Peer Review File

Manuscript Title: SARS-CoV-2 infection and persistence in the human body and brain at autopsy

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

In the manuscript by Stein et al., the authors present a comprehensive analysis of 44 COVID-19 autopsies from April 2020 to March 2021, including histopathology, SARS-CoV-2 spike RNA ISH, viral IHC/IF, ddPCR, sgRNA qRT-PCR, viral culture, and viral spike gene sequencing. Overall the study was well designed with appropriate use of statistics, and the manuscript was well written with sufficient references. The general autopsy findings largely confirm previously published results from other large general autopsy or organ-specific studies, as well as the detection of viral RNA in multiple extrapulmonary organs. The viral sequencing results suggesting compartmentalization were perhaps the most interesting novel findings of this study, and could be more emphasized.

The major concerns for this manuscript center around the findings reported in the 11 cases that included brain examination. In contrast to the vast majority of the peer-reviewed literature, the authors here report widespread ISH/IHC staining for SARS-CoV-2 RNA/antigen in the brain of multiple subjects, which would be a novel finding. However, it is difficult to assess these conclusions as there are no RNA ISH images of brain in the manuscript, and it was unclear if these were included in Supplementary Data 3 since I was unsuccessful in accessing https://halo.cancer.gov. Illustrating examples of all the cell types listed in Extended Data Table 4 is necessary to support the localization of virus in human tissues (at least in the brain).

The IF images are also difficult to interpret without COVID-19 positive and negative controls, which are crucial when using custom made antibodies. It would be helpful to know why custom antibodies rather than any of the numerous commercially available ones were used for these experiments, and to clarify whether the target was NP1 per line 600 in the methods or spike protein per line 226 in the results. Would also be useful to see IF staining in multiple brain areas besides cerebellum and in multiple patients.

Other comments:
1) Was viral isolation attempted for any brain specimens as a method to confirm intact replicating virus?
2) The hypothesis that extrapulmonary viral presence due to residual blood within tissue is in general supported by these results, but concluding this possibility has been “ruled out” is a bit strong (line 273).
3) While this cohort predates emergence of Delta variant, should comment on possible differences in tissue distribution or duration due to viral variants
4) Should also include SARS-CoV-2 vaccination status for cases after December 2020

Referee #2 (Remarks to the Author):

General Comments: Stein et al provide a comprehensive analysis of SARS-CoV-2 RNAs within different organs from 44 autopsy cases. The authors quantify viral RNA in a range of tissues beyond the respiratory tract and propose wide-spread virus replication at multiple extrapulmonary sites including the brain. Inspection of tissues showed limited evidence of inflammation or cytopathology outside the lungs, leading the authors to conclude that SARS-CoV-2 infection is systemic and viral RNA persists months after clearance of symptoms.

Specific points listed below

Page 7 lines 146-50. Where RNA samples from the different tissues analysed for integrity and VL measurements normalized to a panel of housekeeping genes since the study includes diverse tissue types. The authors list the mean value for the N gene RNA copies/ng RNA in different tissues and yet show no variability for these measurements (SD or interquartile ranges) across different sites both within and between subjects. There is limited statistical analysis on the intra- or inter-patient VL estimates that would be required to support the conclusions in this paper.

No discussion on the limitations of the study and one obvious item is the potential for sampling bias.

Page 7 second paragraph. The authors should define the parameters of 'high' or 'low' VL and cite published literature to support their definitions. Not appropriate to use such descriptive terminology without providing such information. The authors frequently state in the paper 'the persistence of low level viral RNA' (eg line 154), I question if the data provided in this study allows one to infer the persistence of viral RNA that would surely require kinetic measurements of RNA half-life, encourage authors to remove such descriptive statements.

Page 8 lines 160-165. The authors should provide comparative VL data with stat analysis to support their conclusions.

Page 8 lines 166-170. sgRNA is considered a marker of virus replication, although its predictive value is controversial with studies reporting variable ratios of gRNA to sgRNA. Indeed data in the Supplementary Excel spreadsheet shows variable gene expression at different tissue sites that should be analysed, allowing comparison with published studies. The authors have an opportunity to address several important questions for example: Is there an association between N gene inferred VL and sgRNA levels and do they differ across tissues and stage of disease (early, mid or late).

Page 8 lines 171-173. Isolation of infectious virus was attempted in a small subset of subjects (n=9) and the majority were negative. Where quality controls included in these isolation studies to help interpret the many failures. Supplementary data suggests a limited association with the detection of infectious virus and sgRNA levels, however, this data is not analysed or discussed by the authors. How where samples selected for the evaluation of infectious virus levels– current data set and limited description feels somewhat ad-hoc.

Page 9, line 190- 192. This section concludes that there is limited evidence for genetic compartmentalization and suggests 'no need for alterations in receptor utilization to permit extrapulmonary dissemination'. This conclusion seems premature and raises questions as to whether the authors analysed Ace2 and Tmprss2 transcript levels in the different tissues that were sequenced to ascertain whether there is any association of receptor levels with VL? Ideally the authors should extend this analysis to include viral entry receptors along with reported host factors and inflammatory gene activity to all patients and to compare selected extrapulmonary tissues to complement the section on viral tropism.

Page 9 Can the authors extend these studies to probe for sgRNA and Ace2/Tmprss2 transcripts to inform the reader about sites of active replication and role of receptors in defining viral tropism, respectively.

Page 12 Discussion. For reasons stated above the authors should refrain from over-interpreting their data on SARS-CoV-2 replication and RNA persistence. This reviewer commends the authors on the collection of these important clinical samples and would encourage them to fully analyse their data on N/sgRNA and to confirm bulk-population studies with ISH of different SARS-CoV-2 RNAs. The authors could select representative subjects and tissues to probes (RNA or antibodies) host factors such as receptors to inform us about pathways that are essential for SARS-CoV-2 replication at extrapulmonary sites.

Referee #3 (Remarks to the Author):

In "SARS-CoV-2 infection and persistence throughout the human body and brain", Stein et al provide a thorough and comprehensive autopsy series of patients who died with COVID-19 in an attempt to quantify distribution, replication and cell type specificity of the virus across the human body. In comparison with previous autopsy reports, the authors claim to have used a stringent process to optimize retrieval of viral particles/RNA from the organs along with showing 'persistence' demonstrating that a patient can have virus in the body up to 230 days following symptom onset.

Overall, the study is important, completed well and builds on many previous studies about SARS-CoV-2, reaffirming certain ideas that have been presented before but also bringing to light— experimentally— certain SARS-CoV-2 biology that was only hypothesized. Following concerns should be addressed prior to publication.

Major issues

-The current version of the article written accurately, but at the same time gives readers the impression that the findings could translate to most COVID19 cases or even connects to Long-COVID directly. Given that the population studied is mostly people who died with or from COVID in a hospitalized setting, it is misleading to hint at an idea that these findings apply to PASC patients.

Such claims throughout the article need to be toned down. This includes adding the word autopsy in the title.

-It is true that the many of the plasma samples had no virus, but it could be that the virus is hematologically spread through RBCs, immune cells etc and that the plasma isolation no longer has detectible virus. Do the authors have data on peripheral blood cells? This can also reconcile some organs with very low detectible levels but no obvious cellular tropism.

-a key point of this article is that the study used more sensitive techniques to determine the presence of virus, along with replication-competent virus. Currently, the authors include the ddPCR along with sgRNA data in one table, but have the important virus isolation in vitro data in the supplement in a form that is hard to access. These data make this study unique compared to some of the other autopsy studies and should be made into main figure.

Related to above, all the samples that were tested for virus isolation should be made into a graphic/heatmap along with the corresponding PCR/sgRNA data.

More information on how the authors picked the certain organs for viral isolation will be also important. Was sample selection enriched for in a certain way to have the highest chance of getting a replicating virus? Did the authors try any brain samples or ocular samples (we see one choroid sample) that seem to have had higher amounts of virus by qpcr/sgRNA?

Minor points

-In figure 1 and sup fig 2, the black bars for sgRNA are not visible against the blue boxes. A different color scheme or a change in depiction for the sgRNA is needed

-the sequence mapping in different organs provide a valuable insight and interesting hypothesis as the authors discuss in their discussion. Some of these ideas have been shown experimentally and would be worth citing: PMID 34626548, 33969321, 34083450

-The authors did a great job reconciling all of their relevant findings in the discussions with good comparison against current literature. If the authors can expand upon their findings regarding the lack of immunological responses in the different compartments and put forth some hypotheses in a similar manner that would round out the article better.

