure of schools and workplaces in March 2020 followed by nonpharmaceutical interventions (e.g., masks) to reduce the transmission of SARS-CoV-2 resulted in an extremely steep drop in both influenza infection and influenza-like illnesses from March – December 2020[4]. Hence while it is possible that some study participants were infected with other pathogens during the period between acute COVID-19 (after March 2020) and study enrollment, public health surveillance data suggests the likelihood is low. As discussed above in our response to common question #2, our study design includes a matching healthy group drawn from the same geographic region, which should help control for the extent of potential infection encounters in the area.

Please note that as our data demonstrate, while COVR subjects appeared clinically healthy (aside from residual symptoms such as anosmia in some individuals – see Extended Data Table 1), the molecular and cellular states of their immune system and their functions can be different from those of matching healthy controls who never had COVID-19.

Point-by-point responses to reviewers

Referee #1:

We initial read with enthusiasm the title and abstract of this paper and we commend the authors on the amount of work that was put into this study. Unfortunately, there are too many major issues with the analyses that dampen enthusiasm.

We thank the reviewer for the enthusiasm and appreciation for our work. We apologize for any confusion caused by the lack of clarity of the previous presentations of our data and analyses. As discussed in the common response above, in addition to addressing the reviewers' comments with new data and analyses (especially multimodal single cell data and analyses), we have streamlined and substantially improved the presentation to clarify and better highlight our most salient findings.

Conceptual
The authors do not adequately justify why they chose influenza vaccination as an immune challenge or address confounding that could have been caused by pre-existing immunity to influenza. As is well-documented in the literature, prior exposure to influenza through either infection or vaccination is nearly universal in middle-aged individuals, leading to heterogeneous levels of pre-existing immunity, which is a fundamental predictor of the response to vaccination. Despite some effort to statistically correct for vaccination history (which is never addressed in the main body of the text), the authors do not adequately address this.

We thank the reviewer for raising these important questions. Please see common responses #1 and #2 above.

For example, the low percentage of 'responders' in figure 2B is likely due to high levels of pre-existing immunity precluding a measurable 4-fold change.

We agree that the level of pre-existing (pre-vaccination) antibody titers against influenza is known to correlate negatively with the fold-changes in titers following vaccination. Please see common response #2 above that addresses this issue.

We have since removed previous Fig. 2b for clarity since the stacked bar plot was difficult to understand based on the feedback of multiple reviewers; in addition to the improved presentation on the humoral responses (see current Fig. 2g-i, Extended Data Fig. 3d-g), for comprehensiveness the raw titer values and their fold-changes are included in Extended Data Table 9 and Fig. 2h-i.

While one justification for using influenza as an immune challenge could be that the response to influenza vaccination is very well-characterized in the literature, the authors transform the data in such a way that it cannot be directly compared to the literature. As a result, we think that the interpretations of transformed data by the authors are not completely accurate and become so difficult for the reader to assess that we found ourselves questioning the validity of many statements by the authors.

We apologize for the confusion. The antibody titer data was generated in a standardized and rigorous process using a well-described microneutralization assay performed by our US Food and Drug Administration (FDA) co-authors. This assay has been used extensively for vaccine and infection studies, including clinical trials (e.g., see refs[5–9]).

To improve clarity and address the reviewer's concern, we now present the fold-changes (D28/D0) and the baseline (D0) antibody titers for each vaccine strain (Fig. 2h,i and Extended Data Fig. 3e,f; raw antibody titer data can be found in Extended Data Table 9). We use both antigen-specific cellular data (i.e., influenza-specific plasmablasts; Fig. 2g) and antibody titer data to arrive at the conclusion that COVID-19-recovered males (COVR-M) had elevated humoral responses to the seasonal influenza vaccine. Finally, as mentioned above in common response #2, we adopted a statistical model for evaluating antibody responses (also used by the Food and Drug Administration to evaluate vaccine trials) that includes age and the pre-vaccination influenza antibody titer as covariates.

The updated presentation and results are consistent with the results shown in the previous manuscript using the "transformed" metric (as referred to by the reviewer), which was a measure of the maximum response among all four influenza strains in the vaccine (after scaling/standardization across the strains as their titer values could span different scales and dynamic range – data now shown in Extended Data Fig. 3e, f; see also Methods). We and others found in earlier studies that this approach of using the maximum titer response among the strains was helpful for dissecting early cellular and transcriptional correlates, including blood transcriptomic and single cell surface protein and gene expression data[1,2,10,11], which are often measured in an antigen-agnostic manner.

Text-specific

• Line 15: [5th-95th percentile:58 -235] days after diagnosis need clarification. It is a little confusing what the authors meant '5th - 95th percentile' and range of days.

We have added a figure (Extended Data Fig. 1a) with the distribution of the days since diagnosis and have removed the language about the 5^{th} – 95^{th} percentile to avoid confusion.

• Line 20 -21 - Authors categorized CD71hi B cells as a subset including memory B cells and plasmablasts (as described in Supplementary Figure 2B). Authors need further clarification and/or explanation why CD71hi B cells (including influenza-specific subsets) increased before vaccination in COVID-19-recovered male subjects. Other very well-respected vaccine immunologists define IgD- and CD71- as memory B cells in the context of influenza vaccination (doi: 10.1038/ni.3533).

As shown in previous Supplementary Figure 2b, the $CD71^{hi}$ memory B-cells we gated are IgD- (to exclude naïve cells), CD38- and $CD20^{hi}$. This excludes plasmablasts (which are $CD38^{hi}$ and CD20-), leaving the $CD20^{hi}CD71^{hi}$ memory B-cells only. The (Ellebedy *et al*, 2016) paper mentioned by the reviewer is indeed informative and one that we previously cited as motivation for assessing memory B-cells with a $CD71^{hi}$ phenotype. Ellebedy *et al* define a population of IgD-$CD20^{hi}CD71^{hi}$ B-cells as activated B-cells (ABCs) that emerge after influenza vaccination or natural infection with influenza or Ebola virus. This B-cell subset is distinct from antibody-secreting (ASCs) and naïve B-cells, and remains detectable after ASCs disappear in blood, especially in natural infections. These ABCs also have elevated mRNA expression of antigen presentation genes (including both HLA class I and II genes) compared to ASCs and naïve and memory B-cells (MBCs). We used our single cell CITE-seq data to confirm that HLA genes are indeed elevated in our $CD20^{hi}CD71^{hi}$ memory B-cells compared to other B-cells in our cohort (see table below), thus further supporting that our cells phenotypically resemble those ABCs defined in the Ellebedy study. This also gives us additional confidence that these cells we identified are not classical MBCs.

Table R1: differential expression of HLA transcripts and surface proteins between $CD20^{hi}CD71^{hi}$ memory B-cells compared to other B-cells.

	CITE-seq molecule type	Adjusted p value	Average log2FC	% expression in $CD71^{hi}$ B-cells	% expression in other B-cells
HLA-A	mRNA	2.48E-156	0.44	99.8	99.3
HLA-C	mRNA	4.24E-107	0.37	99.8	99
HLA-B	mRNA	8.55E-80	0.30	100	99.9
HLA-F	mRNA	5.07E-46	0.26	62.8	45.3
HLA-DPB1	mRNA	4.27E-44	0.29	99.1	99.2
HLA-DQA2	mRNA	9.95E-15	0.30	49.4	41.4
PROT-HLA-ABC	surface protein	0	3.15	99.3	98.8
PROT-HLA-DR	surface protein	4.45E-14	0.34	88.5	93.1

It is a compelling question as to why the $CD71^{hi}$ memory B-cells (including the influenza-specific subsets we define) were higher in COVR-M at baseline before influenza vaccination. Our hypothesis is that in addition to antigen-specific ABCs (as shown in the Ellebedy paper), "bystander activation" following viral infections such as SARS-CoV-2 could result in the expansion of memory B-cells with specificities beyond the infecting pathogen (e.g., including influenza and potentially other specificities reflective of past exposures). Since males tend to have a more inflammatory responses during acute COVID-19, bystander activation of cells could have been more extensive compared to females. Antigen non-specific, bystander B-cell activation has been previously reported with other viral infections[12]. Although beyond the scope of our paper, this phenomenon of B-cell bystander activation is an area of active research.

Given that bystander activation of B-cells is not a focus of our revised paper, we have removed the $CD71^{hi}$ related observations in the text and figures.

• Line 67: Authors need to identify the time of symptom assessment. Also, the number of total subjects and the percentage are important.

We have added information to the Methods section to clarify that persistent symptoms of past SARS-CoV-2 infection were documented at the time of study screening/enrollment. We have also added percentages to Extended Data Table 1.

• Line 77: Was there any significant difference between male and female subjects.

From our TSD association model (see Methods), we did not detect a significant difference ($p = 0.865$) between male and female subjects for the association of SARS-CoV-2 titers to time since diagnosis.

• Line 82: the red cell distribution' should be 'the red blood cell distribution'.

We thank the reviewer for catching this error. Given their peripheral nature to the main theme of our work, we have removed the analyses of the red blood cell distribution width (RDW) from the revised manuscript.

• Line 82 - 84: The normal range of RDW is from 11.8 to 15.6 percent, which tends to increase with age. The authors showed the elevation and decrease of RDW in COVID-19 subjects, but the changes are within the normal range. In this case, the authors should have a healthy control (HC) group as a baseline.

Line 84 -87: The decrease appears to be 1.5 -2 % which is within the normal range. Thus, the data do not support the statements made.

We agree that these values of the COVID-19 recoverees are within the normal range; however, the negative correlation with time since diagnosis in the COVID-19 group is significant (after accounting for age in the regression model when both sexes are included or in males alone), suggesting that RDW continues to evolve after acute disease (previous Fig. 1e). However, we have removed the analyses of the RDW from the revised manuscript given their peripheral nature.

• Line 88 - 92: The decrease of plasmablast over time is a normal immunological reaction, which is not specific to COVID-19. The authors should have looked at plasmablast between Day 0 and Day 20, which is more sensitive to demonstrate the change in plasmablasts (Extended Figure 1 a -c).

Previous Extended Data Fig. 1b and 1c showed negative correlation between plasmablast frequency (and corresponding gene module score) and time since COVID-19 diagnosis, suggesting that plasmablast frequencies decreased over time post-COVID-19. The plasmablast frequencies were measured at baseline enrollment in the study (i.e., prior to seasonal influenza vaccination). We agree that this decrease over time is expected as a sign of recovery given that plasmablasts tend to be elevated during acute COVID-19.

Independent of the above, the other plasmablast data that we show is the response following seasonal influenza vaccination, including the frequencies of antigen-specific cells (new Fig. 2g and Extended Data Fig. 3d.)

To avoid potential confusion and given that the observation concerning plasmablast frequencies before influenza vaccination is not an important point of our manuscript, the previous Extended Data Fig. 1a-c are now removed from the revised manuscript.

• Line 113 -117: COVR-M has a higher innate immune activation and metabolic process than COVR-F. But T cell surface signature and T cell differentiation in COVR-M are lower than in COVR-F. These data suggest that the adaptive immune response might be better in COVR-F than in COVR-M, which is different from the conclusions drawn by the authors.

This impression is mainly due to the confounding of whole blood transcriptomic results by cell frequency. The blood transcriptomic enrichments cited by the reviewer were based on bulk blood RNA expression analysis, which can be driven by a combination of changes in cell composition and transcriptional changes within cell types. The higher innate immune activation enrichment in COVR-M could be attributed to the elevated frequency of monocytes (see new Fig. 1d and Extended Data Fig. 1e). The lower “T cell surface” and “T cell differentiation” signatures in COVR-M could be driven by the correspondingly lower frequency of T-cells in blood because the myeloid components of PBMCs (including monocytes) are higher in COVR-M. Indeed, these enrichments cited by the reviewer are no longer statistically significant once peripheral blood monocyte % was included in the statistical model (see lines 122-127; data not shown.)

Given the potential confounding of cell frequencies, in our revised manuscript we use the bulk transcriptomic data only as a first step to suggest which biological processes might be relevant. We then further analyze these processes using single cell CITE-seq data to assess cell type-specific differences. This approach revealed, for example, that CD8+ effector memory (EM) T-cells have an elevated activation signature and monocytes have a repressed immune receptor signature in both COVR-M and COVR females (COVR-F) relative to their healthy counterparts (Fig. 1f-i). However, there were also sex specific differences, including a subset of CD8+

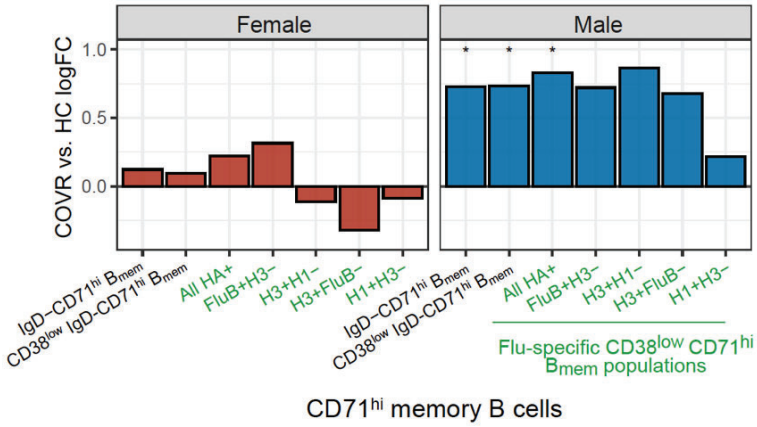
EM T-cells with elevated frequency in COVR-M at baseline (Fig. 3e) and these cells had a significant increase in IFNG transcript levels following influenza vaccination (Fig. 3f and Extended Data Fig. 4e and 4f). These cells thus contributed to the increased IFNγ-related responses in COVR-M (Fig. 2). See also the "Overview of our revised manuscript" section in the common response above for additional information.

• Line 150 - 157: The author should elaborate on the meaning of 'true' and 'false' in Figure 2b. It is unclear whether the authors combine the titers of different strains. and how the normalization was processed. In healthy subjects, the difference between male and female subjects is present.

We agree with the reviewer that this figure is confusing. We have replaced previous Fig. 2b with plots (new Fig. 2h and 2i) that are easier to interpret. We now show the raw day 0 titer data in Extended Data Figs. 3e and 3f and associated statistics in Extended Data Table 9. As shown in Extended Data Fig. 3e and 3f, the influenza strain-specific neutralization titers at baseline (D0) are similar between COVID-19-recovered individuals and age- and sex-matched HCs. We have de-emphasized evaluation of "high" vs "low" responders as we believe that the raw data provides a fuller and clearer picture of the humoral response to the vaccine.

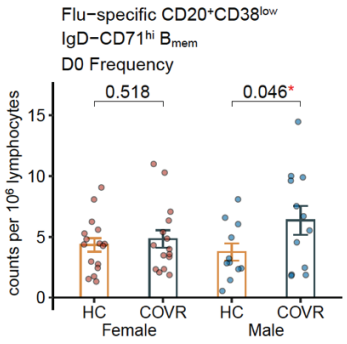
• Line 160-163: Additional clarification is needed for a better understanding of sex difference in Flu-specific memory B cell response. Influenza A -specific B cells (H3+H1-, H3+Flu-B-, H1+H3-) in HC show a higher fold change than those in COVR-F. Influenza B specific memory B cells showed the opposite trend. Also, all memory B cells in males show higher levels than those in females. The representation of the data is so transformed that it was particularly difficult to interpret.

To clarify, those lines (160-163 in the previous manuscript) and the corresponding figure (previous Fig. 2e – copied below) were referring to the differences in the frequency of CD20+CD71^{hi} memory B-cells between COVR-M vs. HC-M and COVR-F vs. HC-F at baseline, prior to influenza vaccination (shown as log fold-changes below).



*Previous Fig. 2e. Bar plots showing the fold change in D0 (prior to influenza vaccination) frequency of CD71^{hi} memory (CD71^{hi}IgD-) B-cells (as fraction of lymphocytes) as well as influenza-specific cells (with green label) between COVR and HC subjects, separately for females (left) and males (right). Significance was determined by linear regression models accounting for age, race, and vaccination history. * p ≤ 0.05.*

The raw (untransformed) frequencies of CD71^{hi} memory B-cells including the influenza-specific (all HA+) subsets are significantly higher in COVR-M than HC-M prior to influenza vaccination. See figure below:



Box plot showing the counts (per million lymphocytes) of influenza-specific CD71^{hi} memory (CD20+CD38^{low}CD71^{hi}IgD-) B-cells from flow cytometry (y-axis) for COVR-F (n = 15), HC-F (n = 16), COVR-M (n = 12), and HC-M (n = 11) at D0, including subjects aged 65 and above. Significance of differences is determined by linear regression models accounting for age, race, and vaccination history (see Methods).

