Your manuscript has been checked for clarity and against journal policies and formatting style. The issues listed below must be addressed; failure to do so will cause delays in acceptance.

For further information, please see our [formatting instructions](https://www.nature.com/documents/ncomms-formatting-instructions.pdf).

Please highlight all changes in the manuscript text file, either using the track changes feature in Microsoft Word or coloured highlighting in LaTeX.

Please include your response to these requests in the space provided and return this checklist with your final submission.

|  |  |
| --- | --- |
| **EDITORIAL REQUESTS:** | **AUTHOR RESPONSE:** |
| **POLICIES AND CHECKLISTS** | **POLICIES AND CHECKLISTS** |
| An updated editorial policy checklist must be completed and uploaded as a related manuscript file with the revised manuscript. All points on the policy checklist must be addressed; if needed, please revise your manuscript in response to these points. Please note that this form is a dynamic 'smart pdf' and must therefore be downloaded and completed in Adobe Reader, instead of opening it in a web browser. https://www.nature.com/authors/policies/Policy.pdf | **Done** |
| Please find attached a reporting summary that includes comments on how to revise it in line with our policies and requests the addition of further information in the text. An updated reporting summary must be completed and uploaded as a supplementary information file with the revised manuscript. This checklist is published alongside your manuscript online. Please note that this form is a dynamic 'smart pdf' and must therefore be downloaded and completed in Adobe Reader, instead of opening it in a web browser. https://www.nature.com/authors/policies/ReportingSummary.pdf | **Done (including edits to manuscript where necessary)** |
| **TITLE PAGE (page 2 of our formatting instructions)** | **TITLE PAGE (page 2 of our formatting instructions)** |
| The abstract should briefly discuss the background and context of the work, followed by the major results and conclusions of the paper. The discussion of the current work should begin with "Here we report" or an equivalent phrase and should be written in the present tense. Please ensure your abstract conforms to this structure. Please also ensure that the abstract does not exceed 150 words. | **Edited a sentence in the abstract to begin ‘Here we illustrate’. Now 150 words** |
| **MAIN TEXT (pages 1 to 3 of our formatting instructions)** | **MAIN TEXT (pages 1 to 3 of our formatting instructions)** |
| The following sections must be present in the manuscript and must begin with a heading: Abstract, Introduction, Results, and optional Discussion and Methods. All other main text section headings should be removed or renamed. | **This is not possible given the current format of the paper, as we explained when we resubmitted the paper. Please do contact me about this to discuss if it is a problem.** |
| Please divide the Results section into subsections, each with a title of 60 characters or fewer including spaces. | **We do not have a results section (see above)** |
| We allow only one level of subheadings in the Results section. Please remove secondary subheadings. | **We do not have a results section (see above)** |
| Please provide a Discussion section in the main manuscript file. | **We have changed ‘Implications’ to ‘Discussion’** |
| **METHODS AND DATA (page 3 of our formatting instructions)** | **METHODS AND DATA (page 3 of our formatting instructions)** |
| A full Methods section, divided into subsections and subheadings, must be provided in the main manuscript. There is no word limit to this section. | **This has now been added** |
| Please ensure the Methods section is named as 'Methods'. | **Done** |
| Please provide more explicit details of the methods in the main manuscript; we discourage referencing published work and ask authors to explicitly discuss all methods, including computation approaches, in the manuscript. | **Done** |
| All manuscripts must include a Data Availability statement as a separate section after the Methods section but before the References. For acceptable examples, see: https://www.nature.com/documents/nr-data-availability-statements-data-citations.pdf The Data Availability statement should include:  - Accession codes with hyperlinks for deposited data   - Other unique identifiers (such as DOIs and hyperlinks for any other datasets)   - At a minimum, a statement confirming that all relevant data are available from the authors   - If applicable, a statement regarding data available with restrictions  - If a dataset has a Digital Object Identifier (DOI) as its unique identifier, we strongly encourage including this in the Reference list and citing the dataset in the Data Availability Statement | **Done, linking to a document that describes in full how the data were generated.** |
| In particular, please specify the UK BioBank exact datasets that you analysed in the 'Data Availability' section of the manuscript. The information should be clear and enable independent researchers to access the same datasets that you have analysed in this manuscript. |  |