Referee #4 (Remarks to the Author):

Summary

In this clearly presented article, Stein and colleagues perform 44 whole-body autopsies (11 including brain tissue) on individuals diagnosed with COVID-19 prior to death. They find SARS-CoV-2 viral RNA at varying levels in a wide array of tissue types. They detect evidence of active viral replication in extra-pulmonary tissues in some cases. In addition, they detect spike RNA and spike protein across a variety of tissues and tissue types. The authors perform limited viral sequencing of spike using a high-throughput, single-copy sequencing approach, finding some evidence of anatomical compartmentalization of genetically distinct viruses. The authors conclude that SARS-CoV-2 is widely anatomically distributed, even amongst patients with asymptomatic-to-mild SARS-CoV-2 infections, and claim their "data prove that SARS-CoV-2 causes systemic infection and can persist in the body for months."

Overview of comments

This is an important, comprehensive study that provides a substantial amount of data on extra-pulmonary detection of SARS-CoV-2, and the authors should be congratulated on assembling such a large and multidimensional dataset. While we agree that the data presented indicate that SARS-CoV-2 is capable of replicating at extrapulmonary sites, we are concerned that the conclusions as presented here, particularly in the abstract and introduction, are broader and more definitive than the data can support. Without sufficient acknowledgement of caveats, this study could be interpreted to mean that extrapulmonary virus replication is a feature of virtually all SARS-CoV-2 infections. While that possibility cannot be ruled out, neither can these data fully support it. The key caveats as we see them are explained in greater detail below.

Major points

1. The vast majority of patients described appear to have severe COVID-19, requiring intubation, ECMO, etc. The authors state "we show SARS-CoV-2 disseminates across the human body and brain early in infection at high levels, and provide evidence of virus replication at multiple extrapulmonary sites during the first week following symptom onset" - this should be caveated so it is clear this is in the setting of severe disease, systemic inflammation, high-intensity medical intervention, and in the context of high levels of pulmonary and plasma viral RNA. These settings cannot be directly extrapolated to healthy individuals with milder COVID-19 disease. The patients here also had a high burden of comorbidities. The authors do note (line 129) that 95.5% had at least one comorbidity, but it seems important to further note in the introductory text that over 65% had 3 or more.

2. The few deaths with apparently mild/asymptomatic COVID-19 infection (e.g. patient 3, patient 26, patient 42) have very low levels of vRNA and no sgRNA. The authors also specifically point to patients 28 and 36 as examples of mild or asymptomatic COVID-19 in spite of widespread SARS-CoV-2 dissemination. They state patient 28 had a mild respiratory infection, but also list diffuse alveolar disease as an immediate cause of death, so it may not be fair to describe this patient's disease as "mild". Patient 36 died from cerebral ischemia secondary to a seizure event, yet does appear to have moderate amounts of vRNA and sgRNA in multiple tissue compartments. At first glance, this may

support the authors' claim that SARS-CoV-2 is widely distributed even amongst patients with asymptomatic COVID-19. However, the patient appears to have experienced a seizure event during the acute phase of infection. Severe cerebral ischemia can result in dysregulated immune functions and systemic inflammation, which permitting viral RNA and even replicating virus to traffic systemically. Thus, most of the cases presented here, including patients 28 and 36, have extenuating factors that complicate our ability to extrapolate findings from these individuals to the broader population. The authors should be careful not to make broad conclusions about the ability of SARS-CoV-2 to disseminate widely in the context of mild or asymptomatic infections based on these two cases. We would also advocate for a more thorough description of these cases in the main text so readers can easily grasp the circumstances surrounding these deaths.

3. The authors should clarify the rationale for, and results of, virus isolation experiments and the use of subgenomic (sg)RNA. The authors extensively use the presence of sgRNA molecules as a proxy for the presence of infectious SARS-CoV-2. While it is true that sgRNA is generated as part of the viral replication cycle inside cells, several studies (e.g., PMIDs 33431879, 33870434, 33247099) have reported evidence that the presence of sgRNA alone is not necessarily a reliable proxy for detection of infectious virus. The authors should cite these studies and comment on how the issues raised affect the confidence with which they can use sgRNA presence as a proxy for viral infectivity. Moreover, on lines 171-173 they summarize virus culture data, but to us this passage does not sufficiently convey the relatively limited scope of virus culture experiments. We appreciate the complexity and limitations of these sorts of studies, and understand that it is not feasible to attempt virus culture from all specimens. But, if we interpret the supplemental data correctly, virus isolation was attempted from 22 specimens across the entire study, and virus replication was only detected (called unequivocally positive) in 7 of these specimens. Detection of replication-competent virus in tissues is an important finding to corroborate sgRNA results, so these results should be quantitatively described in the text and presented to readers in a main display item, not only in supplemental data. In addition, it would be helpful and informative for the authors to explore the relationship between sgRNA levels in tissues and recovery of infectious virus, and to explain the rationale for choosing samples from which to attempt virus isolation. There are several tissues with very high sgRNA burdens for which isolation was not attempted. It also seems that no single tissue type was tested in more than one patient.

4. Detection of spike protein/RNA by IHC/RNAScope provides important evidence for bona fide virus replication in tissues. Is it possible to add some quantitative analyses of these data? It would be useful to understand how intracellular spike RNA/protein quantities detected by these methods compare with SARS-CoV-2 vRNA quantities detected by PCR in tissues and plasma.

5. As the authors note, others have suggested that detection of SARS-CoV-2 in tissues could be due to perfusion of tissues with viremic blood and not necessarily indicative of local virus replication. While IHC/RNAScope data provide evidence of local virus replication, it is notable that in the extended data file, patients with detectable sgRNA in a large number of tissues also tend to be patients with detectable vRNA in plasma and/or very high pulmonary viral loads. Can the authors comment on this and perhaps quantify the relationship among viral loads in respiratory tissues and plasma, and the extent of genomic/sgRNA detection at extrapulmonary sites?

Minor points

Line 231: It would be helpful to specify which of the 5 patients were not considered to have died of COVID-19.

Please make sure all acronyms are defined. Eg. "NeuN" on line 478.

Line 553: Please define which viral gene region is targeted by the sgRNA assay.

-It is unlikely given the timeframe in which these autopsies were conducted, but were any of these individuals vaccinated against COVID-19? If so, please provide these data.

-The addition of deep sequencing data in a few patients is quite interesting. It would be informative to include information in Extended Data Figure 3 about the viral load of each sample, in order to give a sense of the depth to which sequences were sampled.

Author Rebuttals to Initial Comments:

Referees' comments:

Referee #1 (Remarks to the Author):

In the manuscript by Stein et al., the authors present a comprehensive analysis of 44 COVID-19 autopsies from April 2020 to March 2021, including histopathology, SARS-CoV-2 spike RNA ISH, viral IHC/IF, ddPCR, sgRNA qRT-PCR, viral culture, and viral spike gene sequencing. Overall the study was well designed with appropriate use of statistics, and the manuscript was well written with sufficient references. The general autopsy findings largely confirm previously published results from other large general autopsy or organ-specific studies, as well as the detection of viral RNA in multiple extrapulmonary organs. The viral sequencing results suggesting compartmentalization were perhaps the most interesting novel findings of this study, and could be more emphasized.

The major concerns for this manuscript center around the findings reported in the 11 cases that included brain examination. In contrast to the vast majority of the peer-reviewed literature, the authors here report widespread ISH/IHC staining for SARS-CoV-2 RNA/antigen in the brain of multiple subjects, which would be a novel finding. However, it is difficult to assess these conclusions as there are no RNA ISH images of brain in the manuscript, and it was unclear if these were included in Supplementary Data 3 since I was unsuccessful in accessing https://halo.cancer.gov. Illustrating examples of all the cell types listed in Extended Data Table 4 is necessary to support the localization of virus in human tissues (at least in the brain).

We thank the reviewer for these comments and have now organized an extensive multi-panel PDF of SARS-CoV-2 RNA ISH-positive images, including images from multiple brain regions, in Supplementary Data 3. This PDF annotates all SARS-CoV-2 RNA ISH-positive cell types and tissues listed in Extended Data Table 3 (previously 4). We added a column to Extended Data Table 3 indicating the panel in Supplementary Data 3 displaying each SARS-CoV-2 RNA ISH-positive cell type and tissue. We also added SARS-CoV-2 RNA ISH-positive images of multiple brain regions from several cases to Figure 2, so that these data are available in a primary display element.

The IF images are also difficult to interpret without COVID-19 positive and negative controls, which are crucial when using custom made antibodies. It would be helpful to know why custom antibodies rather than any of the numerous commercially available ones were used for these experiments, and to clarify whether the target was NP1 per line 600 in the methods or spike protein per line 226 in the results. Would also be useful to see IF staining in multiple brain areas besides cerebellum and in multiple patients.