In our updated manuscript, we have removed the CD71^{hi} memory B-cell results given their peripheral nature to the main messages of the manuscript.

• Line 310: Authors stated that the adaptive immune response is elevated in COVR-M. However, supportive data were not presented in the manuscript.

The adaptive responses we referred to include influenza-specific plasmablast (day 7 following vaccination) and antibody titer (day 28) responses (see current Fig. 2g-i). Please also note that in the revised manuscript we now present results on both elevated early (day 1) innate (IFNγ-related) and adaptive (humoral) responses in COVR-M in Fig. 2.

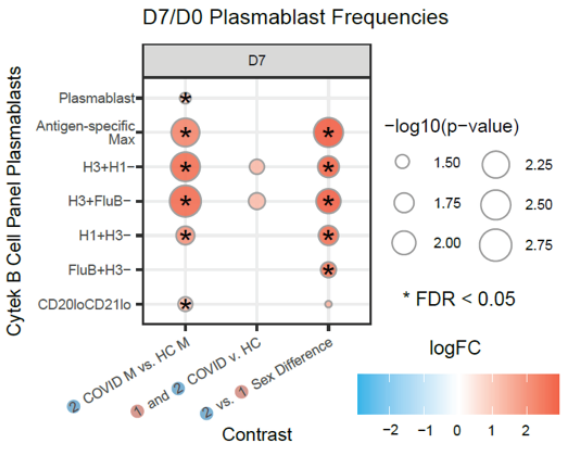
• Line 339: Authors stated the plasmablasts increased in COVR-M. However, Figure 2b shows the comparison of frequency (%) [we believe that the reviewer is referring to previous Fig. 2d]. We suggest that the comparison should be conducted using ratio, not the frequency.

• 1Figure 2d: It appears that the authors subtract frequency on Day 0 from Day 7. Subtraction of % (frequency) should be replaced with actual cell count numbers. The strain of influenza should be provided. The authors stated 'Flu-specific plasmablast' in Figure 2d. Sex differences in flu-specific plasmablasts must be evaluated.

Previous Fig. 2d showed the difference between the day 7 and day 0 plasmablast frequencies, which are a fraction/ratio of the plasmablasts over total lymphocytes [in addition, we used a summary metric (maximum frequency among the influenza strains)]. We have replaced this with Fig. 2g in the revised manuscript, which explicitly shows data from both days 0 and 7 to illustrate the change in influenza-specific plasmablast frequency following influenza vaccination. The actual cell count numbers (per 1000 B-cells) are shown as recommended by the reviewer.

As can be seen in new Fig. 2g, the frequency of flu-specific plasmablasts at day 0 is nearly zero and can potentially be influenced by noise. Hence the day 7/day 0 ratio (fold-change) would be extremely high and may be difficult to interpret. We believe that our new Fig. 2g, which follows the approach often used in vaccinology papers, is clearer and more interpretable than the old Fig. 2d.

The strain-specific plasmablast frequency comparisons (day 7/day 0) are shown below (previous Extended Data Fig. 3d). These data are consistent with what we show in new Fig. 2g.



Bubble plot showing differential levels of the plasmablast frequencies at day 7 compared to day 0. Each column represents a specific comparison and each row is a plasmablast population, including the H1+, H3+, and B+ flu-specific plasmablasts, and the maximum of those flu-specific populations ("Antigen-specific Max"). A population is included if the difference is significant ($p < 0.05$) in at least one of the comparisons shown. The color of each bubble indicates the effect size of the comparison and its size represents the degree of statistical significance.

The sex-differences requested by the reviewer are also available in the Extended Data Table 8.

Figure-specific

• All stacked bar graphs (1B, 2B, 4F) are difficult to interpret and do not support the conclusions made in the text.

We thank the reviewer for this feedback and agree the stacked bar graphs are confusing. We avoided using the stacked bar graphs in the revised manuscript. They are either no longer needed given our revised narrative or we referred to the test statistics directly in the text (e.g., lines 92-94).

• 1D/E: Unclear why correlation plots here. There is a clear dependent (nAb/RBC) and independent variable (TSD). Why not use regression analyses with a time by sex interaction term, which would allow direct testing of the sex different in trends? Also for RBC where are control patient data?

Regarding previous Fig. 1e (relationship between RDW and TSD), we have removed the analyses of RDW from the revised manuscript given their peripheral nature.

Our manuscript previously included the analysis suggested by the reviewer (Extended Data Table 2), where we used a regression model to test for association between SARS-CoV-2 antibody titer or RDW with TSD while accounting for age and race (see *Association with time since COVID-19 diagnosis* section in Methods for additional details). We apologize for not including the SARS-CoV-2 titer association results in that version of Extended Data Table 2. The TSD association results for the RDW and SARS-CoV-2 titer are listed below as a reference:

Table R2: Regression model results for association of SARS-CoV-2 (WA-1) titer and RDW with time since diagnosis (TSD).

Group	Dependent variable	Model TSD coefficient	Effect size	P value
COVR-M	WA-1 IC50	-5.25246	-2.16454	0.035648
COVR-M & COVR-F	WA-1 IC50	-4.77466	-2.3791	0.02156
COVR-F	WA-1 IC50	-4.29687	-1.39148	0.170772
Sex difference	WA.1_IC50	-0.95558	-0.24892	0.804529

Group	Dependent variable	Model TSD coefficient	Effect Size	P value
COVR-M & COVR-F	RBC.Distribution.Width %	-11.9886	-3.47999	0.001722
COVR-F	RBC.Distribution.Width %	-14.0961	-2.6737	0.012463
COVR-M	RBC.Distribution.Width %	-9.88101	-2.34404	0.027075
Sex difference	RBC.Distribution.Width %	4.215094	0.637937	0.52895

As shown the tables above, the sex difference is not significant for either SARS-CoV-2 (WA-1) titer or RDW. However, please note that we were not trying to explicitly test for or show sex differences in the previous Fig. 1e, but to show that these two parameters (SARS-CoV-2 antibody titer and RDW) are correlated with TSD.

• 2B: typo in Male panel, says that OR =0.

Because there were no healthy male high responders, the OR is zero. However, we have removed this stacked bar plot to improve clarity. We now show the titer response data in Fig. 2h,i as fold-changes (day 28/day 0).

• 2C: unclear if statistics account for that fact that the timepoints are not independent of each other. I do not think this panel was addressed in the results section.

We agree with the reviewer that previous Fig. 2c is confusing, and that the post-vaccination titer can be influenced by and correlated with the baseline titer level.

We have replaced this figure with the day 28/day 0 fold-changes separately for each of the four strains (Fig. 2h,i). In addition, the p-values shown are from regression models that include day 0 titers as a covariate (see also common response #2 above). We further show that the day 0 influenza titers were similar across the groups at day 0 (Extended Data Fig. 3e,f). Together these data show that the antibody responses to influenza vaccination are higher in COVR-M than HC-M.

• 3: unclear how samples at day 1 OR day 7 were used. Did the same individuals contribute samples at both timepoints? Or did participants contribute samples at either day 1 or day 7? Time-course analyses (C and E) suggest that these are the same people over time. Could small sample sizes in at day 7 explain the dips seen at this timepoint in C and E?

All enrolled individuals were followed longitudinally with samples collected at days -7, 0, 1, 14, 28, 70, and 100, followed by optional monthly visits (see Fig. 1b). All subjects contributed samples to day 7 and the sample sizes were the same at each timepoint (except for one subject who left the study before day 28 and thus does not have day 28 titer data). We have added additional clarifying text to the manuscript when introducing Fig. 1b (lines 85-90), and in the Fig. 1b legend.

• Figure 3b-d: Figure 3b (transcriptomic analysis) shows little difference in interferon-gamma response between COVR-F and COVR-M. However, Figures 3c and d show that significantly higher IFN-gamma response in COVR-M compared to that in COVR-F. Additional information would be helpful for a better understanding.

We regret an error in the previous legend stating that the FDR cutoff was 0.05 in previous Fig. 3b; the actual FDR cutoff used was 0.01 and thus the enriched gene sets with an FDR between 0.01 and 0.05 were not shown previously (we used the 0.01 cutoff due to space limitation.) The Hallmark Interferon Gamma Response was indeed enriched in COVR-M compared to COVR-F (third box labeled "IFN signaling" in previous Fig. 3b: "2 vs. 1 sex difference") with an enrichment score (NES) of 3.38 and FDR of 0.04 (therefore missed the 0.01 cutoff). We have modified the panel (new Extended Data Fig. 3a) to use 0.05 as the FDR cutoff as we had originally intended. In addition, these results are consistent with current Fig. 2b-e, including the new single cell CITE-seq data showing that the IFNγ response signature in major immune cell types is higher in COVR-M than other groups, including COVR-F.

• Figure 3e: While interesting, absolute monocyte counts presented in this figure are confusing compared to Figure 1g data (frequency). In Figure 1g, the frequency of monocytes was similar between HC_M and COVR-M on Day 0 whereas a marked difference was presented in absolute monocyte counts between HC-M and COVR-M in Figure 3e. While the frequency and absolute count are different measures, additional description and explanation would be helpful.

We agree that this is confusing. To avoid confusion, we have removed previous Fig. 3e and no longer highlight this peripheral point.

• Figure 4d: It is unclear how many subjects were included in the gene heatmap. HC appears to have 3 subjects and COVR appears to have 8 subjects. However, there seems a notable variation within the group.

We apologize for not including the number of subjects in previous Fig. 4d. We have replaced this figure with data from our expanded CITE-seq analysis. We have also included the number of subjects with data for each assay type in Fig. 1b.

• Figure 4c: If a sex difference is present in this case, then separate figures for males and females should be included.

We have since removed Fig. 4c and replaced this figure with data from our expanded CITE-seq analysis with males and females shown separately as suggested by the reviewer (see Fig. 4b-e and Extended Data Fig. 5).

• The entirety of Figure 4 and the use of a new population of people from Italy without having the HC day 1 control group, which makes all of these comparisons confusing and, in some cases, incorrect.

We apologize for this confusion. We have simplified Fig. 4 with data from the expanded CITE-seq analysis. We now show the data from Brescia, Italy earlier in Extended Data Fig. 1i. The goal of the analysis using data from the independent Italian cohort of acute COVID-19 remains the same: to suggest that the depressed immune receptor and signaling signature in monocytes (see new Fig. 1f, g) in COVR-M and COVR-F could have originated from the acute response during COVID-19 (see lines 150-155).

Referee #2:

Sparks et al analyzed the long-term impact of mild SARS-CoV-2 infection on immune status at baseline and following seasonal Flu vaccination. The study identifies shifts in immune signatures between healthy and following mild COVID19 that are different among sexes. In terms of effects of mild COVID19 on baseline immune parameters, sex differences were observed with increased monocytic, oxidative phosphorylation and fatty acid metabolism in males vs increased T cell processes in females. Following influenza vaccination, a proportion of males had higher Ab titers at day 28 post vaccination than females, together with higher frequency of flu-specific B memory populations and plasmablast frequency. Vaccination induced shifts from baselines in a sex-dependent manner, with males showing significant up-regulation of IFN signalling and IFNγ transcriptomic responses and IFNγ levels at day 1 following vaccination when compared to females. Other features that were differentially observed between the sexes were increased immune activation in males and increased neutrophil, NK and T cell modules in females at day 1 post vaccination, and increased B/plasmablast and neutrophil signatures in males at day 7 post vaccination. Finally, the authors identify a depressed immune state in COVR that is partially resolved upon vaccination, with genes relating to oxidative phosphorylation, monocyte and cell cycle processes partially returning to baseline in females, and processes relating to plasmablasts returning to baseline levels in males after 28 days post vaccination. Vaccination induces transcriptional changes in classical monocytes at day 1, similar to baseline. Interestingly, a larger fraction of females had genes whose expression reset by day 28 when compared to males suggesting females are able to return to closer to a baseline more effectively than males.

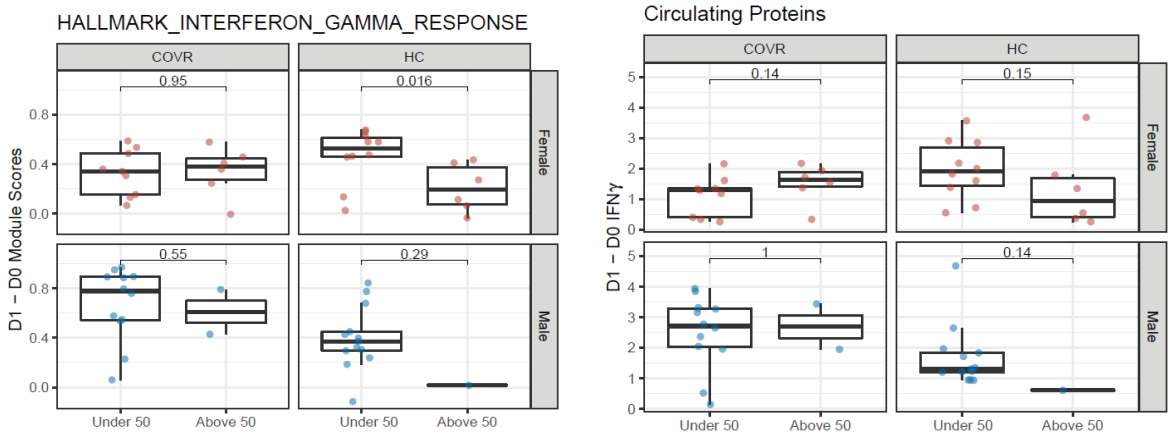
This manuscript offers several interesting new perspectives on immune recovery after acute infection in general and COVID-19 in particular and thus opens up several new avenues for future studies.

We thank the reviewer for their appreciation of our study and findings, e.g., that they open new avenues for future work.

General comments and suggestions
- Subjects with variable ages were included in the study. Because sex differences were observed for many features, it would be interesting to indicate how sex differences occur when comparing subjects pre and post andropause/ menopause; if data is not available, perhaps separating by age groups.

We thank the reviewer for this interesting question. While we did not collect andropause/menopause status data, per the reviewer's suggestion, we investigated the Day 1 IFNγ response (which shows major differences between COVR-M and COVR-F) between subjects under and above 50 years of age. We found that in healthy females (HC-F), the responses trend lower in those older than 50 (see figure below). This trend is not apparent in the other groups, although this could be due to lack of statistical power.

We agree with the reviewer that in general that it would be informative to assess the effects of andropause/menopause, but we likely do not have sufficient statistical power to divide the cohort into two groups based on age (as shown in the IFNγ example here). We did, however, include age as a continuous covariate in our statistical models to better delineate differences among the groups after accounting for age effects. Please also note that because the vaccination responses in older subjects/elderly can be substantially different from younger individuals[2], we excluded participants 65 or older from our analyses.



Box plots showing the distribution of day 1 response (D1 – D0; y-axis) in (left) Hallmark IFN Gamma Response gene set module score and (right) serum IFNγ protein level from the OLINK platform in COVID-19-recovered and healthy control males and females aged <50 years and greater than or equal to 50 years. Significance of group difference is determined by two-tailed Wilcoxon test.

- More insights on how sex differences in immune and metabolic findings relate to one another?

Sex differences in metabolic-related transcriptional signatures, such as oxidative phosphorylation, fatty acid metabolism, and MTOR, revealed by whole blood transcriptomics were not prominent at the individual cell type level using new CITE-seq single cell data. This suggests that these may be related to cell frequency differences. Indeed, these signatures are no longer significant after accounting for the frequency of major cell subsets in blood (neutrophils, CD4, CD8, B-cells, and monocytes; data not shown). As discussed above, in the revised manuscript we use the bulk transcriptomic data only to point us toward what signatures to examine further in a cell type-specific manner using the single cell data (the bulk data are now shown in Extended Data Fig. 1 instead of Fig. 1). We have removed the mention of metabolic-related sex differences in the revised manuscript.

- Fig 2: More discussion is warranted on why male influenza antibody titers are higher in COVR vs HC in males while the opposite happens in females? Any difference in prior immunization history between groups?