| Nature Research policies (https://go.nature.com/data-availability-AIP) strongly encourage deposition of research data in public repositories. In some cases this is mandatory, and you may have been previously advised if that was the case. If you need help depositing and curating your research data (including raw and processed data, text, video, audio and images) you should consider: - Contacting Springer Nature’s Research Data Helpdesk (https://go.nature.com/helpdesk-AIP) for advice - Finding a suitable data repository (https://go.nature.com/RD-policies-AIP) for your data - Uploading your data to Springer Nature’s Research Data Support service (https://go.nature.com/RDS-AIP) Research Data Support is an optional Springer Nature service. There are fees (https://go.nature.com/RDS-pricing-AIP) for using this service, however, if you receive funding from the Wellcome Trust or are affiliated to a Wellcome Centre you can use Research Data Support at no cost. See https://go.nature.com/wellcome-RDS-AIP for more information. If you choose to use Research Data Support, please do not submit your revised manuscript until you have been supplied with the DOI for your data by the Research Data Support team. Please also ensure that you update your Data Availability statement with this DOI and the information provided by the team. Please provide a unique identifier for the data (for example a DOI or a permanent URL) in the data availability statement, if possible. If the repository does not provide identifiers, we encourage authors to supply the search terms that will return the data. For data that have been obtained from publicly available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be included in the reference list and cited where relevant. Alternatively, include the data in the Supplementary Information. For datasets for which mandatory deposition is not required and the data can only be shared on request, please explain why in your Data Availability Statement and in your response here.  Please refer to our data policies here: http://www.nature.com/authors/policies/availability.html | **We explored this with UK Biobank and it is not possible.** |
| Please separate the 'Code Availability' and 'Data Availability' sections in the manuscript. | **Done** |
| A reference to the source data file should be added in the 'Data Availability' section, using the text “Source data are provided with this paper.” | **Source data for Figure 3 is available in Supplementary Table 1.** |
| Please note that you have marked the clinical data section as “involved in the study” in the reporting summary and in the editorial policy checklists; however, no clinical data are provided. Please ensure that the checklists accurately reflect this. | **Updated this field to NA** |
| Similarly, you have not marked the field for “Human research participants” under additional policy considerations section of the editorial policy checklist, although information on the corresponding section has been marked in the reporting summary. | **Updated the editorial policy checklist** |
| Within the manuscript, please provide an explicit statement as to whether ethical approval was waived, or unnecessary, with any relevant exemption codes provided. This is especially important as you have referenced the use of COVID-19 tested patient’s data obtained from UK biobank, in the manuscript. | **Added a statement of ethical approval to the methods.** |
| **DISPLAY ITEMS (pages 4 and 5 of our formatting instructions)** | **DISPLAY ITEMS (pages 4 and 5 of our formatting instructions)** |
| The use or adaptation of previously published images is strongly discouraged. If this is unavoidable, please request the necessary rights documentation to re-use such material from the relevant copyright holders and return this to us when you submit your revised manuscript. Please check whether your manuscript or Supplementary Information contain third-party images, such as figures from the literature, stock photos, clip art or commercial satellite and map data. | **No previously published images used** |
| In particular, please indicate whether any of the figures were not created by any of the co-authors in this manuscript. | **All figures were created by the authors in this manuscript** |
| All figures must be cited in the order in which they appear in the text, figures, tables and boxes. If a figure needs to be referred to out of order, please use the text 'see below' instead of a citation. | **Done.** |
| Any abbreviations, symbols or colours present in your figures must be defined in the associated legends. | **Done.** |
| **SUPPLEMENTARY INFORMATION (page 5 of our formatting instructions)** | **SUPPLEMENTARY INFORMATION (page 5 of our formatting instructions)** |