Per the Reviewer’s request we have included fluorescent IHC images of SARS-CoV-2 infected and non-infected Vero cells as positive and negative controls for our custom antibody in Supplementary Data 3. We created this custom antibody soon after SARS-CoV-2 peptide sequences were published and prior to commercial availability of alternative antibodies. Our custom antibody has been previously validated in non-human primates and extensively

published, including in three *Nature* manuscripts (PMIDs: 32516797, 32396922, 33090972, 32731258, 34678071, 33431511, 34963391). We have added citations for the first three manuscripts to Methods (Line 683). The target of our custom Ab is SARS-CoV-2 nucleocapsid protein (NP1), which is now corrected in Results (Line 234). Fig. 3 has been expanded to include additional fluorescent IHC images from multiple brain areas and patients including P42 cerebellum, P38 hypothalamus, P40 frontal lobe, P40 basal ganglia, P43 temporal lobe, P43 corpus callosum, and P42 cervical spinal cord.

Other comments:
1) Was viral isolation attempted for any brain specimens as a method to confirm intact replicating virus?

We attempted viral isolation on thalamus, hypothalamus, midbrain, and cervical spinal cord of P38 and dura mater of P36, but no cytopathic effect was observed. Extended Data Fig. 3 now summarizes all tissues where virus isolation was attempted and the result.

2) The hypothesis that extrapulmonary viral presence due to residual blood within tissue is in general supported by these results, but concluding this possibility has been "ruled out" is a bit strong (line 273).

We have softened this sentence to say "unlikely" (Lines 285-288).

3) While this cohort predates emergence of Delta variant, should comment on possible differences in tissue distribution or duration due to viral variants

42 of the 44 patients in our cohort had COVID-19 symptom onset prior to January 1, 2021, with the remaining two reporting symptoms in mid-January and early February 2021. Variants of concern were either not reported to be circulating in the US at that time or were reported to account for only a small proportion of circulating strains (PMIDs: 34111060, 33595644). In the discussion section (Lines 375-377) we now acknowledged that our results might not be generalizable to variants of concern.

4) Should also include SARS-CoV-2 vaccination status for cases after December 2020.

We agree with the reviewer and now clearly state all patients in our study were unvaccinated, both in the "Autopsy cohort overview" section (Lines 128-129) and Methods (Lines 581-582).

Referee #2 (Remarks to the Author):

General Comments: Stein et al provide a comprehensive analysis of SARS-CoV-2 RNAs within different organs from 44 autopsy cases. The authors quantify viral RNA in a range of tissues beyond the respiratory tract and propose wide-spread virus replication at multiple extrapulmonary sites including the brain. Inspection of tissues showed limited evidence of inflammation or cytopathology outside the lungs, leading the authors to conclude that SARS-CoV-2 infection is systemic and viral RNA persists months after clearance of symptoms.

Specific points listed below

Page 7 lines 146-50. Where RNA samples from the different tissues analysed for integrity and VL measurements normalized to a panel of housekeeping genes since the study includes diverse tissue types. The authors list the mean value for the N gene RNA copies/ng RNA in different tissues and yet show no variability for these measurements (SD or interquartile ranges) across different sites both within and between subjects. There is limited statistical analysis on the intra- or inter-patient VL estimates that would be required to support the conclusions in this paper.

We confirmed the integrity of the RNA extraction process for each sample tested, to avoid false-negative results for detection of SARS-CoV-2 RNA. To accomplish this, we used the human ribonuclease P/MRP subunit p30 (RP) gene target available in the Bio-Rad SARS-CoV-2 ddPCR kit to confirm successful extraction of RNA from each sample. We have clarified this detail in Methods (Lines 613-615).

We pursued absolute, rather than relative, quantification of SARS-CoV-2 N gene copies in our tissues, which is an intrinsic strength of the Bio-Rad ddPCR platform. In this platform, a single sample is partitioned into up to 20,000 droplets, each of which undergoes an independent PCR reaction. Poisson distribution analysis of the fraction of positive droplets provides an accurate and reproducible determination of gene target copy number/well. To validly compare results across tissues and patients, we normalized SARS-CoV-2 N gene copy number in each sample to the total RNA inputted into that sample. Total RNA concentration for each sample was determined by Nanodrop.

While it might also be enticing to evaluate our data by normalizing SARS-CoV-2 N gene copies to the human RP gene included in the Bio-Rad assay, which as mentioned above is meant to assure integrity of the RNA extraction process, this gene is known to be differentially expressed in cell types across the human body (https://www.proteinatlas.org/ENSG00000148688-RPP30). Differential expression of human “housekeeping” gene(s) based upon different cell types or physiologic conditions, is additional rationale for why we elected to pursue absolute, rather than relative, quantification of SARS-CoV-2 N gene copies in our study.

To better communicate the average SARS-CoV-2 N gene copies/ng RNA across tissues and variability for these measurements, we revised Extended Data Table 2 (previously 3) to include median SARS-CoV-2 N gene copies/ng RNA and interquartile ranges for each of our 21 tissue groups among early, mid, and late cases. These tissue groups were, in part, re-organized from our prior submission to allow for more anatomically appropriate comparisons between tissue groups.

We value the Reviewer’s request for statistical analyses to support our conclusions that SARS-CoV-2 RNA levels differ between respiratory and non-respiratory tissues and between early, mid, and late cases. Linear mixed model regression analysis was performed and showed statistically significant differences in SARS-CoV-2 RNA levels in respiratory compared with non-respiratory tissues in early, mid, and late cases (Extended Data Fig. 2a). Separate linear regression analysis of SARS-CoV-2 RNA levels, using illness duration as a continuous variable, showed significantly different slopes between respiratory and non-respiratory tissues (Extended

Data Fig. 2b-c). Y-intercepts and slopes for each of our 21 tissue groups were evaluated and are now reported.

No discussion on the limitations of the study and one obvious item is the potential for sampling bias.

We thank the reviewer for this comment and have added study limitations to the Discussion (Lines 372-378). Here, we acknowledge that our findings might not be applicable to younger, healthier, vaccinated individuals or to variants of concern given that our study sample largely included older, sicker, unvaccinated individuals prior to the widespread circulation of variants of concern.

Page 7 second paragraph. The authors should define the parameters of 'high' or 'low' VL and cite published literature to support their definitions. Not appropriate to use such descriptive terminology without providing such information. The authors frequently state in the paper 'the persistence of low level viral RNA' (eg line 154), I question if the data provided in this study allows one to infer the persistence of viral RNA that would surely require kinetic measurements of RNA half-life, encourage authors to remove such descriptive statements.

To our knowledge, we are the first group to systematically quantify SARS-CoV-2 RNA N gene levels across the human body and brain. Absent reliable historical comparators, we have removed the use of terms "high" and "low" related to SARS-CoV-2 RNA levels in tissues.

We respectfully disagree with the reviewer that we "infer" persistence of viral RNA. For each case, we identify the date of COVID-19 symptom onset and report the interval from symptom onset to death. We directly quantify SARS-CoV-2 RNA N gene levels in tissues and propose a conservative definition of persistence as detection of SARS-CoV-2 RNA in tissues >30 days.

Page 8 lines 160-165. The authors should provide comparative VL data with stat analysis to support their conclusions.

To better communicate the average SARS-CoV-2 N gene copies/ng RNA across tissues and variability for these measurements, we revised Extended Data Table 2a (previously 3) to include median SARS-CoV-2 N gene copies/ng RNA and interquartile ranges for each of our 21 tissue groups among early, mid, and late cases. In Extended Data Table 2b we summarize the percentage of ddPCR+ and sgRNA+ cases for each tissue category among early, mid, and late cases.

To support our conclusions that SARS-CoV-2 RNA levels differ between respiratory and non-respiratory tissues and between early, mid, and late cases we performed linear mixed model regression analyses with significant findings now summarized in Lines 149-156 and displayed in Extended Data Fig. 2a.

Page 8 lines 166-170. sgRNA is considered a marker of virus replication, although its predictive value is controversial with studies reporting variable ratios of gRNA to sgRNA. Indeed data in the Supplementary Excel spreadsheet shows variable gene expression at different tissue sites that should be analysed, allowing comparison with published studies. The authors have an

opportunity to address several important questions for example: Is there an association between N gene inferred VL and sgRNA levels and do they differ across tissues and stage of disease (early, mid or late).

We appreciate the reviewer's suggestion that we pursue further analyses to improve understanding of the association between our ddPCR and sgRNA assay results in respiratory and non-respiratory tissues among early, mid, and late cases. To accomplish this, we performed spearman correlations between these 2 assays with results now summarized in Lines 171-177 of the text and displayed in Extended Data Fig 2d-e. Briefly, we observed a strong correlation between ddPCR and sgRNA assays across all tissues tested (N=1025, ρ=0.76; 95% CI: 0.73-0.78) with the strongest correlation observed in early cases (N=496, ρ=0.88; 95% CI: 0.85-0.89) and in respiratory tissues (N=369, ρ=0.86; 95% CI: 0.84-0.89). Among the subset of tissues that were positive by both assays, an even stronger correlation was observed (N=302, ρ=0.91; 95% CI: 0.88-0.93).