To clarify, there are no significant differences between COVR-F and HC-F in either day 7 influenza-specific plasmablast frequencies or day 28 titer increases (new Fig. 2g,h).

For COVR-M, they have elevations in both the innate (Fig. 2b-f) and humoral (Fig. 2g-i) responses compared to HC-M. The higher plasmablast response (day 7) is a key contributor to the higher antibody titers (day 28). The plasmablast response is in turn correlated with the IFNγ response on day 1 (new Extended Data Fig. 3d). In Fig. 3 we use CITE-seq single cell data to identify a population of GPR56+ effector memory CD8+ T-cells as a source of the increased day 1 IFN response in COVR-M relative to HC-M. We have added discussion to suggest that the higher inflammatory response (e.g., higher circulating cytokine levels) during acute disease in males relative to females might have led to the higher frequency of these cells in COVR-M (lines 440-444).

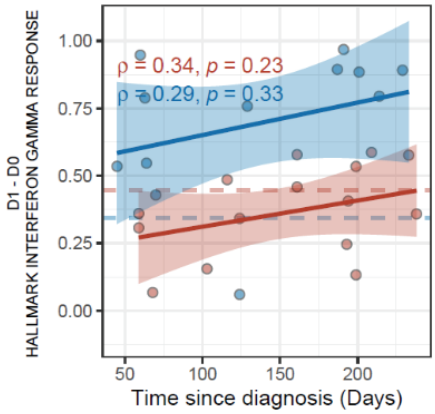
We now include self-reported prior influenza vaccination history data in Extended Data Table 1. There is no significant difference among the groups (Kruskal-Wallis p= 0.95). The vaccination history was also included as a covariate in our statistical models for assessing antibody titer responses. Additionally, antibody titers at baseline (day 0) are similar among the groups (Extended Data Fig. 3e, f). We have also included baseline influenza antibody titers as a covariate in our statistical model. See common response #2 for additional details.

- Were there reported days following SARSCoV2-specific vaccination and/ or differences between the sexes and groups of COVR vs HC?

See common response #3. This study was conducted in Fall 2020, prior to the widespread availability of COVID-19 vaccines. No one in our cohort had received COVID-19 vaccination at the time of the study. We have revised Fig. 1b to highlight both the calendar season/year during which the study took place and that no subjects had received the COVID-19 vaccine. We have also added to the Methods that no study subjects had been vaccinated against COVID-19 at the time of our study.

- Some of the individuals were samples quite recently after a prior COVID-infection. Is there any difference in relation to time from prior COVID in these subjects?

Indeed, our COVR cohort spans a large interval since their COVID-19 diagnosis including some that recovered relatively recently from the infection, but there were no detectable difference in the time since diagnosis (TSD) between male and female recoverees (see new Extended Data Fig. 1a). This large range in TSD helped us identify parameters that were potentially still evolving following recovery from COVID-19, and enabled us to focus our analyses and results on parameters that are stable at baseline (i.e., they do not correlate with TSD). We have also determined that the vaccine response (e.g., day 1 IFNγ increase) is not linked to the TSD (see figure below); the same is true for the baseline level of GRP56+ CD8+ effector memory T-cells (new Extended Data Fig. 4d).



Scatterplot showing the correlation between D1 – D0 Hallmark IFN Gamma Response gene set module score (y-axis) and the time since diagnosis (x-axis) for COVR subjects. Dashed lines show the median frequency for the healthy control groups. Spearman's rank correlation and p values are shown.

- The observation of CD71 hi B-cells after COVID-19 is curious and I wonder whether any particular plasma protein correlated with the abundance of these cells?

Per the suggestion of the reviewer, we analyzed correlation between serum proteins and influenza-specific $CD71^{hi}$ B-cells. While no proteins passed the FDR < 0.05 cutoff, listed below are the most significant associations.

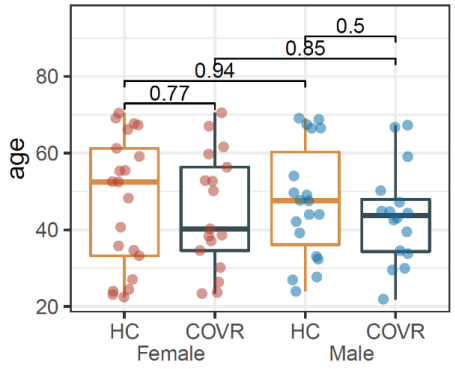
Table R3: Association between OLINK serum protein levels and CD71^{hi} B-cell frequencies in the COVR-M group at day 0 (before influenza vaccination).

group	protein	association coefficient	std error	statistic	P Value	FDR
COVR-M	CXCL12	-0.93225	0.312356	-2.98459	0.004718	0.16631
COVR-M	PLAU	-2.5491	0.837909	-3.04221	0.004038	0.16631
COVR-M	TNFSF12	-0.74565	0.229666	-3.24664	0.002297	0.16631
COVR-M	VEGFA	1.039663	0.34562	3.008108	0.004428	0.16631

The positive correlation between the CD71^{hi} B memory cells and VEGFA in COVR-M is potentially interesting. VEGFs have been shown to be elevated in COVID-19 infection[13]. This association is consistent with our hypothesis that the CD71^{hi} B-cells may be activated in a bystander fashion, possibly secondary to endothelial activation in the lungs.

In our updated manuscript, we have removed the CD71^{hi} memory B-cell results given their peripheral nature to the main messages of the manuscript.

Minor points:
- Clarify age range and distribution among groups
We have included this information in Extended Data Table 1. In our study design we matched the age and sex between the COVR and HC groups to the best extent possible (see plot below; Wilcoxon test p values shown).



Box plot of age distribution in COVID-19-recovered and healthy control males and females. Significance determined by two-tail Wilcoxon test.

- Clarify in text whether average time from diagnosis was accounted for in analysis.

In our analysis, we distinguish between evolving vs. stable parameters (e.g., the immune parameters we present in Fig. 1 are stable). The former is defined as those that exhibit significant correlation with time since diagnosis (TSD) at baseline (before influenza vaccination). The focus of the paper is on the stable parameters (i.e., those that do not correlate with TSD) and how those (e.g., GPR56+ effector memory CD8+ T-cells) may impact responses to influenza vaccination. Most of our analyses (including the CITE-seq analyses) focus on the stable parameters at baseline.

In our statistical models involving healthy controls (who never had COVID-19 at the time of analysis), the TSD was not explicitly included as a covariate because TSD is undefined for healthy controls who never had COVID-19 at the time of our study.

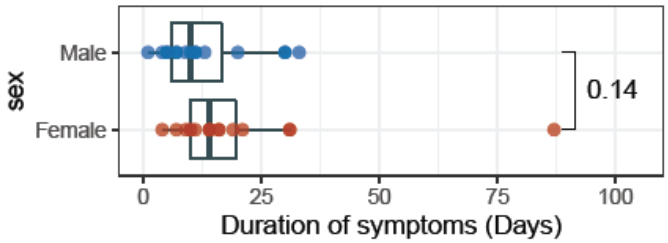
We have now added text to clarify these points (lines 106-113).

- Specify range of self-reported average length of illness 20.8 days and whether counted for in analysis

The length of illness was not accounted for in our analyses because:

1. We could not include the length of illness in models that used the healthy control group as a reference as the length of illness is undefined for healthy controls.
2. We are not confident that self-reported average length of illness was quantitatively accurate, particularly for individuals who were the farthest out from their illness.

Below we show that the self-reported length of illness is not significantly different between the COVR males and females.



Box plot of COVID-19 symptom duration in COVID-19-recovered males and females. The p value is shown on the right with significance determined by the two-tail Wilcoxon test.

- Clarify Figure 1b legend in terms of Y N bars relating to residual symptoms

As reviewer #1 also pointed out, the stacked bars are confusing. We have removed this (and all other) stacked bar graphs in the revised manuscript and referred to the test-statistics in the text directly (see lines 92-94 and Extended Data Table 1).

- Clarify in text that Fig 1d only show significant anticorrelation between time of diagnosis and neutralizing antibody titers in males: Lines 78, 93, 94

We have changed the language to improve clarity (lines 106-110).

- Clarify normal range for RDW in Fig 1e and comparison with uninfected males and females

Regarding previous Fig. 1e (relationship between RDW and TSD), we have removed the RDW results from the revised manuscript due to their peripheral nature.

- Line 109 conclusion based on males, but Figure 1g suggests no changes in monocytes in males between healthy vs COVR, whereas females have decreased % monocytes.

We apologize that the sentence (line 109) in the original manuscript was confusing. Please see current Fig. 1d and lines 116-119. In summary, monocyte frequency (%) is significantly higher in COVR-M than COVR-F (trending in the same direction between COVR-M and HC-M). Monocyte frequency is lower in COVR-F than HC-F.

- Information/ differences in vaccinations between subjects who got covid/ did not?

None of the subjects in the study had received the COVID-19 vaccine (including any experimental vaccines) or other recent vaccines. We have added clarification to both the Methods and Fig. 1b. We have also evaluated the influenza vaccination histories between groups and there is no significant difference between groups (Kruskal-Wallis p= 0.95). The vaccination history of the participants can now be found in Extended Data Table 1.

See also common response #3.

- Verify Fig 2b male OR=0?

Because none of the healthy males were high responders, the OR is zero. However, we have removed this stacked bar chart in the revised manuscript to improve clarity.

- Fig 2e indicates higher FC COVR vs HC in males but not females; are there differences in baseline values (rather than FC) between males and females?

We now show the baseline titers for all four vaccine strains in the new Extended Data Fig. 3e,f. There is no significant difference in the baseline titers between COVR-M and HC-M. For the A/Hong Kong/2671/2019 strain, COVR-M actually trends higher than HC-M at baseline, which could translate to lower FC given that baseline titers are known to negatively correlate with FC. We have also included the baseline titer as a covariate in our statistical model for evaluating antibody response differences, so the p values shown in the new Fig. 2h,i already account for the baseline titers.

See also common response #2.

- Figure 3 indicates that vaccine induced shifts from baseline states are quite between sexes; hard to interpret

In current Fig. 2 we show that the vaccine induced shifts from the baseline, including both innate and adaptive responses, are different between the sexes. See also Fig. 3 in which we identify cells at baseline (prior to vaccination) linked to the elevated IFN response in COVR-M.

- Figure 4 indicates partial reset of a previously depressed immune state in COVR subjects following influenza vaccination. More discussion on why would influenza vaccine reset post covid immune states back to healthy controls? And why sex effects are impacting this reset?

We have simplified the presentation of Fig. 4, enabled by additional single cell CITE-seq data (covering more subjects and new day 28 data). The differences between COVR-M and COVR-F continue to be present based on the new single cell analysis, and the results are more quantitatively robust with the larger sample sizes. Both COVR-M and COVR-F experience an overall increase in average gene expression of depressed gene signatures in classical and non-classical monocytes identified in Fig. 1f,g (Fig. 4b), but the reversal is more pronounced in COVR-F especially in non-classical monocytes. COVR-F also have a larger fraction of genes that demonstrate reversal (Fig. 4c-e).

We have added more discussions per the reviewer's suggestion (see lines 426-436). Briefly, we speculated that mechanisms facilitating trained innate immunity (in which innate immune cells demonstrate altered – could be either enhanced or repressed – responses to stimulations after a previous inflammatory exposure) may

come into play in 1) establishing the depressed gene signatures in monocytes after mild COVID-19, and 2) partially reversing this depressed state following influenza vaccination[14,15]. These mechanisms include epigenetic and metabolic modifications and may occur centrally in progenitor cells residing in the bone marrow. In general, females are known to mount stronger innate/inflammatory responses to vaccination than males and therefore that could have contributed to the stronger reversal detected in COVID-19-recovered females.

- Overall observation of a depressed immune state following COVID. Are these effects speculated to result from SARS-CoV-2-specific effects or due to overall effects of a respiratory infection? As for the HC and COVR, did any of the subjects have any symptomatic respiratory infection, other than COVID19, that could be accounted for in some of the analyses?

We agree that some of the imprints we observed can be similar among different types of viral infections, yet some could be unique to SARS-CoV-2 infections, as shown in our preliminary comparison with natural influenza infection (Extended Data Fig. 2). We imagine that features such as the route of the infection and tropism are potentially important determinants. However, a detailed comparison with influenza or other viral infections is outside the scope of the current work. A future study could examine this issue following parallel cohorts with distinct prior infections. See also lines 445-450.

Regarding the occurrence of symptomatic respiratory infections between COVID-19 and influenza vaccination, please see common response #3 above.

Referee #3:

This is an ambitious attempt to address a series of important questions: first, whether mild to moderate COVID-19 leaves a long-term imprint in the immune system that can be detected by -omics approaches of the (potentially) new steady state or after a normalized, antigenically different challenge of the immune system, namely influenza vaccination. Second, whether any of the observed imprints are sex-specific. Overall, this paper is well written and holds together (at least in most places) the high complexity of analysis. Also, the graphic displays are helpful to guide the reader through the complex data, and overall, the discussion is fair.

We thank the reviewer for their positive comments on our work and enthusiasm for the importance of the questions asked.

My main concerns with this manuscript are that the observed phenotypes are very subtle, the n numbers are not very high, the phenomena described are slightly unlinked, and due to the nature of the study, correlative data is combined to develop a narrative that remains however speculative. This in combination with the fact that only blood data is available for analysis, while both in SARS-CoV-2 infection and in vaccination, local processes are crucial, it remains difficult to judge the biological meaning of the phenomena described here.

Thank you for these insightful comments. One of the reasons why some of the key differences and potential connections we highlighted in the previous manuscript might appear "unlinked" and less concrete is that the bulk blood transcriptomic data are confounded by contributions from changes in cell frequency and cell state. We therefore could not quite pinpoint the cellular sources of our signals in our original manuscript.

In our revised manuscript we anchor our narrative using the new multimodal single cell CITE-seq data (measuring 138 surface proteins, transcriptome, and V(D)J simultaneously in single cells), which provide high-resolution single cell profiles of 40 subjects over days 0 (baseline/pre-vaccination), 1, and 28 (post-influenza vaccination). The single cell data together with other improvements (see also common response above) allowed us to streamline the presentation and pinpoint cell type-specific transcriptional signatures to concretely connect our major findings, including:

1. Identification of temporally stable (i.e., not correlated with time since diagnosis cross-sectionally) cell type-specific imprints in Fig. 1, including a novel depression signature in both classical and non-classical monocytes. In Fig. 4, this depressed monocyte signature is followed using single cell data from days 1 to 28 post vaccination to arrive at the conclusion that influenza vaccination partially reverses this transcriptional state in monocytes towards that of healthy controls.

2. An integrated presentation of the sex dimorphic differences in both innate (IFN and antigen-presentation related signatures) and adaptive (plasmablast and antibody titers) responses to influenza vaccination in Fig. 2. These observed differences are robust, as demonstrated separately within the bulk transcriptomic, single cell CITE-seq, and circulating protein data. The COVR-M specific elevation of the early, day 1 response involving IFNγ is correlated with the increased humoral responses (Extended Data Fig. 3d). In Fig. 3, we utilize single cell data to identify baseline, pre-vaccination differences in the frequency of virtual memory-like CD8+ effector memory T-cells as a source of the elevated IFNγ response. This thus establishes a potential cellular link between prior mild viral infection and functional responses to an antigenically distinct immunological challenge in humans. The surface phenotype of these memory cells is consistent with the hypothesis that they emerged from bystander activation during acute COVID-19. Bystander activation during COVID-19 has been described in the literature[16],

but to the best of our knowledge the persistence and functional relevance of these cells in recovered individuals are not known.

Our paper thus represents a unique contribution revealing the long-term (post-clinical recovery) inflammatory/cellular imprints of (even mild) viral infections and their impacts on immune response outcomes following heterologous antigen exposure.

We agree with the reviewer and appreciate the importance of tissue level responses. However, sample collection from the injection site and local lymph nodes is logistically and clinically challenging even during normal times and was impossible at the NIH Clinical Center during the height of the pandemic in 2020. In addition, we had to establish and obtain IRB approval for our protocol within a short time frame to start our study at the beginning of the 2020-21 influenza vaccination season (September 2020). We have added discussion about these study limitations to the paper (lines 505-507).

Major concerns:

1. Inter-individual variation in response to influenza vaccines is strongly determined by the individual pre-vaccination exposure history. A vaccine that all study participants were naïve to would have been a better choice and would have facilitated some of the analysis.

Please see common response #1 and #2 above.