| We do not edit Supplementary Information files; they will be uploaded with the published article as they are submitted with the final version of your manuscript. Any tracked changes should be removed from the file and the file should be provided as a PDF file. Supplementary Figures do not need to be provided separately. | **Supplementary Note file is now a separate PDF** |
| Supplementary Information must be provided as a single separate PDF file, not within the manuscript file. | **Done** |
| All Supplementary Information items (e.g. Supplementary Figures, Supplementary Tables, Supplementary Methods, Supplementary Notes, Supplementary Discussion, Supplementary References) must be included in one PDF document. Only Supplementary Movie, Audio, Data and Software files should be submitted separately from the Supplementary Information. | **Supplementary Table 1 is submitted as an Excel file, as this is over 2500 lines and would not be accessible as a pdf file.** |
| Supplementary References should appear at the end of the Supplementary Information file, and must be self-contained and numbered from 1. References mentioned in both the main text and the Supplementary Information should be part of both reference lists so that the Supplementary Information does not refer to the reference list in the main paper and vice versa. | **Done** |
| Please ensure that a Source Data file is included with your resubmission. Within the Source Data file, the relevant raw data from each figure or table (in the main manuscript and in the Supplementary Information) should be represented by a single sheet in an Excel document, or a single .txt file or other file type in a zipped folder. Uncropped blots and gel images should be pasted in and labelled with the relevant panel and identifying information such as the antibody used. An example of the Source Data file is available demonstrating the correct format: https://www.nature.com/documents/ncomms-example-source-data.xlsx The file should be labelled 'Source Data', with the title and a brief description included in your response here, and should be mentioned in all relevant figure legends using the template text below: "Source data are provided as a Source Data file." | **Source data that we can share are the results in Supplementary Table 1, that are used to generate Figure 3. We have added a statement to this effect.** |
| **PUBLICATION** | **PUBLICATION** |
| Your paper will be accompanied by a two-sentence Editor's summary, of between 250-300 characters including spaces, when it is published online. I have drafted the summary below. If you would like to make changes to this, please provide me with a suitably edited version. | **Great thank you** |
| Many published studies of the current SARS-CoV-2 pandemic have analysed data from non-representative samples from populations. Here, using UK BioBank samples, Gibran Hemani and colleagues discuss the potential for such studies to suffer from collider bias, and provide suggestions for optimising study design to account for this. | **Approved** |
| As part of our efforts to communicate our content to a wider audience, we endeavour to highlight papers published in Nature Communications on the journal’s Twitter account (https://twitter.com/NatureComms). If you would like us to mention authors, institutions or lab groups in these tweets, please provide the relevant twitter handles. | **@Garethjgriffith, @bristimtom, @tudballstats, @AnnieSHerbert, @Lindsey\_Pike, @ammegandchips, @mendel\_random, @Luisa\_Zu, @nm\_davies, @explodecomputer** |

|  |  |  |
| --- | --- | --- |
|  | **EDITORIAL REQUESTS:** | **AUTHOR RESPONSE:** |
| **1.** | **Data presentation:** Please ensure that data presented in a plot, chart or other visual representation format shows data distribution clearly (e.g. dot plots, box-and-whisker plots). When using bar charts, please overlay the corresponding data points (as dot plots) whenever possible and always for n ≤ 10. (Please see the following editorial for the rationale behind this request and an example <https://www.nature.com/articles/s41551-017-0079>). | |
| **Panels requiring revision:** |  |
| **2.** | **Statistics**: Wherever statistics have been derived (e.g. error bars, box plots, statistical significance) the legend needs to provide and define the n number (i.e. the sample size used to derive statistics) as a precise value (not a range), using the wording “n=X biologically independent samples/animals/cells/independent experiments/n= X cells examined over Y independent experiments” etc. as applicable. | |
| **Legends requiring revision:** |  |
| **3.** | Statistics such as error bars, significance and p values cannot be derived from n<3 and must be removed from all such cases.  *[SNTPS team: this sentence should always be in the template, but here, you can mention panels where there are clearly error bars for n<3 if you find any.]* | |
| We strongly discourage deriving statistics from technical replicates, unless there is a clear scientific justification for why providing this information is important. Conflating technical and biological variability, e.g., by pooling technically replicates samples across independent experiments is strongly discouraged. (For examples of expected description of statistics in figure legends, please see the following <https://www.nature.com/articles/s41467-019-11636-5> or <https://www.nature.com/articles/s41467-019-11510-4>). | |