Through regression analyses we found that ddPCR has excellent predictive power for positive sgRNA with a receiver operating characteristic (ROC) area under the curve (AUC) of 0.965 (95% CI: 0.953-0.977, Extended Data Fig. 2f). When we weighed sensitivity and specificity as equally important, a ddPCR value of $\geq$ 1.47 N copies/ng RNA predicted a positive sgRNA result with 93.0% sensitivity and 91.6% specificity. These results are summarized in line 178-182 of the text and displayed in Extended data Fig 2f.

Page 8 lines 171-173. Isolation of infectious virus was attempted in a small subset of subjects (n=9) and the majority were negative. Where quality controls included in these isolation studies to help interpret the many failures. Supplementary data suggests a limited association with the detection of infectious virus and sgRNA levels, however, this data is not analysed or discussed by the authors. How where samples selected for the evaluation of infectious virus levels– current data set and limited description feels somewhat ad-hoc.

While SARS-CoV-2 growth in cell culture is specific for detection of replication-competent virus, the assay is not sensitive for this purpose (PMID 34346713). Consequently, it is not surprising that we were unable to culture virus from many tissues with higher versus lower sgRNA qPCR values. As a quality control measure to assure that cytopathic effect (CPE) observed in cell culture was specific to SARS-CoV-2 replication, we performed SARS-CoV-2 sgRNA qPCR on the tissue homogenate and cell culture supernatant from all wells. In all cases that CPE was observed, a decrease in Cq value between the tissue homogenate and the cell culture supernatant confirmed SARS-CoV-2 replication. These data are presented in Extended Data Fig. 3.

In our original submission our aim was to isolate virus from selected non-respiratory tissues simply to validate that replication-competent virus could be detected outside of the respiratory tract. Based upon the reviewers' helpful suggestions, we pursued virus isolation on an additional 35 tissues (55 in total) across four sgRNA Cq intervals to explore the relationship between the SARS-CoV-2 sgRNA levels and successful virus isolation. Results are summarized in the text in Lines 185-189 and in Extended Data Fig. 3a. Additionally, we performed logistic regression

analysis to predict CPE based upon SARS-CoV-2 RNA levels by ddPCR and sgRNA assays. These results are summarized in the text Lines 189-194 and in Extended Data Fig. 3b and c.

Page 9, line 190- 192. This section concludes that there is limited evidence for genetic compartmentalization and suggests 'no need for alterations in receptor utilization to permit extrapulmonary dissemination'. This conclusion seems premature and raises questions as to whether the authors analysed Ace2 and Tmprss2 transcript levels in the different tissues that were sequenced to ascertain whether there is any association of receptor levels with VL? Ideally the authors should extend this analysis to include viral entry receptors along with reported host factors and inflammatory gene activity to all patients and to compare selected extrapulmonary tissues to complement the section on viral tropism.

We appreciate the reviewer bringing this line of the text to our attention. Given that we did not evaluate SARS-CoV-2 receptors at the transcriptional or protein level, we re-phrased this sentence to read "Overall, these findings suggest no need for alterations in the spike sequence to permit dissemination of SARS-CoV-2 to non-respiratory tissues…" (Lines 210-213).

Our HT-SGS data showed that while intra-individual virus diversity did occur, changes to the spike sequence were not required for SARS-CoV-2 to infect and replicate within non-respiratory tissues, as evident by multiple patients with either the identical haplotype, haplotypes with synonymous mutations, or the same major haplotype within the lung also dominant across all tissues sequenced. Our HT-SGS approach is targeted to only amplify and sequence specific pathogen targets and does not reveal host targets. While we appreciate the reviewer's proposal to analyze these targets in our cohort, we feel that analysis of viral entry receptors along with reported host factors and inflammatory gene activity is outside of the scope of this manuscript which focuses on comprehensively investigating the pathogen.

Page 9 Can the authors extend these studies to probe for sgRNA and Ace2/Tmprss2 transcripts to inform the reader about sites of active replication and role of receptors in defining viral tropism, respectively.

We thank the reviewer for the suggestion, but as noted above feel that pursing ACE2/TMPRSS2 transcription levels across tissues collected in our autopsy cohort is beyond the scope of this manuscript.

Page 12 Discussion. For reasons stated above the authors should refrain from over-interpreting their data on SARS-CoV-2 replication and RNA persistence. This reviewer commends the authors on the collection of these important clinical samples and would encourage them to fully analyse their data on N/sgRNA and to confirm bulk-population studies with ISH of different SARS-CoV-2 RNAs. The authors could select representative subjects and tissues to probes (RNA or antibodies) host factors such as receptors to inform us about pathways that are essential for SARS-CoV-2 replication at extrapulmonary sites.

We again thank the Reviewer for their suggestions to more fully analyze our data and have now incorporated multiple additional analyses into our revised manuscript, that have significantly strengthened our observations. Briefly, further analyses now include 1) regression analysis of

SARS-CoV-2 N gene copies/ng RNA in respiratory and non-respiratory tissues among early, mid, and late cases and by DOI as a continuous variable (Extended data 2a-c) 2) multiple Spearman correlations between samples tested by ddPCR and sgRNA assays including among early, mid, and late cases and respiratory and non-respiratory tissues (Extended data 2d-e) 3) ROC curve of logistic regression using log_{10}ddPCR in tissues to predict the detection of sgRNA in tissues (Extended data 2f) 4) tissue culture on an additional 35 tissues across four sgRNA Cq value ranges to better define the relationship between sgRNA levels and replication-competent virus (Extended Data Fig. 3a) 5) ROC curves of logistic regression using log_{10}ddPCR and log_{10}sgRNA to predict presence of replication-competent virus (Extended Data Fig. 3b-c) 6) extensive SARS-CoV-2 spike RNA ISH imaging and cell type annotation across the body and brain (Fig 2 and Supplementary Data 3) 7) digital image analysis to quantify ISH-positive cells from the interventricular septal tissue of 16 cases correlated with ddPCR results from the same tissues (Supplementary Data 3.) 8) additional fluorescent IHC imaging of brain targeting SARS-CoV-2 nucleocapsid protein across multiple cases and brain regions. (Figure 3).

Referee #3 (Remarks to the Author):

In “SARS-CoV-2 infection and persistence throughout the human body and brain”, Stein et al provide a thorough and comprehensive autopsy series of patients who died with COVID-19 in an attempt to quantify distribution, replication and cell type specificity of the virus across the human body. In comparison with previous autopsy reports, the authors claim to have used a stringent process to optimize retrieval of viral particles/RNA from the organs along with showing ‘persistence’ demonstrating that a patient can have virus in the body up to 230 days following symptom onset.

Overall, the study is important, completed well and builds on many previous studies about SARS-CoV-2, reaffirming certain ideas that have been presented before but also bringing to light—experimentally— certain SARS-CoV-2 biology that was only hypothesized. Following concerns should be addressed prior to publication.

We thank the Reviewer for their thoughtful comments regarding the importance of the findings from our COVID-19 autopsy cohort.

Major issues

-The current version of the article written accurately, but at the same time gives readers the impression that the findings could translate to most COVID19 cases or even connects to Long-COVID directly. Given that the population studied is mostly people who died with or from COVID in a hospitalized setting, it is misleading to hint at an idea that these findings apply to PASC patients. Such claims throughout the article need to be toned down. This includes adding the word autopsy in the title.

We understand and appreciate the Reviewer's comments to be careful about how we apply our findings to all COVID-19 cases and particularly to PASC patients. Accordingly, we have added a limitations paragraph to the Discussion (Lines 372-378) where we now clearly state that our results are derived from an autopsy cohort that was largely composed of older unvaccinated individuals with pre-existing medical conditions who died from severe COVID-19. We also state that our study was not designed to determine the contribution of viral persistence to PACS. Per the reviewer's suggestion we have added the word autopsy to our title.

-It is true that the many of the plasma samples had no virus, but it could be that the virus is hematologically spread through RBCs, immune cells etc and that the plasma isolation no longer has detectible virus. Do the authors have data on peripheral blood cells? This can also reconcile some organs with very low detectible levels but no obvious cellular tropism.

Existing literature suggest that leukocytes are either not infected by SARS-CoV-2 or do not support viral replication (PMID: 35385861). Per the reviewer's suggestion, however, we identified 12 cases for which sufficient PBMCs were available, 6 with SARS-CoV-2 RNA detected in plasma and 6 without, who were between 4-18 days post symptom onset. We detected SARS-CoV-2 RNA in three of 12 samples at levels just above the LOD for our ddPCR assay (Supplementary Data 1e). These results indicate that SARS-CoV-2 trafficking in PBMCs is unlikely to meaningfully contribute to viral RNA detected in tissues. While we are unable to run ddPCR on RBCs, we performed extensive ISH imaging across the body and brain, including many cases with significant pulmonary hemorrhage, and observed no SARS-CoV-2 ISH-positive signal in RBCs.