2. It is a pity that true longitudinal data, i.e. different time points since diagnosis from the same donor, is not part of this study – again this would have strengthened the analysis.

We agree that following individuals longitudinally from the acute phase to recovery (so that there are multiple sampling time points since COVID-19 diagnosis from the same donor) would have been informative. However, as mentioned above, this study was logistically and technically challenging given the inability to recruit individuals with mild infections early in the pandemic (also because in-person clinical research at the NIH Clinical Center was severely limited). The timing of our study is unique in that it allowed us to catch the first wave of COVID-19 recoverees (infected early 2020) who were eligible for seasonal influenza vaccination during the Fall of 2020. Repeating this study later (e.g., in 2021) but with a design to include sampling starting from acute disease would largely catch individuals who have had COVID-19 vaccination and/or additional infections. See paragraph 3 of common response #1 for additional discussions related to this point.

Please see also common response #3 above.

We have added the above points to the study limitation section of the Discussion (lines 486-495).

3. An important point that is at least missing in the main text is the question of participant infections between COVID and vaccination – all assumptions here are based on the absence of such confounding factors – are we sure that participants did not have anything, not even a cold in that period?

Please see common response #3 above.

4. Following on from point 3: COVID infection of the participants was not the first, but the nth infections of their life. I understand from the comparison made with flu infection data that similar imprints were found

after other infections (and this is what I would expect). So the concept of a sort of persisting immune-paralysis that is partially or fully relieved is difficult to reconcile with a long sequence of infections – which one depresses, which one relieves the immune depression? Or would it be specific to COVID as compared to other infections to induce immunodepression, and of vaccination as compared to infection to be able to reinstate to original state? Apart from flu data, the authors cite literature on vaccinations to have immune-depressive effects, so the answer to both questions above seem to be no. I think it is important not only to discuss this, but to re-consider the basic assumptions of the study in the light of sequential immune-exposure that started long before COVID for these individuals.

We appreciate these thoughtful comments from the reviewer. We agree that any infection or other immune challenge may change the immune status – a future study to follow multiple immunological encounters within individuals could help clarify this important question. We expect that the immune statuses across individuals prior to their first SARS-CoV-2 encounter are already variable and shaped by previous exposures. Those pre-COVID-19 states may impact SARS-CoV-2 infection responses and resolution, which in turn can lead to the variation in baseline immune statuses (prior to influenza vaccination) that we observe in our study. For example, the frequency of GPR56+ effector memory CD8+ T-cells at baseline is highly variable within the COVR-M group (Fig. 3e).

Considering these complexities, however, our study's unique design contributes valuable data and conceptual advances to help reveal underlying principles regarding what happens after two well-defined immunological encounters in humans: mild COVID-19 as a natural experimental perturbation and influenza vaccination as a controlled and timed intervention. The latter could also be considered a "natural perturbation" insofar as a sizable fraction of the population we sampled would have had this intervention independent of our study. Thus, our data provide a valuable basis for the development of hypotheses regarding more complex scenarios including what happens over longer timescales with additional encounters.

Furthermore, we enrolled COVID-19-recovered subjects together with matching HCs who never had COVID-19 at the time of influenza vaccination. It is reasonable to believe that the extent of heterogeneity in the exposure history within each group is, on average, quite comparable between the COVR and HC groups. Thus, while a particular individual may have unique imprints after COVID-19 as a function of their prior exposure history before the infection, the group-level features shared across individuals (e.g., depressed monocyte signatures or the heightened IFN response after influenza vaccination in COVR-M) that we were able to detect should be largely independent of differences in personal exposure history. This approach thus allowed us to robustly uncover both pre-vaccination (baseline) and immune response features associated with prior mild COVID-19.

While some specific imprints due to prior infection are partially reversible as in the case of the depressed monocyte signature, they are still not fully reversed by vaccination or natural resolution. Thus, while monocyte statuses may be reshaped and temporally stable over the timescale of months after vaccination, they may be constantly shaped over the timescale of years due to other encounters. This hypothesis is consistent with the extensive heterogeneity seen in monocyte phenotypes in the human population. Another consideration is the quantitative strength of the perturbation. For example, a symptomatic infection is generally expected to be a stronger inflammatory perturbation than vaccination. A stronger perturbation may result in longer lasting effects, even if they are partially reversible or further reshaped by the next perturbation. Some effects, such as the bystander expansion and phenotypic alteration (e.g., lowering of surface CD5 levels[17]) of virtual memory-like CD8+ T-cells, may potentially be less reversible.

Finally, we agree that some of the imprints we observed are shared among different types of viral infections, yet some could be unique to SARS-CoV-2 infection, as shown in our preliminary comparison with natural influenza infection (Extended Data Fig. 2). We imagine that features such as the route of the infection and viral tropism are potentially important determinants. However, a detailed comparison with influenza or other viral infections is outside the scope of the current work. A future study could examine this issue following parallel cohorts with distinct prior infections.

We have added discussion to highlight the above points (lines 450-457 and 476-485).

5. Line 60 why include 2 asymptomatic subjects? Do you know when they were infected, probably not? How do you establish TSD? How do the analyses look without them?

We do not know when the asymptomatic individuals were infected. For this reason, we could not determine the time since diagnosis, and these individuals are thus not included in the analyses that relied on TSD to determine which parameters were still resolving. To clarify this point, we have added information to the Methods, the figure legends, and the manuscript (line 109).

6. Along the same lines, it is not clear to me why in different analyses, different n numbers appear, when figure 1c suggests that the key analyses were done on all 71 study participants (e.g. 2b, c, also later figures).

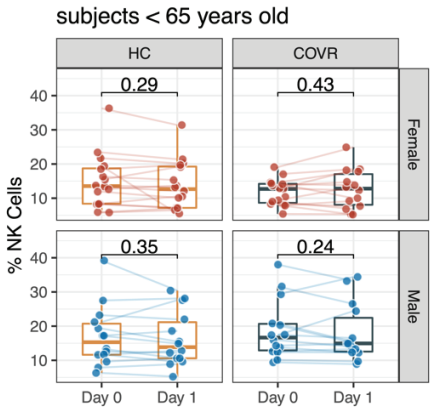
For the analyses of the immune response following seasonal influenza vaccination, we present results from individuals younger than 65 who received the Flucelvax vaccine. We did this as older individuals (age $\geq$ 65) received the high-dose Fluzone vaccine and we had too few older individuals to evaluate them independently as a group. In addition, not all subjects who had titer data collected (total n = 72) were included in flow cytometry analysis (total n = 54; the sample size is smaller primarily because we only included subjects who were younger than 65). These two factors resulted in different number of subjects listed in previous figure panels. We now show the number of subjects included in each assay in Fig. 1b.

7. It is a pity that no PBMCs were frozen to allow testing of some of the findings more directly – for instance, the authors link higher serum IFNg on d1 post vaccination to the increased presence of CD8 effector cells in M convalescents per-vaccination – it could be tested whether these cells are prone to IFNg production if such samples were available. It could also be tested whether increased NK cells on d1 post vaccination in F would not produce higher IFNg. Such minimal functional experiments would strengthen some of the links that the authors plausibly draw but which remain speculative.

Please see Fig. 3 in which we use our new single cell data to address the reviewer's questions. We feel that these observations, from measurements of *in vivo* statuses and responses, provide biological insights, including the finding that a population of GPR56+ virtual memory-like effector memory CD8+ cells contributed to the heightened IFN response early after vaccination.

We thank the reviewer for the *in vitro* experiment suggestion. Although we have limited amounts of extra cells banked, we felt that the new single cell analyses have addressed the reviewers' questions and thus decided to save the limited cells we have for future work or to address additional questions/suggestions from the reviewers. If this reviewer feels that specific *in vitro* assays would strengthen our findings, we will certainly consider pursuing those suggestions.

The increase in the NK cell-related signature in COVR-F on day 1 was based on bulk transcriptomic analysis (previous Fig. 3b). However, NK cell frequency was not changed significantly between days 1 and 0 in COVR or HC (in either sex) based on clinical flow cytometry data – see figure below.



Percent of NK cells measured in the T- and B-Lymphocyte and Natural Killer cell phenotyping panel for subjects under 65 years of age at days 0 and 1 relative to the time of influenza vaccination. Same subject at those two timepoints are connected by a line. P-values shown are determined by paired two-tailed Wilcoxon test.

We used single cell CITE-seq data to further assess the "enriched in NK cells (I) (M7.2)" signature (which does not appear to be IFNγ-related) in NK cell subsets (for the day 1 vs. day 0 response comparison). Details can be found below, but in summary we did not find evidence of a higher day 1 IFNγ response in NK cells in COVR-F. In fact, some NK cells might have contributed to the higher day 1 IFNγ response in COVR-M compared to HC-M and COVR-F.

Using CITE-seq data, we detected an increase (day 1 vs. day 0) in the "enriched in NK cells (I)" signature in $CD16^{hi}$ NK cells in COVR-F vs. HC-F (FDR = 0.06; Extended Data Table 6 – "COVR-F vs. HC-F vax response" tab). A similar trend was seen in COVR-M vs. HC-M (FDR = 0.16; Extended Data Table 6 – "COVR-M vs. HC-M vax response" tab). In the "sex difference" comparison (difference between day 1 responses of COVR-M and HC-M vs. day 1 responses of COVR-F and HC-F; Extended Data Table 6 – "Sex difference vax response" tab), only the proliferating NK cell subset (marked by KI67 mRNA expression) shows marginal enrichment for this signature in the negative direction (same direction as the bulk data – i.e., lower response in males; FDR = 0.1, with leading edge genes such as IL2RB, KLRB1, ARL4C but no apparent enrichment for IFNγ-related genes). Thus, we conclude that the sex-dependent day 1 NK cell signature from bulk expression analysis in previous Fig. 3b did not reflect cell frequency changes but might capture complex cell state changes in certain NK cell subsets. We also examined IFNG transcript levels in the two major NK cell subsets ($CD16^{hi}$ and $CD16^{lo}$ NK cells). No difference was detected in the former. In the latter subset (comprising ~1% of cells), COVR-M had significantly higher increases (day 1 vs. day 0) than HC-M ($p < 0.05$; Extended Data Table 5 – "COVR-M vs. HC-M vax response" tab), while the opposite was true for females: COVR-F had smaller increases than HC-F ($p < 0.05$; Extended Data Table 5 – "COVR-F vs. HC-F vax response" tab).

8. In figure 1 and related data, stable and evolving changes in immune set points are analysed. Subtle differences in monocytes and T cells are reported, partly reminiscent of the reported more dramatic changes during acute and severe COVID-19. Many of the cellular, protein and Cytek analyses do not reach an FDR< 0.05 significance value.

In previous Extended Data Fig. 1d, we tried to provide a more comprehensive view of the data by showing all cell frequency/count and circulating protein parameters with p value < 0.01 (and highlighted those with FDR < 0.05 with a star). We have removed that panel in the revised manuscript to avoid distraction for the readers and instead refer them to the Extended Data Tables for the complete results. Furthermore, our new Fig. 1 places emphasis on the single cell data (complete results in Extended Data Tables 5 [gene level] and 6 [GSEA/enrichment analysis]), which provides greater resolution and cell type specificity in transcriptional signatures regarding baseline states than the bulk blood transcriptomic data shown in previous Fig. 1 and Extended Data Fig. 1. In the revised manuscript the bulk transcriptomic data was used only as an initial guide to suggest what biological processes should be further examined using the single cell data.

In the related data in Extended Data Fig. 1a, of five pathways shown, the only one with a significant (FDR <0.05 – not a very stringent cut-off, we hardly ever use anything less stringent than < 0.01) enrichment factor in F does not correlate significantly with TSD, and vice versa for the four other ones. I appreciate the authors' argument that the (significant) loss over time in F but not in M leads to (significant) sex differences overall in a subset of these categories, but overall, my impression is that the long-term imprints detected here are quite subtle. The fact that we don't really know what a change in e.g. DC representation in the blood means, given the tissue- or lymph node restricted role of these cells, makes it difficult to develop a clear narrative of the changes we observe and their biological meaning.

In our previous analyses, there were two ways that we used the TSD: 1) parameters that were stable (i.e., not correlated with the TSD across subjects) and different between COVR and HC in each sex, and 2) parameters that were correlated with TSD but potentially resolving at different rates between COVR-M and COVR-F. The latter are often more subtle and would require additional subjects in order to achieve greater statistical robustness. Additionally, this aspect of COVID-19 immune recovery is not a major focus of our paper, which is why we included it in the Extended Data Figures previously. Our updated, simplified manuscript further de-emphasizes this aspect and focuses on features that are stable.

We agree that the relationship between circulating immune cell frequencies and those in lymph nodes and tissues is an area of additional research (see lines 505-507).

Line 129 I agree that M/F combined analysis yields less significant differences, but could this be due to increased n number for analysis? Concretely, the stats shown if EDFig1F refer to sex differences, while the figure is used for an argument about sex-independent differences (line 130), while the Cytek analysis EDFig1d does not show sex-independent differences for pDCs.

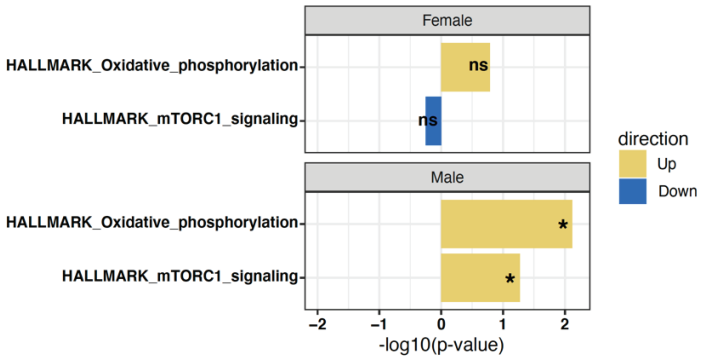
The primary reason why the combined analysis had fewer significant results is that the females and males in the COVR group exhibit different characteristics compared to their HC counterparts for some of the parameters we tested. For example, since COVR-M had elevated frequency of monocytes compared to HC-M but COVR-F showed the opposite trend, these effects canceled out in the combined analysis. Such diverging trends between COVR-M and COVR-F increase the variance within the COVR group and this outweighs the increase in sample size (which would only help if the trends were consistent between males and females), thus yielding fewer significant differences in the combined analyses.

We apologize for the confusion about the pDCs. Previous Extended Data Fig. 1f shows sex-independent difference in pDC frequencies between COVR and HC, but it did not pass the 0.05 FDR cutoff in the sex-specific analyses (as indicated in previous Extended Data Fig. 1d). We mentioned this pDC observation in the text previously because differences in pDC frequency have also been seen in acute COVID-19. We have removed

this point and the associated figures in our new manuscript to avoid confusion and because this is peripheral to the key messages of our revised manuscript.

9. In figure 2, the Ab and B cell response to influenza vaccine is analysed, and recovered M appear to have a slightly stronger Ab response that correlates with higher frequencies of plasmablasts and higher baseline levels of a specific memory B cell subset, which in turn correlates with an mTORC-driven gene signature. If I understand correctly, the metabolic signature is derived from bulk analysis of PBMCs, so this may or may not reflect metabolic changes in this B cell subset. I do not see that the B-cell specific metabolic argument is picked up upon once the scRNAseq in fig.4 is done, and EDFig.5B does not seem to show detectable differences for B cells in males.

We thank the reviewer for these helpful comments. We agree that the bulk data was not as helpful in pinpointing the cellular origin of the metabolic signals we detected. Our expanded CITE-seq single cell data allowed us to assess these metabolic signatures in the CD71^{hi} memory B-cells. The figure below shows increased enrichment of two metabolic gene sets in COVR-M compared to HC-M within these cells.



*Bar plot showing the -log10 p-value multiplied by the direction of the difference between COVR and HC (up=higher in COVR; down=lower in COVR; shown separately for males and females). GSEA analysis of the metabolic gene sets shown was conducted using the pseudobulk gene expression data derived from the CD71^{hi} memory (CD20+CD71^{hi}IgD-) B-cells in our CITE-seq data. P values determined by GSEA. * Benjamini-Hochberg (BH) adjusted p < 0.05.*

In our updated manuscript, we have removed the CD71^{hi} memory B-cell results as it became apparent that they were peripheral to the key messages of our revised manuscript.