| All error bars need to be defined in the legends (e.g. SD, SEM) together with a measure of centre (e.g. mean, median). For example, the legends should state something along the lines of “Data are presented as mean values +/- SEM” as appropriate.  All box plots need to be defined in the legends in terms of minima, maxima, centre, bounds of box and whiskers and percentile. | |
| **Legends requiring revision:**  *[SNTPS team: please describe to the author exactly what is missing (measure of centre (mean or median), the error bar definition, or the characteristics of the box plots)]* |  |
| **4.** | The figure legends must indicate the statistical test used. Where appropriate, please indicate in the figure legends whether the statistical tests were one-sided or two-sided and whether adjustments were made for multiple comparisons.  For null hypothesis testing, please indicate the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P values noted.  Please provide the test results (e.g. P values) as exact values whenever possible and with confidence intervals noted. | |
| **Legends requiring revision:** |  |
| **5.** | **Reproducibility:** Please state in the legends how many times each experiment was repeated independently with similar results. This is needed for all experiments, but is particularly important wherever results from representative experiments (such as micrographs) are shown. If space in the legends is limiting, this information can be included in a section titled “Statistics and Reproducibility” in the methods section. | |
| **Legends requiring revision:** |  |
| **6.** | **Data availability:** This journal strongly supports public availability of data and custom code associated with the paper in a persistent repository where they can be freely and enduringly accessed or as a supplementary data file when no appropriate repository is available. If data and code can only be shared on request, please explain why in your data Availability Statement, and also in the correspondence with your editor. For more information, please refer to <https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data> | |
| Please ensure that datasets deposited in public repositories are now publicly accessible, and that accession codes or DOI are provided in the "Data Availability" section. As long as these datasets are not public, we cannot proceed with the acceptance of your paper. For data that have been obtained from publicly available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section. |  |
| **7.** | **Gels and Blots:** Quantitative comparisons between samples on different gels/blots are discouraged; if this is unavoidable, the figure legend must state that the samples derive from the same experiment and that gels/blots were processed in parallel.  Vertically sliced images that juxtapose lanes that were non-adjacent in the gel must have a clear separation or a black line delineating the boundary between the gels. Loading controls (e.g. GAPDH, actin) must be run on the same blot.  Sample processing controls run on different gels must be identified as such in the figure legends, and distinctly from loading controls.  All blots and gels must be accompanied by the locations of molecular weight/size markers. Blots should be cropped such that at least one marker position is present.  Please also supply uncropped and unprocessed scans of the most important blots in the Source Data file or as a supplementary figure in the Supplementary Information. This should be cited once in the Methods section. For an example of presentation of full scan blots, see the Source Data file of <https://www.nature.com/articles/s41467-020-16984-1#Sec35> and for more information, please refer to <https://www.nature.com/nature-research/editorial-policies/image-integrity> | |
| **Panels requiring revision:**  Molecular weight markers are missing for panels  Full scans are missing for panels |  |
| **8.** | Micrographs: Please ensure that all micrographs include a scale bar and this scale bar is defined on the panels or in the figure legends. | |
| **Panels requiring revision:** |  |
| **9.** | **Flow cytometry data:** Please provide a Supplementary Figure to graphically account for all FACS sequential gating/sorting strategies, or provide gating/sorting strategies in-figure. If the former, please be sure to indicate, in the Supplementary Figure legend, which gating panel(s) correspond to which FACS data panel(s) in the manuscript figures. (For an example, please see <https://www.nature.com/articles/ncomms15067#supplementary-information>).  *[STNPS team: Flag separately to the editor when this figure and the reporting summary module is missing and the manuscript contains flow cytometry data; if you are unsure if the flow cytometry figure is provided, please flag to the editor that this may need checking]* | |
| **10.** | **Additional Notes** |  |