-a key point of this article is that the study used more sensitive techniques to determine the presence of virus, along with replication-competent virus. Currently, the authors include the ddPCR along with sgRNA data in one table, but have the important virus isolation in vitro data in the supplement in a form that is hard to access. These data make this study unique compared to some of the other autopsy studies and should be made into main figure.

Related to above, all the samples that were tested for virus isolation should be made into a graphic/heatmap along with the corresponding PCR/sgRNA data.

We agree with the Reviewer that the virus isolation data was difficult to access. Based upon feedback from this and other reviewers, we attempted viral isolation on an additional 35 tissues (55 in total) across four sgRNA Cq intervals. The virus isolation data is now summarized in detail in Extended Data Fig. 3 along with ROC curves for logistic regression models of ddPCR and sgRNA qPCR to predict the presence of replication-competent virus. We appreciate the reviewer requesting this important data be made into a main figure, but feel it is best kept separate from the heat map in Fig. 1 as the data presented there only accounts for the 11 patients from whom CNS tissue was collected.

More information on how the authors picked the certain organs for viral isolation will be also important. Was sample selection enriched for in a certain way to have the highest chance of getting a replicating virus? Did the authors try any brain samples or ocular samples (we see one choroid sample) that seem to have had higher amounts of virus by qpcr/sgRNA?

We appreciate the request to clarify our approach to selection of tissues for virus isolation. With the first submission, our primary intention for virus isolation was to provide conclusive evidence for SARS-CoV-2 replication outside of the respiratory system. Therefore, we mainly selected a single sample for each non-respiratory tissue with a sgRNA qPCR Cq <30 as this is a labor- and time-intensive diagnostic tool. Unfortunately, samples from P6, P18, P19, and P27 that met the criteria for viral culture were delayed in air transit between investigators due to inclement weather and thawed, making them unfit for culture. This loss of samples limited our selection of non-respiratory tissues which met our virus isolation criteria for the manuscript.

In response to a comment from reviewer 2, we attempted virus isolation on an additional 35 tissues to further explore the relationship between the level of sgRNA detected and successful virus isolation. Overall, we attempted virus isolation on a total of 55 tissues with the following Cq ranges: 15 to <20 (n=11), 20 to <25 (n=14), 25 to <30 (n=16), 30 to <35 (n=14). These results are summarized in Extended Data Fig. 3. Included in these samples were thalamus, hypothalamus, midbrain, and cervical spinal cord from P38 as well as dura mater from P36, but we did not observe CPE from any. In addition to the choroid/sclera of P32 which had CPE, we performed virus isolation of optic nerve from P37, but did not observe CPE.

Minor points

-In figure 1 and sup fig 2, the black bars for sgRNA are not visible against the blue boxes. A different color scheme or a change in depiction for the sgRNA is needed

We agree with the reviewer that the black bars for sgRNA were difficult to visualize against the darker blue colored boxes, and have modified the fill bars to be white to enhance the contrast.

-the sequence mapping in different organs provide a valuable insight and interesting hypothesis as the authors discuss in their discussion. Some of these ideas have been shown experimentally and would be worth citing: PMID 34626548, 33969321, 34083450

We are appreciative to the reviewer for calling these manuscripts to our attention and have cited them in the Discussion section (Line 328).

-The authors did a great job reconciling all of their relevant findings in the discussions with good comparison against current literature. If the authors can expand upon their findings regarding the lack of immunological responses in the different compartments and put forth some hypotheses in a similar manner that would round out the article better.

We thank the reviewer for these comments and have expanded our discussion on lack of immunological responses (Lines 350-359).

Referee #4 (Remarks to the Author):

Summary

In this clearly presented article, Stein and colleagues perform 44 whole-body autopsies (11 including brain tissue) on individuals diagnosed with COVID-19 prior to death. They find SARS-CoV-2 viral RNA at varying levels in a wide array of tissue types. They detect evidence of active viral replication in extra-pulmonary tissues in some cases. In addition, they detect spike RNA and spike protein across a variety of tissues and tissue types. The authors perform limited viral sequencing of spike using a high-throughput, single-copy sequencing approach, finding some evidence of anatomical compartmentalization of genetically distinct viruses. The authors conclude that SARS-CoV-2 is widely anatomically distributed, even amongst patients with asymptomatic-to-mild SARS-CoV-2 infections, and claim their "data prove that SARS-CoV-2 causes systemic infection and can persist in the body for months."

Overview of comments

This is an important, comprehensive study that provides a substantial amount of data on extra-pulmonary detection of SARS-CoV-2, and the authors should be congratulated on assembling such a large and multidimensional dataset. While we agree that the data presented indicate that SARS-CoV-2 is capable of replicating at extrapulmonary sites, we are concerned that the conclusions as presented here, particularly in the abstract and introduction, are broader and more definitive than the data can support. Without sufficient acknowledgement of caveats, this study could be interpreted to mean that extrapulmonary virus replication is a feature of virtually all SARS-CoV-2 infections. While that possibility cannot be ruled out, neither can these data fully support it. The key caveats as we see them are explained in greater detail below.

We thank the reviewer for their thoughtful overview of our manuscript. Based upon input from this and other reviewers, we have made modifications to the manuscript to make clear that our findings may not be translatable to younger and healthier individuals who are not well represented in our study, or vaccinated individuals who we did not study. In the abstract we now state “We show that SARS-CoV-2 is widely distributed, predominantly among patients who died with severe COVID-19…” (Lines 75-77) and in the discussion we clearly outline study limitations (Lines 372-378).

Major points

1. The vast majority of patients described appear to have severe COVID-19, requiring intubation, ECMO, etc. The authors state "we show SARS-CoV-2 disseminates across the human body and brain early in infection at high levels, and provide evidence of virus replication at multiple extrapulmonary sites during the first week following symptom onset" - this should be caveated so it is clear this is in the setting of severe disease, systemic inflammation, high-intensity medical intervention, and in the context of high levels of pulmonary and plasma viral RNA. These settings cannot be directly extrapolated to healthy individuals with milder COVID-19 disease. The patients here also had a high burden of comorbidities. The authors do note (line 129) that 95.5% had at least one comorbidity, but it seems important to further note in the introductory text that over 65% had 3 or more.

Per the reviewer's suggestion, we added a paragraph on study limitations in the Discussion (Lines 372-378) and the percentage of cases with three or more comorbidities (Line 134).

2. The few deaths with apparently mild/asymptomatic COVID-19 infection (e.g. patient 3, patient 26, patient 42) have very low levels of vRNA and no sgRNA. The authors also specifically point to patients 28 and 36 as examples of mild or asymptomatic COVID-19 in spite of widespread SARS-CoV-2 dissemination. They state patient 28 had a mild respiratory infection, but also list diffuse alveolar disease as an immediate cause of death, so it may not be fair to describe this patient's disease as "mild". Patient 36 died from cerebral ischemia secondary to a seizure event, yet does appear to have moderate amounts of vRNA and sgRNA in multiple tissue compartments. At first glance, this may support the authors' claim that SARS-CoV-2 is widely distributed even amongst patients with asymptomatic COVID-19. However, the patient appears to have experienced a seizure event during the acute phase of infection. Severe cerebral ischemia can result in dysregulated immune functions and systemic inflammation, which permitting viral RNA and even replicating virus to traffic systemically. Thus, most of the cases presented here, including patients 28 and 36, have extenuating factors that complicate our ability to extrapolate findings from these individuals to the broader population. The authors should be careful not to make broad conclusions about the ability of SARS-CoV-2 to disseminate widely in the context of mild or asymptomatic infections based on these two cases. We would also advocate for a more thorough description of these cases in the main text so readers can easily grasp the circumstances surrounding these deaths.

Based on the reviewer's feedback, we have provided additional relevant clinical information for each case in Supplementary Data 2a, including the presenting illness upon final hospitalization prior to death, if an individual died with vs. from COVID-19, and relevant "notes" that help clarify the clinical course of selected cases. We have narrowed our discussion of cases who reported only mild or no respiratory symptoms to P36 and P42 (Lines 337-343) and have added a paragraph on study limitations, including that our cohort largely represents older unvaccinated individuals with pre-existing medical conditions, thus limiting our ability to extrapolate findings to younger, healthier, or vaccinated individuals (Lines 372-374).

3. The authors should clarify the rationale for, and results of, virus isolation experiments and the use of subgenomic (sg)RNA. The authors extensively use the presence of sgRNA molecules as a proxy for the presence of infectious SARS-CoV-2. While it is true that sgRNA is generated as part of the viral replication cycle inside cells, several studies (e.g., PMIDs 33431879, 33870434, 33247099) have reported evidence that the presence of sgRNA alone is not necessarily a reliable proxy for detection of infectious virus. The authors should cite these studies and comment on how the issues raised affect the confidence with which they can use sgRNA presence as a proxy for viral infectivity. Moreover, on lines 171-173 they summarize virus culture data, but to us this passage does not sufficiently convey the relatively limited scope of virus culture experiments. We appreciate the complexity and limitations of these sorts of studies, and understand that it is not feasible to attempt virus culture from all specimens. But, if we interpret the supplemental data correctly, virus isolation was attempted from 22 specimens across the entire study, and virus replication was only detected (called unequivocally positive) in 7 of these specimens. Detection

of replication-competent virus in tissues is an important finding to corroborate sgRNA results, so these results should be quantitatively described in the text and presented to readers in a main display item, not only in supplemental data. In addition, it would be helpful and informative for the authors to explore the relationship between sgRNA levels in tissues and recovery of infectious virus, and to explain the rationale for choosing samples from which to attempt virus isolation. There are several tissues with very high sgRNA burdens for which isolation was not attempted. It also seems that no single tissue type was tested in more than one patient.