10. In fig. 3, further analysis of the anti-vaccine response is performed, and strong early IFN responses were found in recovered M. The authors link IFNg reponses to higher base-line CD8 effector cells levels in M, and I wonder why increased NK-frequencies in F cannot provide a similar level IFNg. As mentioned above, testing function of cells from frozen aliquots would come in handy here.

Please see our response to point 7 above.

11. Finally fig.4 explores a vaccination-driven immunological “reset”. Apart from the conceptual issues (explained in point 4) to consider, his figure is not easy to understand. In particular, for Fig. 4B the legend talks about four lines, including the HC base line, but this fourth line is missing – is the HC base line the comparator and therefore would be the zero value on the x axis? Based on this initial issue, the whole argument of vaccination releasing immune paralysis post COVID is difficult to follow. The x-axis annotation (abs(effect size))

is not helpful to understand this, and the precise selection of genes that are channelled into this analysis is difficult to follow. It is also not helpful for the comparison that the scale of the x axis for M and F is not the same. The authors conclude from the data analysis in this and associated figures that indeed, at d1 post vaccination of convalescents, there is a release of depressed genes, in particular in monocytes, which persists up to day 28 more in F than in M. Early monocyte activation post vaccination is perhaps unsurprising, so to talk about a relief of an immune-depressed state may be pushing it a little – but it does show that we ae not in front of an immune-paralysis that cannot be overcome.

We thank the reviewer for this helpful feedback on previous Fig. 4. Thanks to the new CITE-seq data covering additional subjects and the day 28 time point, we were able to simplify the analysis and presentation of the data. Previous Fig. 4b, 4g, 4h relied on bulk gene expression data to show the partial reversal of the depressed monocyte signature. However, our new single cell data allowed us to look directly within specific monocyte populations across days 0, 1, and 28 to directly assess the reversal phenomenon that we showed in the original manuscript, thus giving us greater confidence in our findings. The depressed monocyte signature is also now presented early in Fig. 1 as one of the key imprints we highlight using single cell data. In Fig. 4 we ask specifically what happens to the depressed signature after vaccination.

Additionally, in Extended Data Fig. 5a,b (CITE-seq single cell parallel of previous Fig. 4b that used bulk data), the days are now presented from top to bottom (day 0 to day 28) for ease of interpretation; we only show genes from the depressed signature in the classical and non-classical monocytes at baseline identified in Fig. 1f and g, and we show the gene expression difference from the baseline (day 0) of the HC rather than the absolute effect size. The scale on the x-axis is now the same for females and males.

Minor concerns:

1. Typo EDFig.2c: COVR-F instead of COR-F

We have corrected this.

2. File names (at least of the files extracted from the zip) do not indicate which suppl. Tables are in the excel sheet, often it is not indicated in the excel sheet itself – this makes it difficult to follow the data referenced to tables. Having all tables, well annotated, in on excel sheet may be preferable.

Thank you for this helpful suggestion. We have now clearly labeled all files with a header to improve readability. We agree with the excellent suggestion of putting all tables in a single Excel sheet. Many of our tables are already combined in this manner – for example, Extended Data Table 6 has eight tabs – with each tab corresponding to a specific comparison. Unfortunately, having all the tables in one single file would make the file prohibitively large and not as well organized because there would be an extremely large number of tabs (whose labels have limited length).

References

1. Tsang, J. S. *et al.* Global Analyses of Human Immune Variation Reveal Baseline Predictors of Postvaccination Responses. *Cell* **157**, 499–513 (2014).
2. HIPC-CHI Signatures Project Team & HIPC-I Consortium. Multicohort analysis reveals baseline transcriptional predictors of influenza vaccination responses. *Sci. Immunol.* **2**, eaal4656 (2017).
3. Goodridge, H. S. *et al.* Harnessing the beneficial heterologous effects of vaccination. *Nat. Rev. Immunol.* **16**, 392–400 (2016).
4. Qi, Y., Shaman, J. & Pei, S. Quantifying the Impact of COVID-19 Nonpharmaceutical Interventions on Influenza Transmission in the United States. *J. Infect. Dis.* **224**, 1500–1508 (2021).
5. Langley, J. M. *et al.* Dose-Sparing H5N1 A/Indonesia/05/2005 Pre-pandemic Influenza Vaccine in Adults and Elderly Adults: A Phase III, Placebo-Controlled, Randomized Study. *J. Infect. Dis.* **203**, 1729–1738 (2011).
6. Hagan, T. *et al.* Antibiotics-Driven Gut Microbiome Perturbation Alters Immunity to Vaccines in Humans. *Cell* **178**, 1313-1328.e13 (2019).
7. Faix, D. J. *et al.* Decreased Serologic Response in Vaccinated Military Recruits during 2011 Correspond to Genetic Drift in Concurrent Circulating Pandemic A/H1N1 Viruses. *PLOS ONE* **7**, e34581 (2012).
8. Hsu, J. P. *et al.* Rate of decline of antibody titers to pandemic influenza A (H1N1-2009) by hemagglutination inhibition and virus microneutralization assays in a cohort of seroconverting adults in Singapore. *BMC Infect. Dis.* **14**, 414 (2014).
9. Verschoor, C. P. *et al.* Microneutralization Assay Titres Correlate with Protection against Seasonal Influenza H1N1 and H3N2 in Children. *PLOS ONE* **10**, e0131531 (2015).
10. Kotliarov, Y. *et al.* Broad immune activation underlies shared set point signatures for vaccine responsiveness in healthy individuals and disease activity in patients with lupus. *Nat. Med.* **26**, 618–629 (2020).
11. Avey, S. *et al.* Seasonal Variability and Shared Molecular Signatures of Inactivated Influenza Vaccination in Young and Older Adults. *J. Immunol.* **204**, 1661–1673 (2020).
12. Arneborn, P., Biberfeld, G., Forsgren, M. & von Stedingk, L. V. Specific and non-specific B cell activation in measles and varicella. *Clin. Exp. Immunol.* **51**, 165–172 (1983).
13. Jin, Y. *et al.* Endothelial activation and dysfunction in COVID-19: from basic mechanisms to potential therapeutic approaches. *Signal Transduct. Target. Ther.* **5**, 293 (2020).
14. Netea, M. G. *et al.* Defining trained immunity and its role in health and disease. *Nat. Rev. Immunol.* **20**, 375–388 (2020).
15. Cunningham, K. T., Finlay, C. M. & Mills, K. H. G. Helminth Imprinting of Hematopoietic Stem Cells Sustains Anti-Inflammatory Trained Innate Immunity That Attenuates Autoimmune Disease. *J. Immunol.* **206**, 1618–1630 (2021).
16. Bergamaschi, L. *et al.* Longitudinal analysis reveals that delayed bystander CD8+ T cell activation and early immune pathology distinguish severe COVID-19 from mild disease. *Immunity* **54**, 1257-1275.e8 (2021).
17. White, J. T., Cross, E. W. & Kedl, R. M. Antigen-inexperienced memory CD8+ T cells: where they come from and why we need them. *Nat. Rev. Immunol.* **17**, 391–400 (2017).

Reviewer Reports on the First Revision:

Referees' comments:

Referee #2 (Remarks to the Author):

First I want to thank the authors for their efforts in pushing for greater understanding of postinfectious immune system function and adaptation, a topic of great interest and importance.

After reading the rebuttal and the completely updated manuscript my impression is that many aspects of the study has been improved. The new detailed single-cell multiomics data adds greater resolution and is in agreement with several of the findings from bulk mRNA analyses.

My main concern however remains, namely that effects reported are very subtle and number of subjects studied so small that I remain uncertain whether any of these effects would hold up in a validation cohort. I don't think it is possible to do more with the subjects, samples and data already generated and all analyses are already well performed and reasonably interpreted to support the conclusions drawn.

My suggestion is that the authors try to recruit an additional cohort or samples from another cohort to try and validate some of the suggested mechanisms of immune adaptation. Ideally the authors could include subjects infected by SARS-CoV-2 and also another virus for comparison, such as flu. I also suggest that such a validation cohort be vaccinated with another vaccine to which the subjects would be naive, like Yellow fever, in order to answer the legitimate question of prior immunity to flu and it's possible contribution as a basis for the differences seen in the original cohort.

Referee #3 (Remarks to the Author):

How past exposure effects future immune response is an important question whose importance for developing new treatment paradigms is paramount. However our understanding of this is overwhelmingly comes in the context of antigen specific memory responses. Laboratory studies observed a priming effects that relate one stimuli to another, but connecting these non-antigen specific effects to real world scenarios has been very limited. Here the authors conduct a tour-de-force systems immunology study to look at the relation of past (mild) COVD-19 infections to downstream responses to the seasonal flu vaccine. The logistics of this study are impressive, especially given the situation in which it was conducted – during the pandemic and pre-COVID vaccination. Sample size is respectable for systems vaccinology studies, especially with the addition of the CITE-seq data, and for the first time in the decade+ of the field, there is finally proper attention paid to gender differences in response (to date, most studies have looked at gender either at baseline or as a co-variate one adjusts to).

The authors make 3 major claims in the paper: (1) That a mild COVID infection shifts individual's immune state / set point, thereby affecting their responses to future treatments (i.e., influenza), (2) that this is gender dependent, (3) that the flu vaccine reverses the effects of COVID. Briefly, I believe the first two claims, which are the most important ones, are very well shown in the work and the third, which is more minor, possibly requiring some more investigation or toning down. My specific comments appear below:

1) Blood transcriptional modules (BTMs) – where never developed for single cell data but rather on whole blood where cell composition makes much of the associations between genes. Though there is no immediate alternative I am aware of (asides from clustering and enrichment), if using them for CITE_seq I would be wary of some noise coming from the mismatch.

2) The paper keeps switching between BTMs and Hallmark gene sets, unclear to me why. If consistency is a possibility, better.

3) Unclear to me where the monocyte signature emerged from (i.e. no direct relation to what is seen in acute). I would expect it to have some relation to the acute COVID infection – as in there is some continuum in which it appears. Possibly regress by disease severity in acute studies data ? You may see a shift a gradual shift in cell state.

4) Why run baseline correlations as a single correlation analysis, rather than take baseline features (possibly reduced to only those differing between groups) and perform some form of multi-variate regression / lasso analysis to find the predictive features, there may be more than GPR56+ CD8 tcells.

5) Did you see any difference in GPR56+ CD8 tcells between males and females at baseline ?

6) Was there a relation between reversal and the flu titers?

7) Why are we assuming that reversal is a good thing ? Obviously from a flu titer perspective it is not. As the authors themselves state in the discussion, it may be a trained immunity type phenotype that ideally one would want to keep.

8) The question poised in the last section is “Does the influenza vaccine shift immune states toward normal?” (the authors have this even in the figure cartoon). Yet ultimately the answer is limited to looking at monocytes, where I suspect a more global perspective would have yielded a much broader of this important question view (e.g. dimensionality reduction of cell composition/state?). Moreover I would consider switching the visuals of this so that functionality is highlighted (e.g. place genes on KEGG nodes). The present heatmaps are not very informative in building a picture in ones mind.

Referee #4 (Remarks to the Author):

The manuscript is now much improved, the CITE-seq data help greatly to carve out important differences, and the message has been streamlined and very much simplified. The authors show a new set-point post mild COVID-19 that includes reduced monocyte signatures and increased CD8 activation signatures prior to secondary stimulus. Post vaccination, an increased IFNg response is found that the authors locate to bystander-activated CD8 cells and that may explain the increased monocyte activation and later plasmacyte / Ab response. Most of these latter increases are found predominantly or only in males. In the last figure, the authors show data in line with the notion that the depressed monocyte state is reversed by vaccination. Significant concerns remain that need to be addressed before this study is appropriate for Nature:

1. In figure 1 and related data, bulk FACS and RNA analysis showed an increased monocyte signal in COVR-M that however disappeared in the CITE-seq analysis, suggesting that that the signature was mainly driven by cell frequencies. As increased cell frequencies are a biological finding with potential biological consequences, the authors may be slightly too dismissive about this.
2. The term monocyte signature in 1f and g needs to be filled with more life, as this term per se does

not mean that they are more activated. More generally, when a monocyte cluster has a stronger or weaker monocyte signature it is actually not so obvious what this means, it makes more sense in mixed populations. Please discuss the genes in there in a bit more detail, as I seem to spot also some negative regulators in here.
3. I see in extended fig. 2b a highly significant negative correlation in the F, with a slope that is roughly half that of the (positive) slope in M. The authors therefore cannot claim that in F, influenza infection had no impact, it had the opposite impact of that in M. Please correct these statements and discuss in the context of your data.
4. Line 175-77 where do you show this data? You only state in the text. Please indicate these genes in the plot S2b or do a GSEA as in S2c.
5. Line 179-180, "although in general some of these imprints are likely pathogen-dependent": a politician could not be more vague, what do you mean?
6. The increased IFNg response is intriguing. in S4a, and b, how do these correlations look for HC? Please add this data and discuss. Also, you mention later (line 315) very much in passing your data supporting that also NK cells contribute to increased IFNg, please add this data, maybe here or later, but without this the study is incomplete.
7. Fig S4d, would this be time-correlated with higher n numbers? What does the different direction of travel of the trends between m and f mean?
8. 3f How would this look for gpr56 neg cells? Please add.
9. Source of IFNg: The extended omics analysis leads the authors to conclude that bystander-activated GPR56+ CD8 cells whose phenotype is somewhere between EM and TEMRA CD8 cells produce the early IFNg in an Ag-independent way, rather driven by cytokines. I think it is crucial and feasible to show the in vitro, stimulating PBMC with the cytokine(s) in question, ideally more than one cocktail, and stain by icFACS for all CD8 subsets in question, and in addition NK cell (subsets). This will also the authors to show in a real stimulation experiment that their hypothesis is true, and it will add data on the NK contribution to early IFNg.
10. For me, the most difficult part is figure 4: As a general comment, I think the figure must contain data showing how this signature changes in HC over time post vaccination, as the notion of a reversal of immunodepression makes sense if this signature is not changed in HC, or at least if the d28 time point shows convergence between COVR and HC samples.
11. In this context, while the discussion comparing your findings to papers talking about "training" is extensive and can be shortened, I think it is important to mention that the d28 time point is beyond the expected half life of blood monocytes, so resetting the set-point must happen in precursors or must be maintained by non-monocyte-intrinsic mechanisms.

Author Rebuttals to First Revision:

Response to reviewers

We thank the reviewers for their constructive and helpful comments. We are glad that the reviewers find our conclusions robust and sound (e.g., reviewer #2 – *"all analyses are already well performed and reasonably interpreted to support the conclusions drawn"*). We also appreciate that our work is recognized as important and exciting for the field, (e.g., reviewer #3 – "*How past exposure effects future immune response is an important question whose importance for developing new treatment paradigms is paramount……but connecting these non-antigen specific effects to real world scenarios has been very limited. Here the authors conduct a tour-de-force systems immunology study to look at the relation of past (mild) COVD-19 infections to downstream responses to the seasonal flu vaccine*"). Also mentioned was the recognition that our work reveals sex differences that have thus far not been sufficiently examined by the field.

Please find our point-by-point response to the reviewers' comments below. In this revision, per the suggestion of reviewer #4, we performed *in vitro* cytokine stimulation experiments using baseline PBMC samples from subjects with CITE-seq data (n=40), which provide further support for our earlier observations from CITE-seq and another layer of independent support for our conclusion that COVID-19-recovered males have distinct antigen-agnostic immune set points. Based on reviewers #3 and #4's input, we have also clarified and improved the presentation of the data in Fig. 4 regarding the impact of influenza vaccination on the post COVID-19 gene expression imprints in monocytes. We are optimistic that these additions, together with other updates, both address the reviewers' concerns and have improved the manuscript.

Point-by-point responses

Reviewer #1: no new comments were received for the most recent round of reviews. Our responses to the previous round of review from this reviewer can be found in the attached appendix below.

Reviewer #2:

First I want to thank the authors for their efforts in pushing for greater understanding of postinfectious immune system function and adaptation, a topic of great interest and importance.

After reading the rebuttal and the completely updated manuscript my impression is that many aspects of the study has been improved. The new detailed single-cell multiomics data adds greater resolution and is in agreement with several of the findings from bulk mRNA analyses.

My main concern however remains, namely that effects reported are very subtle and number of subjects studied so small that I remain uncertain whether any of these effects would hold up in a validation cohort. I don't think it is possible to do more with the subjects, samples and data already generated and all analyses are already well performed and reasonably interpreted to support the conclusions drawn.