We appreciate the request to clarify our approach to virus isolation and sgRNA. On initial manuscript submission, our intention for virus isolation was to provide conclusive evidence for SARS-CoV-2 replication outside of the respiratory tract. We chose to run sgRNA qPCR on all ddPCR positive samples as it is high-throughput and we used this information to identify samples from non-respiratory tissues with a sgRNA qPCR Cq <30 to undergo virus isolation. Unfortunately, non-respiratory tissues from P6, P18, P19, and P27 that met the criteria for viral culture were delayed in air transit between investigators due to inclement weather and thawed, making them unfit for culture.

Based upon the reviewers' helpful feedback, we pursued virus isolation on an additional 35 tissues (55 in total) across four sgRNA Cq intervals to explore the relationship between the SARS-CoV-2 sgRNA levels and successful virus isolation. Results are summarized in the text in Lines 185-189 and in Extended Data Fig. 3a. Additionally, we performed logistic regression analysis to predict recovery of replication-competent virus based upon SARS-CoV-2 RNA levels by ddPCR and sgRNA assays. These results are summarized in the text Lines 189-194 and in Extended Data Fig. 3b-c.

We thank the reviewer for bringing our attention to the sgRNA studies which are now discussed in Lines 361-364.

4. Detection of spike protein/RNA by IHC/RNAScope provides important evidence for bona fide virus replication in tissues. Is it possible to add some quantitative analyses of these data? It would be useful to understand how intracellular spike RNA/protein quantities detected by these methods compare with SARS-CoV-2 vRNA quantities detected by PCR in tissues and plasma.

Based upon the reviewer's suggestion, we developed a digital image analysis tool to quantify ISH-positive cells from the interventricular septal tissue of 16 cases correlated with ddPCR results from adjacent tissue samples. Interventricular septum was selected for this analysis given that myocyte morphology was intact and discernable, allowing for accurate ISH-positive cell quantification. Mean SARS-CoV-2 N copies per ng RNA significantly correlated with the median SARS-CoV-2 spike RNA-positive cells (ρ=0.704, 95% CI: 0.320-0.889, p=0.002). These results are summarized in Lines 225-233 and data are presented in Supplementary Data 3. While ISH and ddPCR results significantly correlated in this analysis, differences across assays should be considered when interpreting results. For example, there are slight differences in the specific locations within a tissue tested for ddPCR and ISH despite sampling adjacent sections, our image analysis tool is unable to quantify the number of SARS-CoV gene copies in an ISH-

positive cell, and the LOD for ddPCR makes it potentially less sensitive than ISH to detect very low levels of viral RNA, as has been previously reported (PMID: 34956207).

5. As the authors note, others have suggested that detection of SARS-CoV-2 in tissues could be due to perfusion of tissues with viremic blood and not necessarily indicative of local virus replication. While IHC/RNAScope data provide evidence of local virus replication, it is notable that in the extended data file, patients with detectable sgRNA in a large number of tissues also tend to be patients with detectable vRNA in plasma and/or very high pulmonary viral loads. Can the authors comment on this and perhaps quantify the relationship among viral loads in respiratory tissues and plasma, and the extent of genomic/sgRNA detection at extrapulmonary sites?

We thank the Reviewer for this question. Only 12 cases in the autopsy cohort had detectable SARS-CoV-2 RNA in their perimortem plasma sample with only 2 (P27, P37) of these having detectable sgRNA. The Cq values for the sgRNA qPCR occurred at levels higher than the majority of their tissues at 32.51 and 35.77, respectively. This RNA was extracted from 140 µL of plasma which is much more than can fit into the 25 mg of tissue used for extractions, so the plasma from these two cases is unlikely to be major contributor to the levels we detected. Further, we were unable to isolate SARS-CoV-2 from the plasma of P27. At the request of Reviewer 3 we additionally ran ddPCR on PBMCs from 6 patients with RNA detected in their plasma and 6 without, and found only negligible levels of RNA in three of the samples (Supplementary Data 1a). We additionally have seen no staining of RBCs by ISH in tissue, even in samples with high levels of RNA detected by ddPCR and sgRNA. We comment in the Discussion that we think viremia occurs very early following respiratory tract infection which seeds other tissues (Lines 332-335).

Minor points

Line 231: It would be helpful to specify which of the 5 patients were not considered to have died of COVID-19.

We thank the Reviewer for this clarification. This is now specified clearly in Supplementary Data 2a. After additional review of the medical records and histopathology findings with a pathologist, we made the determination that a sixth patient met the classification of dying with COVID-19 as opposed to dying from it. We have expanded the clinical information provided for each patient which can be found in Supplementary Data 2a, including a brief clinical summary for each of these patients illustrating why this classification was made is included in the “Notes” column.

Please make sure all acronyms are defined. Eg. "NeuN" on line 478.

We thank the Reviewer for calling this to our attention and have amended Lines 683-684 to read “neuronal nuclear protein (NeuN)”.

Line 553: Please define which viral gene region is targeted by the sgRNA assay.

The Envelope gene region is targeted by the sgRNA qPCR. This has been added to the Methods (Line 626).

-It is unlikely given the timeframe in which these autopsies were conducted, but were any of these individuals vaccinated against COVID-19? If so, please provide these data.

All individuals in our autopsy cohort were unvaccinated. This information has been added to the "Autopsy cohort overview" section (Lines 128-129) and Methods (Lines 581-582).

The addition of deep sequencing data in a few patients is quite interesting. It would be informative to include information in Extended Data Figure 3 about the viral load of each sample, in order to give a sense of the depth to which sequences were sampled.

Extended Data Figure 4 (previously 3) contains the depth of sequence reads for each tissue in parentheses next to the tissue name. ddPCR, sgRNA qPCR, and HT-SGS were each performed on the same RNA sample. The viral loads for those samples are listed within Supplementary Data 1b.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

In this revised manuscript, the authors have addressed many of the concerns raised in the original submission; however, significant issues remain in regards to the interpretation of the CNS findings.

1) Most importantly, the choice of controls for SARS-CoV-2 spike RNA ISH and nucleocapsid IF staining are insufficient to demonstrate the specificity of the signal in their respective samples. Inclusion of hypothalamus from P38 for ISH staining with alternative probes does confirm that the assay itself can produce positive (housekeeping gene) and negative (E. coli) signals, but does nothing to prove that the spike ISH signal is specific to viral RNA, and not due to off target labeling, or the presence of non-ISH related pigments (e.g. lipofuscin, hemosiderin, melanin). The inclusion of pre-pandemic tissues from each of the positive brain regions (or at a minimum cerebellum and other areas illustrated in Figure 2) is an essential negative control. A potential positive control for the spike ISH could be SARS-CoV-2 positive FFPE lung sections with adjacent sections showing positive staining with spike and or nucleocapsid IHC.

2) The controls for the nucleocapsid IF are similarly inappropriate, since cultured Vero cells fail to adequately control for potential confounding factors of FFPE brain tissue, which is notorious for autofluorescence that can be misinterpreted as positive signal. As for the ISH, FFPE brain sections from each positively reported brain region in Figure 3 (particularly cerebellum) should be included from pre-pandemic brain sections to demonstrate the specificity of the findings.

3) The interpretation of cell localization and cell types in Figure 2 (particularly CNS regions in a-d) is confusing, since much of the signal appears to overlap the nuclei rather than sit adjacent in a perinuclear Golgi or ER pattern. It's unclear what the basis was for identifying glial cells (i.e. from the Hematoxylin-counterstained images provided, these could represent astrocytes, oligodendroglia, macrophages/microglia, lymphocytes, endothelial cells, or small neurons). Since it seems impractical to co-stain all of these tissues with specific IHC/ISH markers, higher resolution images allowing the readers to assess cell type and subcellular localization of staining would be helpful.

4) The description of multiple cell types in Figure 3 as having glial morphology is not very rigorous, since these could represent NeuN-negative neurons, astrocytes, or microglia, the latter two could easily be confirmed by co-staining with GFAP and CD68, respectively.

5) There is somewhat of a discrepancy between the cerebellar staining illustrated in Figures 2 and 3. In Figure 2c the majority of spike ISH is shown to be in cells within the granule cell layer without any Purkinje cell staining, while Figure 3a-d shows significant staining in Purkinje cells, molecular layer, white matter, as well as some staining in the granule cell layer, further emphasizing the need for appropriate controls to determine whether this staining is real versus artifact.

6) It's curious that cerebellum appears to have the greatest amount of staining compared to other

brain regions in Figures 2 and 3, despite having lower viral loads than many other areas. Could this be due to higher amount of RNA from higher cellularity resulting in lower normalized numbers? Would normalizing to tissue mass used for viral quantification better align with staining?

7) To avoid over-interpretation of these data, it should be specified in the abstract that only 11 cases included brain examination (line 74), and that the conclusion that virus could be detected in brain at 230 days was based on a single case (only other case greater than 100 days did not include brain exam). In addition, it should be better emphasized that while viral RNA was detected in 10/11 brains, sgRNA was not detected past 18 days, and no cultures were positive, and as such the persistence of virus in the brain after 18 days is best characterized as low level RNA without evidence of replication (and with possible antigen if the issues with staining controls are sorted out).

Referee #2 (Remarks to the Author):

The authors provide extensive statistical analysis of their earlier data that strengthens the paper significantly. In addition they provide new figures showing IHC stains of viral antigens in a range of tissue samples and a summary of their attempts to isolate and culture virus from an array of clinical samples.

The authors extend the discussion of their MS to highlight their cohort of unvaccinated elderly subjects and how their results may translate to more diverse populations and the impact of newer VOCs.

Referee #3 (Remarks to the Author):

The revised manuscript by Stein et al put forth a lot of effort in addressing all of the referee's comments to significantly strengthen the evidence for their claims. Working with human samples especially makes it difficult to perform new experiments, so we would like to commend the authors for their great effort in trying to answer many of the critiques with new samples and analysis.

With the new experimental evidence and the rephrasing of their original claims to be toned down/more focused, we believe the manuscript is suitable for publication in Nature.

Minor fixes
The authors refer to Post Acute Sequelae of COVID-19 as PACS (not PASC) in several paragraphs, please fix this typo.

Referee #4 (Remarks to the Author):

Overview

Overall the authors have done extensive additional work in an effort to respond thoughtfully to reviewer critiques, and the paper is strengthened as a result. A few minor issues remain. This work provides an important contribution to our understanding of SARS-CoV-2 infection at extrapulmonary sites.

Major points

None

Minor points

1. The authors should amend some passages that may still give the impression that the study supports a role for disseminated infection in most cases. For example, the last sentence of the abstract should include a statement briefly clarifying that the findings of systemic infection and persistence are most clearly supported for older and/or hospitalized patients with multiple comorbid conditions.

2. Lines 113-114: Here the authors could clarify that virus isolation was attempted for selected tissues and not all tissues.

3. The authors should provide some more context for the viral sequencing results presented in Extended Data Figure 4 and discussed starting on line 196. In order to interpret findings of apparent compartmentalization, it would be very helpful to specify here at what timepoint after symptom onset each patient's virus was sequenced. Similarly, although it is tempting to speculate that nonsynonymous mutations apparently restricted to specific tissues might be under tissue-specific selective pressure, this conclusion cannot be supported with the existing data. It is also possible that tissue-specific differences in viral sequences represent results of stochastic seeding of infection of different sites at different times. In this context it would also be helpful for the authors to clarify what they view as the distinction between "independent" replication vs. "spillover" from one anatomical site to another (lines 321-22).

3. Lines 280ff: This main conclusion statement should be qualified to note that these findings apply to patients with severe SARS-CoV-2 disease and a high burden of comorbidities. It is difficult to extrapolate from these findings that disseminated infection is common in all SARS-CoV-2-infected individuals. The authors note this in lines 337ff, but go on to argue from some unusual cases that disseminated infection might nonetheless be common.

Author Rebuttals to First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

In this revised manuscript, the authors have addressed many of the concerns raised in the original submission; however, significant issues remain in regards to the interpretation of the CNS findings.

1) Most importantly, the choice of controls for SARS-CoV-2 spike RNA ISH and nucleocapsid IF staining are insufficient to demonstrate the specificity of the signal in their respective samples. Inclusion of hypothalamus from P38 for ISH staining with alternative probes does confirm that the assay itself can produce positive (housekeeping gene) and negative (E. coli) signals, but does nothing to prove that the spike ISH signal is specific to viral RNA, and not due to off target labeling, or the presence of non-ISH related pigments (e.g. lipofuscin, hemosiderin, melanin). The inclusion of pre-pandemic tissues from each of the positive brain regions (or at a minimum cerebellum and other areas illustrated in Figure 2) is an essential negative control. A potential positive control for the spike ISH could be SARS-CoV-2 positive FFPE lung sections with adjacent sections showing positive staining with spike and or nucleocapsid IHC.

We appreciate the reviewer's emphasis on controls to ensure the specificity of the antibodies and probes used for immunofluorescence (IF) and *in situ* hybridization (ISH) for detection of SARS-CoV-2, and have undertaken a series of studies, based on the reviewer's recommendations as well as additional virus isolation to provide evidence of SARS-CoV-2 infection in the CNS.

Given that the gold standard for detection of virus in tissue is cell-culture, we made an additional attempt to isolate SARS-CoV-2 from CNS tissue. We selected P38 thalamus and hypothalamus for this repeat attempt, given their relatively high SARS-CoV-2 copy numbers by our ddPCR and sgRNA assays, as well as the availability of remaining frozen specimens. To potentially enhance our chances of culturing SARS-CoV-2 from these tissues, which were previously negative by Vero cell culture, we used a modified Vero cell line that expresses the human ACE2 and TMPRSS2 receptors (Line 639-643). Using this modified Vero cell line, **we successfully isolated SARS-CoV-2 from P38 thalamus** (Lines 198-204) As reported in Extended Data Figure 4f of our prior submission, SARS-CoV-2 spike (S) genome haplotypes (major and minor) detected in P38 RNA-later preserved thalamus tissue were distinct from haplotypes detected in P38 RNA-later preserved lung tissues. To confirm that the SARS-CoV-2 S gene sequence in our isolated virus matched the sequence observed in P38 RNA-later preserved thalamus tissue, another laboratory performed short read, whole genome sequencing (WGS) of SARS-CoV-2 isolated from the supernatant of our modified Vero cell culture of P38 thalamus following observation of cytopathic effect at 48 hours. Indeed, the WGS consensus sequence of the SARS-CoV-2 S genome isolated from the supernatant matched the previously sequenced minor haplotype (Extended Data Figure 4f), confirming that we successfully isolated replication-

competent SARS-CoV-2 from human brain. To our knowledge, we are the first group to provide definitive proof of replication-competent SARS-CoV-2 infection in human brain.

To address the reviewer's concerns regarding the specificity of our ISH assay, we attempted ISH on a series of pre-pandemic samples with appropriate pre-analytic conditions, specifically tissue fixation. The fixation protocol used for our COVID-19 tissues included immediate dissection at the time of autopsy, placement into tissue cassettes, fixation for 24 hours in neutral buffered formalin, and then transfer to 70% ethanol for 48 hours prior to impregnation with paraffin (Lines 573-576). Prior studies have shown that this approach enhances RNA and protein preservation compared with longer fixation times in formaldehyde that mediates damage [PMID 29125916 and 18711211]. Our COVID-19 tissue fixation protocol limited the pre-pandemic CNS samples available to us that were processed in the same way for use as adequate controls. However, we did identify a limited number of archival CNS samples collected in the early 2000's as well as a limited number of CNS samples from a rapid autopsy protocol [PMID 30840888], collected in the last 5 years that were processed in the same manner as our COVID-19 samples. We evaluated each of these tissues by ISH as potential control samples. Our positive control human house-keeping gene probe revealed no signal in the CNS samples collected in early 2000's, consistent with RNA degradation due to prolonged storage [PMID: 19061293, 21411807]. Consequently, these samples were deemed to be inadequate controls for our ISH assay. Among the more recently collected pre-pandemic rapid autopsy protocol samples, we identified 4 tissues including basal ganglia, cortex, CNS not otherwise specified (NOS), and spinal cord with adequate RNA quality for use as pre-pandemic ISH controls. Adequate pre-pandemic cerebellar tissue was not available for use as an ISH control. In each of the four pre-pandemic rapid autopsy tissues our ISH positive control human house-keeping gene probe was positive, our negative control *E. coli* probe was negative, and previously published probes against SARS-CoV-2 S and SARS-CoV-2 nucleocapsid (N) (PMID: 32761270) were negative (Supplementary Data 3, panels yy and bbb). Of note background pigment was detected in the pre-pandemic spinal cord sample, but was distinct from our ISH-positive signal. Per the reviewer's suggestion to use COVID-19 positive FFPE lung tissue for a positive control, we now display pre-pandemic and COVID-19-positive FFPE lung sections adjacent to each other (Supplementary Data 3 panel yy). P18 lung is SARS-CoV-2 positive by ISH using our S and N probes, while the pre-pandemic lung is negative by these assays. Importantly, in the pre-pandemic and P18 lung samples our ISH positive control human house-keeping gene probe was positive, our ISH negative control *E. coli* probe was negative.