My suggestion is that the authors try to recruit an additional cohort or samples from another cohort to try and validate some of the suggested mechanisms of immune adaptation. Ideally the authors could include subjects infected by SARS-CoV-2 and also another virus for comparison, such as flu. I also suggest that such a validation cohort be vaccinated with another vaccine to which the subjects would be naive, like Yellow fever, in order to answer the legitimate question of prior immunity to flu and it's possible contribution as a basis for the differences seen in the original cohort.

We appreciate that this reviewer has continued to find our scientific questions important and that our findings and conclusions are robust and sound (*"all analyses are already well performed and reasonably interpreted to support the conclusions drawn"*).

We also appreciate the reviewer's feedback and suggestions. However, we respectfully disagree that the effects we reported are subtle. The novel set point observations and sex dimorphic effects we report have clearly robust and significant effect sizes (reviewers #3 and #4 also found this to be the case based on their comments). The sex dimorphic responses to influenza vaccination are supported by both innate and adaptive (encompassing both antigen-specific and non-specific) measurements, including antigen-specific plasmablasts on day 7 and independent support from the antibody responses on day 28 (Fig. 2). Furthermore, our new *in vitro* stimulation results in updated Fig. 3 (and Extended Data Fig. 4) provide additional compelling supporting evidence for our conclusions (see our response to point #9 from reviewer #4 below for additional details.)

We also respectfully disagree that our study is based on too few subjects. We have 33 COVID-19-recovered and 40 matching healthy control subjects. With deep phenotyping via extensive and cutting-edge assays, including molecular, transcriptomic, cellular, serological, multiomic single cell, and serum protein data (see Fig. 1b), ours is one of the larger (if not one of the largest) systems vaccinology studies conducted thus far. These also represent the most extensive analyses on post COVID-19 recoverees and the largest multiomics/CITE-seq single cell dataset on vaccination responses (covering n=40 subjects at baseline, innate, and long-term response time-points). These points are also supported independently by Reviewer #3's comments, e.g., *"Here the authors conduct a tour-de-force systems immunology study to look at the relation of past (mild) COVD-19 infections to downstream responses to the seasonal flu vaccine…Sample size is respectable for systems vaccinology studies, especially with the addition of the CITE-seq data."* In addition, our sample sizes are large in comparison to recent cutting-edge system vaccinology studies, e.g., a recent study published in *Nature* on mRNA vaccine responses in healthy subjects had 6 subjects with CITE-seq data out of 31 with bulk transcriptomics (Arunachalam et al. Nature 2021), compared to our n=40 with CITE-seq and 73 (33 COVID-19-recovered vs. 40 matching healthy controls) with bulk transcriptomics. Similarly, a recent single cell study of vaccination effects on innate immune cells in healthy subjects had 21 subjects (Wimmers et al. *Cell* 2021).

Finally, the reviewer suggests recruiting independent validation cohorts and vaccinating them with Yellow Fever and influenza vaccines. This would constitute a completely new study and is beyond the scope of our work. Additionally, Yellow Fever vaccine is not clinically indicated in the COVID-19-recovered population and would require a completely new clinical protocol to be written and approved, followed by detailed immunological and omics analyses. As we mentioned in the previous rebuttal, our population is uniquely suitable for addressing our main scientific question: *what happens to the immune system after mild COVID-19?* Because the recovered subjects in our study had only one mild SARS-CoV-2 infection (and no COVID-19 vaccination) before we administered the influenza vaccine, this allowed for a "clean" comparison with matching healthy controls. Recruiting mild COVID-19 recoverees who have not had subsequent infections or previous COVID-19 vaccination (which, as we've shown in this study, can shape immune statuses) is close to impossible. We also wish to highlight that the type of recall responses that influenza vaccination elicits are important to study in COVID-19-recovered subjects given that they represent the most common type of responses in such populations, especially since many routinely receive the seasonal influenza vaccine.

Reviewer #3:

1. Blood transcriptional modules (BTMs) – where never developed for single cell data but rather on whole blood where cell composition makes much of the associations between genes. Though there is no immediate alternative I am aware of (asides from clustering and enrichment), if using them for CITE_seq I would be wary of some noise coming from the mismatch.

   We agree that currently there are no ideal gene sets for single cell data. BTMs can capture both cell frequency-dependent and -independent (cell intrinsic) signals. In this study, we started our analyses using the whole blood (WB) transcriptomic data and utilized gene sets from multiple sources for increased coverage of molecular and cellular signals. Using the CITE-seq data, we then identified the cell subsets from which the WB enrichment signals might

originate and examined the leading edge genes to better delineate and understand the underlying gene-level signature independent of the gene set/BTM labels.

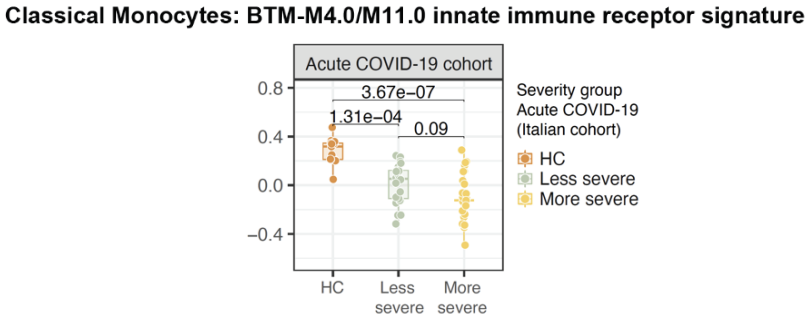
2. The paper keeps switching between BTMs and Hallmark gene sets, unclear to me why. If consistency is a possibility, better.

   Please see our response to point #1 above. We only utilized the Hallmark IFNG response gene signature to assess the bulk (Fig. 2b and Extended Data Fig. 4b) and cell-specific (Fig. 2d,e; Extended Data Fig. 3b; Supplementary Information Fig. 4) IFNG responses.

3. Unclear to me where the monocyte signature emerged from (i.e. no direct relation to what is seen in acute). I would expect it to have some relation to the acute COVID infection – as in there is some continuum in which it appears. Possibly regress by disease severity in acute studies data? You may see a shift a gradual shift in cell state.

   The monocyte signature emerged initially from the analysis of the whole blood transcriptomic data (Extended Data Fig. 1e). We noted that the monocyte-related genes (from BTM gene sets) show sex-specific expression differences. To avoid cell frequency confounding, particularly given that the proportion of monocytes is higher in COVR-M compared to COVR-F (Fig. 1d), we used the CITE-seq data to further evaluate gene expression differences within monocyte subsets. We found that the expression of the leading edge genes from the above enriched monocyte gene sets (identified from the whole blood transcriptomic analysis) were significantly lower in classical and non-classical monocytes in both COVR-M and COVR-F compared to the age- and sex-matched healthy controls (Fig. 1f,g). Collectively, these leading edge genes comprise the monocyte signatures (now called the "innate immune receptor signatures").

   To assess whether these signatures genes had lower expression during acute disease, we utilized an independent patient cohort from Italy with acute infection because we do not have acute samples from the same recovered patients in our study. We have already generated CITE-seq data from this Italian cohort earlier in a separate study[1]. As shown in Extended Data Fig 1i (reproduced below), we indeed found that lower expression of genes from the monocyte signatures is associated with both acute COVID-19 and disease severity in this Italian cohort. Since most of the patients in that cohort were males, we were unable to analyze the trend separately for males and females. However, we expect similar trends in females since this monocyte imprint was found in both female and male recoverees.



Extended Data Fig. 1i from the manuscript. Box plot comparing the classical monocyte pseudobulk module scores of the leading edge genes used in Fig. 1f (union of female (F) and male (M) gene sets) in an acute COVID-19 CITE-seq dataset from Liu *et al*[1]. Both M (n=50) and F (n=9) subjects are included in all three groups (HC n=13, less severe n=21, more severe n=25). Each dot represents a sample. P values from the indicated two-group comparisons are shown. P values were generated using the moderated T statistics from linear model of the limma package in which samples from the same donors were treated as duplicates.

4. Why run baseline correlations as a single correlation analysis, rather than take baseline features (possibly reduced to only those differing between groups) and perform some form of multi-variate regression / lasso analysis to find the predictive features, there may be more than GPR56+ CD8 tcells.

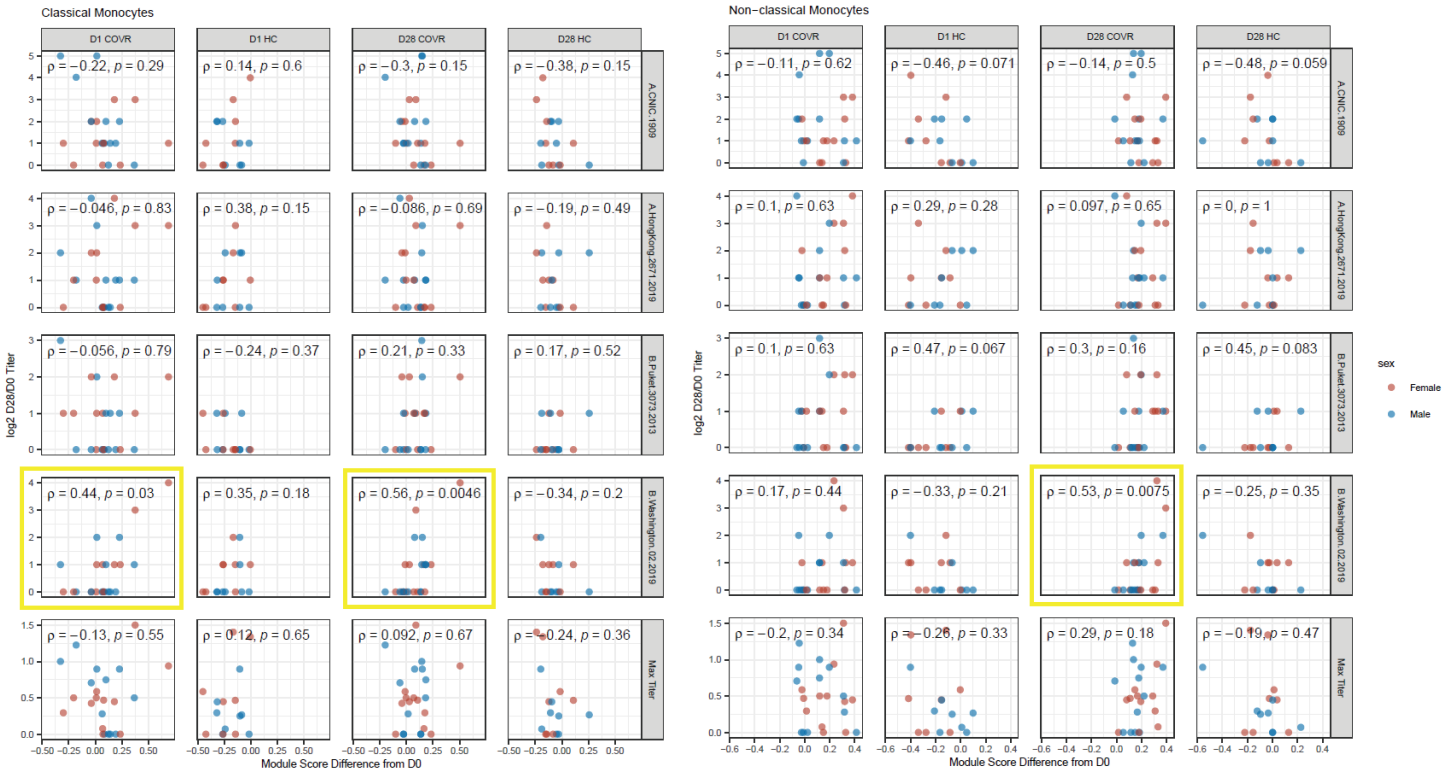
   We thank the reviewer for this suggestion. We have added the results of the multi-variate analysis in Supplementary Information Fig. 4. We ran elastic net models, combining ridge and lasso regression, with Day 0 cell frequencies from both flow cytometry and CITE-seq single cell data as predictors. Consistent with previous results (Extended Data Fig. 4a,b), early effector-like CD8+ T-cell frequency is one of the most significant predictors for the Day 1 IFN response as measured by circulating serum INFG protein concentration or IFNG transcriptional signature score in immune cells. We have added this as lines 257-262. Interestingly, this analysis also reveals CD38+ CD8+ T-cells and CD11c+ DCs as potential baseline correlates of the Day 1 IFNG response.

5. Did you see any difference in GPR56+ CD8 tcells between males and females at baseline?

   As shown in Fig. 3b and 3e, there is significant difference between COVR males and females at baseline (D0) in: 1) the level of surface GPR56 expression on CD8+ T-cells with an effector memory phenotype (CD8 EM), and 2) the frequency of GPR56+ CD8 EM cells. Within these cells, we did not detect any differentially expressed genes or enriched gene sets between COVR males and females that met the FDR cutoff of 0.05. However, we detected significant increases in IFNG mRNA expression in these cells on day 1 post influenza vaccination in COVR-M but not in the other groups (Fig. 3f). Thus, together these cells are a contributor of the heightened IFN response in COVR-M. See also the new *in vitro* stimulation data further supporting this hypothesis (Fig. 3i,j and Extended Data Fig. 4j; lines 302-316).

6. Was there a relation between reversal and the flu titers?

   Per the review's suggestion, we conducted this analysis and found that the extent of the shift in the monocyte signature scores in the COVR group is significantly correlated with the day 28 titer log-fold change (day 28 vs. day 0) for the B/Washington/02/2019 strain (as shown in the Rebuttal Fig. 1 below). The correlation can be detected in classical monocytes as early as day 1 and more significantly in both classical and non-classical monocytes at day 28. Further analyses are needed to validate and better understand the biological underpinnings of this relationship.



Rebuttal Fig. 1. Scatterplots showing the correlation between the innate immune receptor module score difference in classical (top) and non-classical (bottom) monocytes from D0 (x-axis) and the Day 28/Day 0 microneutralization titer log2 fold-change (y-axis) for each of the four strains in the seasonal influenza vaccine (top 4 rows), or the maximum standardized influenza vaccine titer (among the four strains in the vaccine; bottom row). Each column represents a different study time point and subject group combination. The x-axis module score difference is calculated between the date indicated in the column header and D0. Significant ($p < 0.05$) correlations are highlighted with yellow boxes.

7. Why are we assuming that reversal is a good thing? Obviously from a flu titer perspective it is not. As the authors themselves state in the discussion, it may be a trained immunity type phenotype that ideally one would want to keep.

We agree with the reviewer that we currently lack information to assess whether the partial reversal is good or bad functionally. We have added clarifications to the text to avoid giving the impression that reversal is the desired outcome (see lines 395-398; 469-470). We have also carefully checked and finetuned our writing to avoid using "reversal" to improve precision, e.g., we used "partial reversal" or "shifted towards the baseline of healthy individuals" (e.g., lines 466-469).

The question poised in the last section is "Does the influenza vaccine shift immune states toward normal?" (the authors have this even in the figure cartoon). Yet ultimately the answer is limited to looking at monocytes, where I suspect a more global perspective would have yielded a much broader of this important question view (e.g. dimensionality reduction of cell composition/state?). Moreover I would consider switching the visuals of this so that functionality is highlighted (e.g. place genes on KEGG nodes). The present heatmaps are not very informative in building a picture in ones mind.

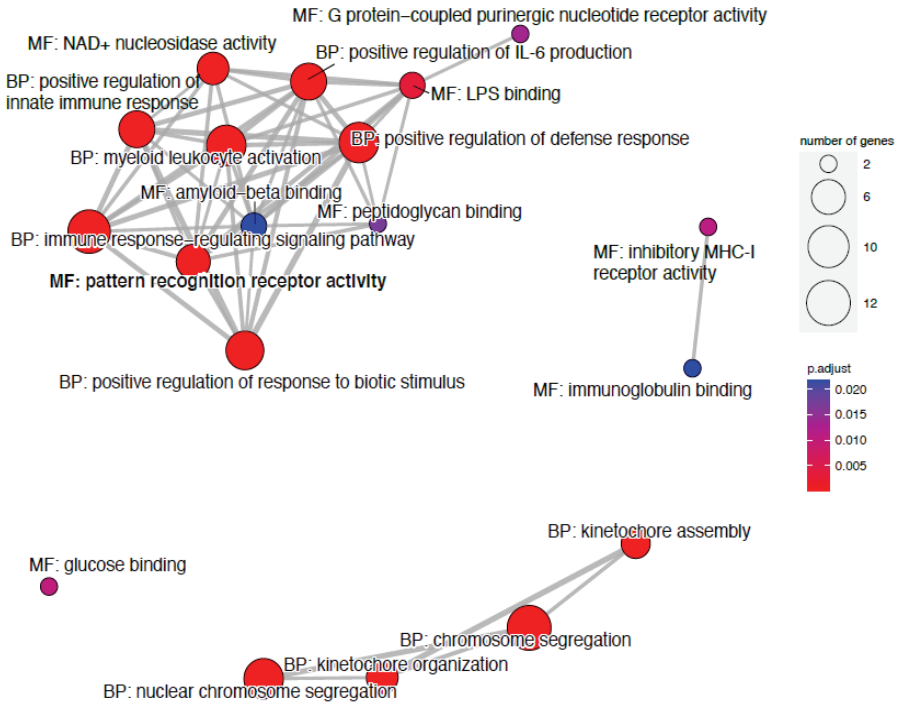
We agree with the reviewer that a more global perspective may yield additional insights, including evaluation of cell types beyond monocytes and additional biological processes. Towards this goal, we provide the differential expression and GSEA results with different statistical contrasts for all cell types in Extended Data Tables 5 and 6,

respectively. Readers can utilize those data to examine additional biological processes/signatures that were: 1) different between the COVR and HC at baseline (before influenza vaccination), and 2) moved closer to the HC baseline after influenza vaccination (COVR-M/F vs. baseline HC-M/F contrasts).

In the manuscript we opted to streamline the presentation and avoid bringing out a list of potential signals. Instead, we focus on monocytes and gene expression changes therein because we had already presented that as a major imprint (Fig. 1f,g), monocytes are known to be capable of stably shifting their statuses after inflammatory inputs ("trained immunity"), and we found both the imprint and its shift towards the baseline of the HCs to be robust in both the bulk transcriptomic and CITE-seq data.

The primary purpose of the heatmaps in Fig. 4 is to show the partially reversed/shifted genes individually in the signature. Based on the reviewer's suggestion, we did over-representation tests on the reversed genes and annotated the genes in the heatmaps with two enriched gene sets (*pattern recognition receptor activity* and *myeloid leukocyte activation* – see Fig. 4d) to provide additional functional information on the genes showing reversal behavior. We have also moved the heatmap for non-classical monocytes to Extended Data Fig. 5 to make Fig. 4 less dense and easier to read.

We also explored a different way of visualizing the partial-reversal signature as shown below: a network of over-represented gene sets from the union of male and female reversed/shifted genes in classical monocytes showing that the signature is driven by innate immune receptor and associated pathway genes. However, we opted to not show this visualization in the manuscript because it can be difficult to interpret for some readers and could become overwhelming if we also provide the names of the shifted genes in the diagram.



Rebuttal Fig. 2. A network diagram of over-represented Gene Ontology (GO) gene sets among classical monocyte-reversed/shifted genes in COVR individuals (both males and females) after influenza vaccination. Reversed/shifted genes are those in the monocyte immune receptor transcriptional signature that moved towards the baseline healthy level at both D1 and D28 after vaccination. Edges represent two gene sets that share one or more reversed genes as their members. MF: Molecular Function; BP: Biological process.

Reviewer #4:

1. In figure 1 and related data, bulk FACS and RNA analysis showed an increased monocyte signal in COVR-M that however disappeared in the CITE-seq analysis, suggesting that that the signature was mainly driven by cell frequencies. As increased cell frequencies are a biological finding with potential biological consequences, the authors may be slightly too dismissive about this.

   We agree that the increased frequencies of monocytes are important biological findings. On lines 116-127 of the previous manuscript, we discussed that the cell frequency differences between different groups may explain the bulk RNA-seq signals we detected, which motivated us to perform single-cell analyses to untangle signals driven by changes in cell frequencies from those resulting from differences in intracellular gene expression. We have added clarification in that section of the manuscript to highlight the increased monocyte frequencies in COVR-M (lines 116-122).

2. The term monocyte signature in 1f and g needs to be filled with more life, as this term per se does not mean that they are more activated. More generally, when a monocyte cluster has a stronger or weaker monocyte signature it is actually not so obvious what this means, it makes more sense in mixed populations. Please discuss the genes in there in a bit more detail, as I seem to spot also some negative regulators in here.

   We agree that a more informative and consistent name should be used for this signature. As the leading edge genes in this monocyte signature include many surface receptors (see lines 142-146 in the revised manuscript), we now refer to it as the "innate immune receptor" signature throughout the manuscript.

3. I see in extended fig. 2b a highly significant negative correlation in the F, with a slope that is roughly half that of the (positive) slope in M. The authors therefore cannot claim that in F, influenza infection had no impact, it had the opposite impact of that in M. Please correct these statements and discuss in the context of your data.

   We apologize for this confusion and have clarified the information provided in this panel with better axis labels. The scatter plots (separately for males and females) show the concordance between two independent influenza seasons (2009 and 2010) because we are interested in knowing whether the difference between the pre- and post-infection statuses are consistent across the two independent cohorts/seasons. Each point in Extended Data Fig. 2b represents a gene and the axes correspond to the difference between before (start of season) and after (end of season) influenza infection for the two cohorts (the x-axis shows the difference in the 2009 cohort and the y-axis shows the difference in the 2010 cohort).

   The males show significant positive correlation, suggesting that there was consistency between the two seasons/cohorts. Even though the correlation in the female group has a significant p value (due to the many genes used to assess the correlation), the correlation is negative (and thus difficult to interpret), and its absolute strength is weaker than that of the positive correlation in the males. We therefore thought that the male signals are more clearly interpretable and thus focused on comparing those to our COVID-19 signatures.

4. Line 175-77 where do you show this data? You only state in the text. Please indicate these genes in the plot S2b or do a GSEA as in S2c.

   We apologize for failing to include the list of genes with lower expression. We have now added an additional tab ("Overlap with Monocyte Signature") in Extended Data Table 7 to show the data.

5. Line 179-180, "although in general some of these imprints are likely pathogen-dependent": a politician could not be more vague, what do you mean?

   We apologize for this unclear statement. We meant to say that although we detected a significant overlap of genes whose expression remain deviated from baseline several months after either a SARS-CoV-2 or influenza A infection, our data also support that each infection is likely to leave unique imprints given that the overlap is far from explaining all the post COVID-19 signatures we observed (e.g., those in Extended Data Fig. 1e and Extended Data Table 4). The biology behind the unique vs. shared imprints of different infections would require more investigation and is beyond the scope of this study. We have revised the text to clarify our point (see lines 182-184).

6. The increased IFNg response is intriguing. in S4a, and b, how do these correlations look for HC? Please add this data and discuss. Also, you mention later (line 315) very much in passing your data supporting that also NK cells contribute to increased IFNg, please add this data, maybe here or later, but without this the study is incomplete.

   We have added an additional supplementary figure to show the correlation between the D0 early effector-like CD8+ cells and day 1 IFNG response for the HCs (Supplementary Information Fig. 5a,b). We did not detect a significant correlation in the HC group. This observation further suggests that the increased expression of IFNG in these CD8+ cells is specific to the COVR individuals, especially the COVR-M.

   In Supplementary Information Fig. 7, we have added panels showing the IFNG transcriptional response in NK cells (similar in part to the GPR56+ panels in Extended Data Fig. 4d). As mentioned in the previous version of the manuscript (previous line 315 as indicated by the reviewer), the $CD56^{hi}$ $CD16^{lo}$ NK cells show significant day 1 increase in IFNG expression compared to day 0 in COVR-M (also in HC-F). Although not statistically significant, the $CD56^{lo}$ $CD16^{hi}$ NK cells exhibit similar trends (data not shown). However, our *in vitro* cytokine stimulation data show insignificant differences in the frequency of IFNγ+ NK cells between COVR-M and other groups after cytokine stimulation (Rebuttal Fig. 4 below). See lines 335-341 where we discuss the NK cell results in the manuscript.

7. Fig S4d, would this be time-correlated with higher n numbers? What does the different direction of travel of the trends between m and f mean?

   As the correlation between time since diagnosis (TSD) and the GPR56+ CD8 cell frequencies is statistically insignificant in COVR-M or COVR-F, we cannot interpret the direction in a definitive way. The opposite direction between males and females could be due to noise. Having more samples could increase robustness, but it is unclear whether it would yield a statistically significant correlation. To avoid potential confusion, we now simply report the statistics in the main text (see lines 277-278) and have removed this figure in the current version of the manuscript.

8. 3f How would this look for gpr56 neg cells? Please add.

   We show the same plots as Fig. 3f for the GPR56- cells in Supplementary Information Fig. 5c.

9. Source of IFNg: The extended omics analysis leads the authors to conclude that bystander-activated GPR56+ CD8 cells whose phenotype is somewhere between EM and TEMRA CD8 cells produce the early IFNg in an Ag-independent way, rather driven by cytokines. I think it is crucial and feasible to show the in vitro, stimulating PBMC with the cytokine(s) in question, ideally more than one cocktail, and stain by icFACS for all CD8 subsets in question, and in addition NK cell (subsets). This will also the authors to show in a real stimulation experiment that their hypothesis is true, and it will add data on the NK contribution to early IFNg.

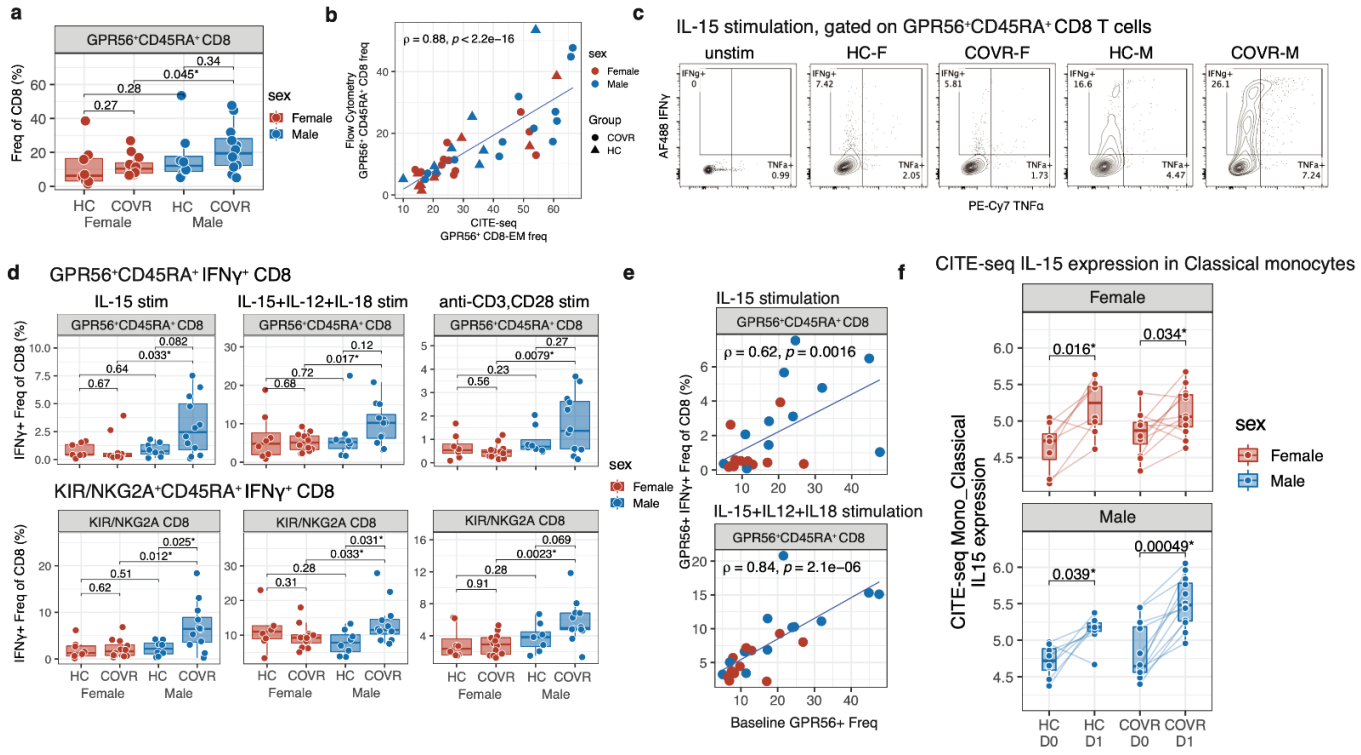
We thank the reviewer for these suggestions. Per the recommendations above, we performed *in vitro* stimulation assays using baseline (D0) PBMCs from all CITE-seq subjects (n=40) to evaluate intracellular IFNγ responses. We used the following four cocktails to evaluate the response to cytokine stimulation (we included anti-CD3/anti-CD28 as a positive control for IFNγ production after T-cell stimulation):

1) IL-15+IL-12+IL-18
2) IL-15
3) IL-18
4) IL12+IL18
5) anti-CD3+anti-CD28 as a control

Consistent with our CITE-seq data, at baseline before stimulation our flow cytometry data indicate a higher frequency of GPR56+ CD45RA+ cells in COVR-M than COVR-F (Rebuttal Fig. 3a). The frequency of GPR56+ cells as measured by flow cytometry is also correlated with that measured by CITE-seq (Rebuttal Fig. 3b). The frequency of IFNγ+ GPR56+ CD8+ T-cells was higher in COVR-M than COVR-F and HC after stimulation with IL-15 or IL15+IL-12+IL-18 (Rebuttal Fig. 3c,d; see also Fig 3i,j). A similar trend was observed in the other stimulation conditions, although IL-18 alone and IL-12+IL18 resulted in less robust differences (data not shown).

Given that the surface levels of GPR56 appeared to decrease upon stimulation (consistent with previous reports[2,3]), we also gated the CD8+ VM-like T-cells using other surface markers identified in the literature: CD3+, CD8+, CD45RA+, KIR+ and/or NKG2A+[4,5] (labeled as KIR/NKG2A CD45RA+ here and in figures in the revised manuscript – Fig 3j). Consistently, this gating showed that COVR-M had a higher frequency of IFNγ+ cells upon stimulation (Rebuttal Fig. 3d bottom row, see also Fig. 3j). Overall, the frequency of GPR56+ CD8+ T-cells prior to stimulation is correlated (or predicts) the frequency of IFNγ+ GPR56+ CD8+ T-cells after stimulation in COVR subjects (Rebuttal Fig. 3e). The correlation appears due primarily to: 1) higher frequencies of GPR56+ cells at baseline prior to stimulation in COVR-M (blue dots in both plots), 2) higher frequencies of IFNγ+ VM-like cells after *in vitro* stimulation compared to COVR-F (Rebuttal Fig. 3d), and 3) a more robust IFNγ response in COVR-M after stimulation of those cells with IL-15 (Rebuttal Fig. 3c). Thus, these together help explain why COVR-M had more robust IFNγ responses early after influenza vaccination.

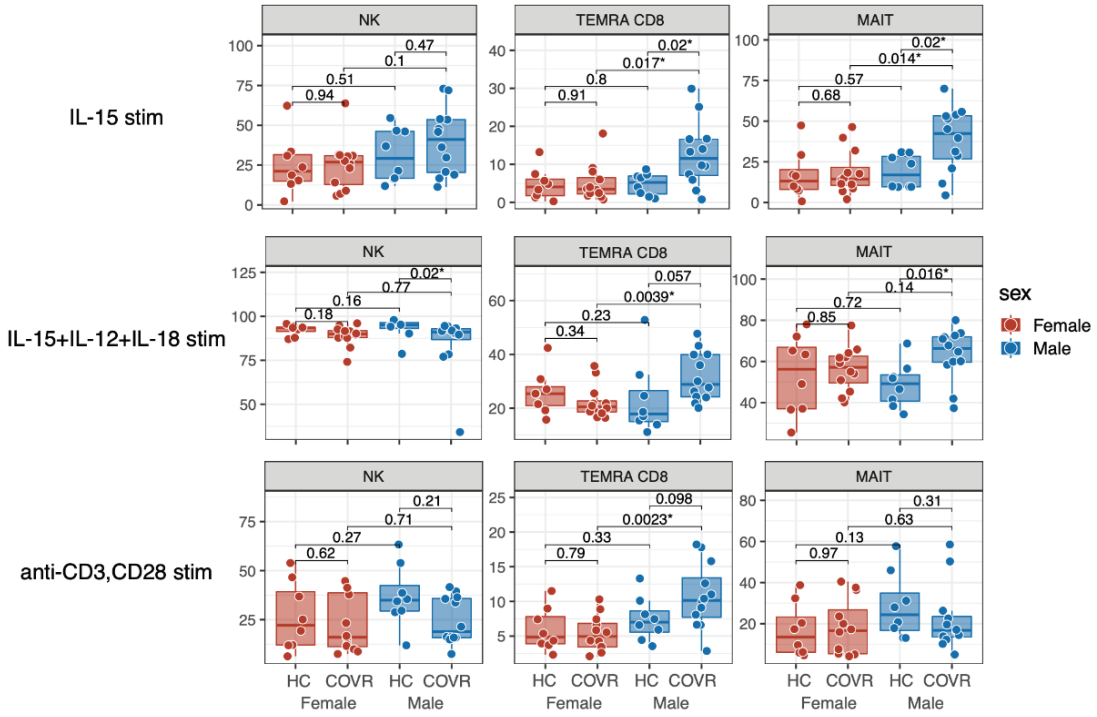
Since IL-15 stimulation revealed robust differences between COVR-M and other groups (Rebuttal Fig. 3d, left column), we also used our CITE-seq data to identify potential *in vivo* cellular sources of IL-15 in response to influenza vaccination. We found increased IL-15 mRNA expression at D1 compared to baseline (D0) in classical monocytes in both males and females (Rebuttal Fig. 3f; also shown in Fig. 3k), although the average magnitude of increase was the highest in COVR-M. This more pronounced increase in IL-15 induced by influenza vaccination could contribute to the higher IFNγ production in the VM-like CD8+ T-cells *in vivo*. Together with the higher frequency of these VM-like cells at baseline and their elevated IFNγ responses upon IL-15 stimulation in COVR-M *in vitro*, our results point to an IL-15-driven circuit that helps explain the more robust IFNγ responses after influenza vaccination in COVR-M.



Rebuttal Fig. 3. **a,** Boxplots showing the flow cytometry frequencies in GPR56+ CD45RA+ VM-like CD8 T-cells in all subject groups at D0 in non-stimulated controls. **b,** Scatterplot showing the correlation between the frequencies of the GPR56+ CD45RA+ VM-like CD8 T-cells from flow cytometry and the GPR56+ CD8 EM cells from CITE-seq at D0. **c,** Flow contour plots showing the representative GPR56+ CD8 T-cell IFNγ secretion after IL-15 stimulation (as an example) in the four subject groups. **d,** Boxplots showing the frequencies of IFNγ+ GPR56+ CD45RA+ VM-like CD8 T-cells (top) and IFNγ+ KIR/NKG2A+ CD45RA+ CD8 T-cells (bottom) after stimulation with IL-15 (left), IL-15+IL-12+IL-18 (center) or anti-CD3/anti-CD28 (right) within the four subject groups. **e,** Correlation of baseline (non-stimulated) GPR56+ CD45RA+ VM-like CD8 T-cell frequency with the frequency of IFNγ+ GPR56+ CD45RA+ VM-like CD8 T-cells after IL-15 stimulation (top) or IL-15+IL-12+IL-18 stimulation (bottom) in COVR-M (blue dots) and COVR-F (red dots). **f,** Boxplots showing the CITE-seq normalized mRNA expression of IL-15 in classical monocytes at D0 before and D1 post influenza vaccination in all four subject groups in females (top) and males (bottom).

As the reviewer suggested, in general NK cells is also a major cell type producing IFNγ *in vivo*. In our *in vitro* stimulation data, we noticed that NK cells did not clearly show higher IFNγ+ responses in COVR-M (Rebuttal Fig. 4, left column). However, we detected higher responses in other candidate cell subsets, including CD8+ T-cells exhibiting a CD45RA+ CD45RO- CCR7- TEMRA phenotype, as well as in MAIT cells (CD3+ TCRVa7.2+) (Rebuttal Fig. 4 center and right columns, respectively). This suggests that in addition to the GPR56+ and VM-like CD8+ T cell subsets, CD8+ TEMRA and MAIT cells are potential contributors to the higher IFNγ response in COVR-M. We added discussion on these points (see lines 334-347, ED Fig. 4j and Supplementary Information Fig. 6b.)

In summary, our *in vitro* stimulation data further support the hypothesis that prior mild COVID-19 increases the baseline frequency of VM-like CD8+ T-cells in males and these cells' propensity to produce IFNγ with cytokine stimulation alone.



Rebuttal Fig 4. Boxplots showing the frequencies of IFNγ+ NK (left), IFNγ+ CD45RA+ CD45RO- CCR7- TEMRA CD8 (center) and IFNγ+ MAIT cells (right) after stimulation with IL-15 (top), IL-15+IL-12+IL-18 (middle) or anti-CD3/anti-CD28 (bottom) within the four subject groups.

10. For me, the most difficult part is figure 4: As a general comment, I think the figure must contain data showing how this signature changes in HC over time post vaccination, as the notion of a reversal of immunodepression makes sense if this signature is not changed in HC, or at least if the d28 time point shows convergence between COVR and HC samples.

We appreciate the reviewer's suggestion and agree that it would be informative to show the changes in HC as well. We have thus added the same set of panels (i.e., Fig. 4 b-d in the previous manuscript) but for the HCs (see new Fig. 4b-d and Extended Data Fig. 5c in the current manuscript). In contrast to the COVRs, the HCs have on average returned to their baseline level (pre-vaccination) by day 28 in both the females and males (Fig. 4b,c). This observation suggests that the baseline HC status is relatively immutable (by influenza vaccination) and thus further supports our approach of using the HC baseline as a reference to assess whether the COVRs are shifting towards "normal" after influenza vaccination.

In addition, our HC analysis reveals that while the monocyte (immune receptor) module scores increase towards the HC baseline in the COVR group at D1, they decrease in the HC group in both sexes (Fig. 4b,c). The trend is similar between the classical and non-classical monocytes. Further investigation is needed to better understand this phenomenon, but we speculate that this may reflect feedback mechanisms: the "normal" monocytes in HCs respond to acute vaccine-induced inflammation by downregulating their immune receptors and associated pathway genes (e.g., a negative feedforward mechanism to avoid overresponding), while the "depressed" monocytes in COVR individuals attempted to shift towards the normal state by increasing the expression of these genes (see lines 372-378). Together, these post-vaccination response differences in monocytes between the HC and COVR groups further illustrate how a prior infection could influence the response to the next perturbation, and how these response differences shape homeostatic immune set points.

11. In this context, while the discussion comparing your findings to papers talking about "training" is extensive and can be shortened, I think it is important to mention that the d28 time point is beyond the expected half life of blood monocytes, so resetting the set-point must happen in precursors or must be maintained by non-monocyte-intrinsic mechanisms.

    We agree and have added additional discussion and references to suggest that the shift may be attributed to long-lived hematopoietic stem and progenitor cells (HSPCs) but could be more detectable in the monocytes given that the frequency of circulating HSPCs is often too low for robust assessment in blood (see lines 481-485).

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #2 (Remarks to the Author):

Overall the second revision version of this manuscript is much improved. I feel that although this study is still tremendously complex and the many associations somewhat difficult to disentangle the authors have done a good job of simplifying things as much as possible without loosing scientific rigor. Now the message of a perturbed immune state, induced by mild COVID-19, impacting males and females differently upon secondary challenge is more understandable than in previous versions of the manuscript. The added data from prior studies suggests generalisable results. My concern regarding a possible role for prior immunity to flu still stands and cannot be overcome but I think the paper still offers valuable insights into human immune system regulation. I think the discussion of these new findings in relation to trained immunity is valuable as I personally consider all of these adaptive changes different versions of a similar biological phenomenon. All in all the new paper is much improved, the study is unique, technologically well executed and well written.

Referee #3 (Remarks to the Author):

The authors have answered all of my questions satisfactorily. I do however wish to convey that they should pay additional attention to the following two points:
1. In many figures the choice to display the significance values in the plots (especially box plots) and occasionally to show outlier samples has resulted in plots where the effect size appears very small whereas in fact, much of the y-axis is showing an irrelevant scale. I would re-consider this.
2. In the analysis of monocyte recovery - the majority of the analysis is performed in longitudinally within each gender, as it should, and shows signal in the females but not the males. If this signal is also observed cross sectionally within a time point, it would be helpful to strengthen the point (i suspect it does not).
3. Would be good to check if there a relation (negative ?) between the magnitude of CD8 responses and ability of monocytes to recover ?

Referee #4 (Remarks to the Author):

The authors have done a great job incorporating suggestions by me and other referees, and the manuscript is improved due to the novel data and analysis. All my major concerns were addressed, in particular the in vitro stimulation and IFNg measurement, adding an important wet-lab confirmation to the ex vivo analysis presented before. I congratulate the authors for this study!

Author Rebuttals to Second Revision:

Point-by-point responses to reviewers

Reviewer #1: no new comments were received for this round.

Reviewer #2:

Overall the second revision version of this manuscript is much improved. I feel that although this study is still tremendously complex and the many associations somewhat difficult to disentangle the authors have done a good job of simplifying things as much as possible without loosing scientific rigor. Now the message of a perturbed immune state, induced by mild COVID-19, impacting males and females differently upon secondary challenge is more understandable than in previous versions of the manuscript. The added data from prior studies suggests generalisable results. My concern regarding a possible role for prior immunity to flu still stands and cannot be overcome but I think the paper still offers valuable insights into human immune system regulation. I think the discussion of these new findings in relation to trained immunity is valuable as I personally consider all of these adaptive changes different versions of a similar biological phenomenon. All in all the new paper is much improved, the study is unique, technologically well executed and well written.

We thank the reviewer for the positive response and appreciate the feedback from this reviewer that made these improvements possible.

Reviewer #3:

The authors have answered all of my questions satisfactorily. I do however wish to convey that they should pay additional attention to the following two points:

Thank you and we have addressed the additional suggestions as follows.

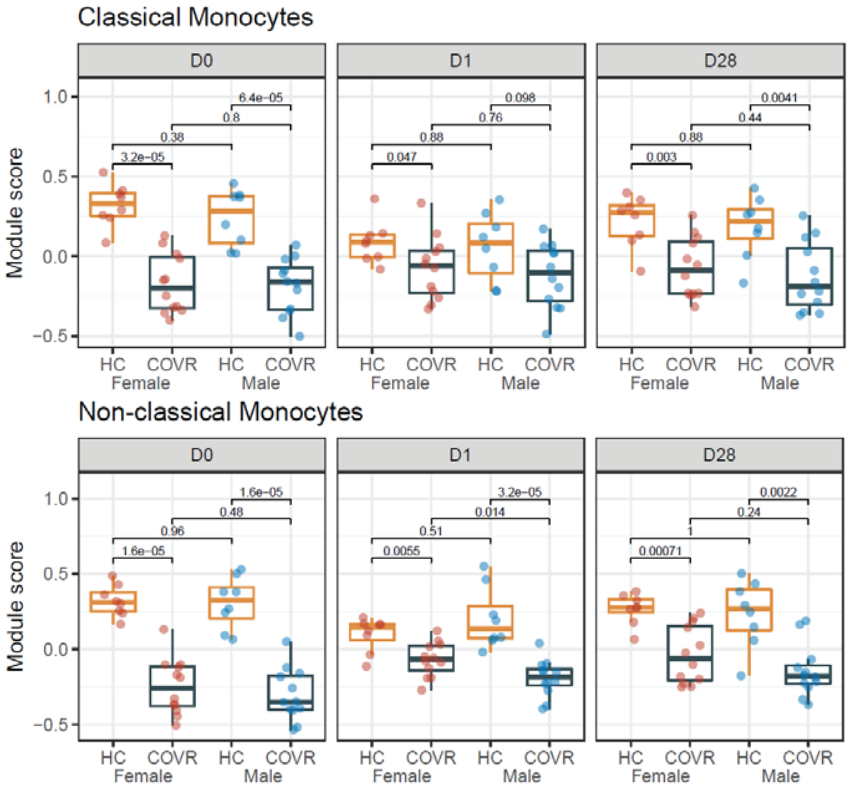
1. In many figures the choice to display the significance values in the plots (especially box plots) and occasionally to show outlier samples has resulted in plots where the effect size appears very small whereas in fact, much of the y-axis is showing an irrelevant scale. I would re-consider this.

We show the statistical significance values and individual points in the plots in accordance with the journal's formatting guidelines. We agree that in some of the plots the group differences were more difficult to distinguish visually because the y-axis was compressed. We have updated the few plots that the reviewer was likely referring to in order to show the effects in a more visually distinguishable manner. In Fig. 1c, we added a break to reduce the gap between the single

outlier and the rest of the points. For the time trend plots (e.g., in Fig. 4b,c), we changed the style to increase the contrast between the boxplot and the individual points.

2. In the analysis of monocyte recovery - the majority of the analysis is performed in longitudinally within each gender, as it should, and shows signal in the females but not the males. If this signal is also observed cross sectionally within a ~~time point, it would be helpful to streng~~then the point (i suspect it does not).

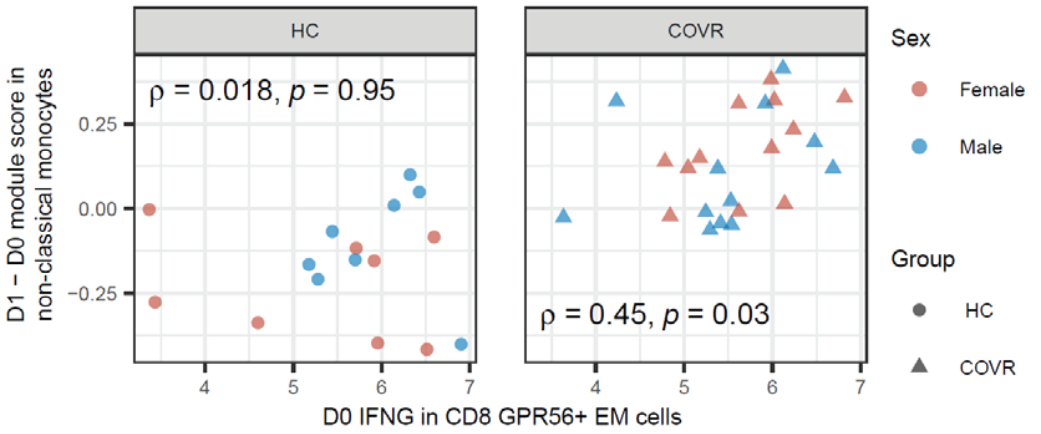
The cross-sectional analysis shown below paints a similar picture to the longitudinal analysis included in our manuscript in which we see both COVR females and males moved towards the healthy controls after vaccination. Additionally, the cross-sectional analysis better illustrates that there is no significant difference between the HC females and males at any of the timepoints. Between the COVR females and males, the D1 increase in non-classical monocytes is more prominent in COVR-F than in COVR-M, resulting in a significant difference between these two groups at day 1.



Comparison of innate immune receptor signature among groups at baseline, day 1 and day 28 after influenza vaccination. P-values shown are from two-tailed Wilcoxon rank-sum test.

3. Would be good to check if there a relation (negative ?) between the magnitude of CD8 responses and ability of monocytes to recover ?

We looked into CD8 EM cells and the GPR56+ EM subset but did not detect a significant correlation between the CD8 day 1 IFNG changes and the recovery of the monocytes at either D1 or D28. However, there seems to be an association between the D0 (baseline, pre-vaccination) IFNG mRNA expression level in CD8 GPR56+ EM cells and day 1 change of the innate immune receptor signature in non-classical monocytes. Further analyses are needed to dissect their potential relationship, but we hypothesize that the baseline IFNG expression in these EM cells may reflect an overall inflammatory state in the COVR individuals that might correlate with how they will respond to future challenges, such as the influenza vaccination they received in this study.



Correlation between day 0 IFNG expression in CD8 GPR56+ EM cells and day 1 change of the innate immune receptor signature in non-classical monocytes. Shown are Spearman's rank correlation and unadjusted p values.

Reviewer #4:

The authors have done a great job incorporating suggestions by me and other referees, and the manuscript is improved due to the novel data and analysis. All my major concerns were addressed, in particular the in vitro stimulation and IFNg measurement, adding an important wet-lab confirmation to the ex vivo analysis presented before. I congratulate the authors for this study!

We are grateful for the excellent recommendations made by the reviewer during the revision process.