To address the reviewer's concerns regarding the specificity of our SARS-CoV-2 N protein IF assay we performed IF on the same pre-pandemic samples used as ISH controls and on a pre-pandemic cerebellum sample deemed appropriate for IHC controls but inadequate for ISH controls due to poor RNA quality. Consistent with the reviewer's prediction, we observed non-specific background autofluorescence on these pre-pandemic control samples by IF. Consequently, we modified our IF assay by adding additional blocking steps (Lines 704-711) to eliminate background autofluorescence. Using this modified IF protocol, our pre-pandemic CNS samples stained negative for SARS-CoV-2 N protein and background autofluorescence. However, when we applied this modified IF protocol on the same COVID-19 CNS tissues displayed in Figure 3 from our prior submission, we saw reduced staining in our previous

samples, with positive SARS-CoV-2 N protein signal only observed in P38 hypothalamus and P42 spinal cord (Fig. 3a, e; Supplementary Data 3 panel zz contains pre-pandemic spinal cord control and basal ganglia as alternative control for hypothalamus, which was not available). In an effort to improve sensitivity and potentially increase the number of positively staining tissues, we attempted IHC using our same anti-SARS-CoV-2 N antibody (Ab) but with a chromogenic DAB detection. The chromogenic IHC protocol is similar to the IF protocol, utilizing the same dried milk blocking, but without the need for the autofluorescence blocking steps on our pre-pandemic CNS tissues (Lines 724-726). Using our chromogenic IHC protocol, we interrogated a pre-pandemic tissue microarray (TMA) that we previously constructed for validation of IHC markers associated with Alzheimer's disease. This platform includes 38 CNS cores encompassing 10 CNS sites, including cerebellum. All the material in the TMA was greater than 9 years old and obtained from routinely fixed autopsy samples, (after 2 weeks of immersion fixation), and was thus not suitable for ISH controls. When we applied our chromogenic IHC assay using our custom anti-SARS-CoV-2 N Ab, almost all of the TMA stained negative for SARS-CoV-2. Three cores (mammillary bodies) demonstrated background reactivity on both the anti-SARS-CoV-2 N Ab and our Rabbit IgG isotype control and could not be interpreted. Additionally, Purkinje cells within the cerebellum displayed the same pattern of background reactivity on the anti-SARS-CoV-2 N Ab and our Rabbit IgG isotype control and could not be interpreted. When we applied our chromogenic IHC protocol to some of the COVID-19 patient CNS tissues displayed in Figure 3 from our prior submission, we observed positive SARS-CoV-2 signal in P38 hypothalamus, P42 spinal cord, P38 cerebellum (with predominant staining in the granular cell layer), and P40 basal ganglia (Figure 3d, j, m, p, respectively). These same tissues were also positive by our ISH assays using SARS-CoV-2 S (Fig. 3b, h, k, and n, respectively) and N probes (Fig. 3c, i, l, o, respectively). For these same tissues, the ISH positive control human house-keeping gene probe was positive, the ISH negative control *E. coli* probe was negative, and the Rabbit IgG isotype control for our anti-SARS-CoV-2 N Ab was negative, (Supplementary Data 3, panels yy and aaa). Per the reviewer's suggestion to use COVID-19 positive FFPE lung tissue for a positive control, we now display pre-pandemic (negative control) and COVID-19-positive (P18) FFPE lung sections (ddPCR positive orthogonal assay positive control) adjacent to each other (Supplementary Data 3 panel yy). P18 lung is SARS-CoV-2 positive by chromogenic IHC using our anti-SARS-CoV-2 N Ab and a Rabbit IgG isotype control for our anti-SARS-CoV-2 N Ab Ab was negative, supporting the specificity of our findings.

The application of two different ISH probes, as well as Ab-mediated detection by both IF and chromogenic IHC methods, with all application controls and evaluation in both pre-pandemic and ddPCR SARS-CoV-2 positive lung and CNS support both the specificity of our findings as well as reproducibility of our findings across multiple microscopy-based assays. The reviewer's concerns about the possibility of false-positive results from ISH related to lipofuscin, hemosiderin and melanin are acknowledged. The presence of pigment in the CNS is not uncommon, with anatomic location, age, previous health issues, and other factors contributing to their presence. All ISH assays were performed with both negative and positive controls and the interpretation is made within the context of these controls. The IHC assay was performed with a concurrent Rabbit IgG isotype control for our anti-SARS-CoV-2 N Ab.

2) The controls for the nucleocapsid IF are similarly inappropriate, since

cultured Vero cells fail to adequately control for potential confounding factors of FFPE brain tissue, which is notorious for autofluorescence that can be misinterpreted as positive signal. As for the ISH, FFPE brain sections from each positively reported brain region in Figure 3 (particularly cerebellum) should be included from pre-pandemic brain sections to demonstrate the specificity of the findings.

We agree with the points that the reviewer has emphasized related to pre-pandemic controls to support specificity of our CNS ISH and IF findings, and as a result we carried out significant additional experimentation to address these concerns. However, as noted above, we were limited in the availability of pre-pandemic ISH controls with appropriate pre-analytic conditions (i.e., adequate RNA quality) across all CNS regions displayed in Fig. 3 in our prior submission. To accommodate this limitation, as noted above, we added a second ISH probe targeting SARS-CoV-2 N, modified our IF assay to prioritize specificity over sensitivity, and implemented a chromogenic IHC assay utilizing our custom anti-SARS-CoV-2 N Ab, with concurrent Rabbit IgG isotype control, which we tested on a broad range of pre-pandemic CNS regions, including cerebellum, to support the specificity of this assay. The RNAscope ISH demonstrates high specificity, and has been widely used for the detection of SARS-CoV-2 in numerous peer reviewed publications (PMIDs: 32761270, 32991819, 34307198). We aimed to enhance the veracity of the ISH signal for SARS-CoV-2 in CNS tissues by adding a second distinct SARS-CoV-2 probe. When we applied our two SARS-CoV-2 ISH probes to the four pre-pandemic tissues deemed to have adequate RNA quality, we observed no signal for SARS-CoV-2, strongly supporting the specificity of these assays across CNS tissues (Supplementary Data 3, panels yy and bbb).

3) The interpretation of cell localization and cell types in Figure 2 (particularly CNS regions in a-d) is confusing, since much of the signal appears to overlap the nuclei rather than sit adjacent in a perinuclear Golgi or ER pattern. It's unclear what the basis was for identifying glial cells (i.e. from the Hematoxylin-counterstained images provided, these could represent astrocytes, oligodendroglia, macrophages/microglia, lymphocytes, endothelial cells, or small neurons). Since it seems impractical to co-stain all of these tissues with specific IHC/ISH markers, higher resolution images allowing the readers to assess cell type and subcellular localization of staining would be helpful.

The reviewer has requested higher resolution images of the CNS images displayed in Figure 2, panels a-d of our prior submission, which we appreciate. We have made every effort to increase the spatial resolution of these images. However, we wish to note that the ISH protocol includes protease treatment to allow probe access to the viral and cellular RNA. In the case of SARS-CoV-2, the presence of a nucleocapsid protein, which packages and protects the viral RNA, is present. We carefully developed and optimized the ISH protocol for sensitivity, but acknowledge this results in less than optimal histomorphology. The protocol in the manuscript accurately reflects in detail the modifications, notably including the longer protease digestion time required.

We were able to accomplish this refinement, based on the ddPCR data, and demonstrated in the previous version of the manuscript that this protocol provides a ρ=0.704 correlation with ddPCR. As such, our ISH assay is unique, compared to others in the literature. Given the limitation of less-than-optimal histomorphology of our ISH images, we now refrain from calling out cell type(s) on most of these CNS samples in the revised manuscript and instead refer to them as CNS NOS. Supplementary Data 3 now contains representative CNS images previously displayed in Figure 2, panels a-d in our prior submission.

4) The description of multiple cell types in Figure 3 as having glial morphology is not very rigorous, since these could represent NeuN-negative neurons, astrocytes, or microglia, the latter two could easily be confirmed by co-staining with GFAP and CD68, respectively.

As noted above, we modified our IF assay to prioritize specificity over sensitivity. Updated Figure 3 now displays positive SARS-CoV-2 signal by IF in P38 hypothalamus and P42 spinal cord in NeuN-positive neurons (Fig. 3a and e, respectively). Co-staining with TMEM119 for microglial cells was attempted, but yielded a negative result (Supplementary Data 3, panel zz).

5) There is somewhat of a discrepancy between the cerebellar staining illustrated in Figures 2 and 3. In Figure 2c the majority of spike ISH is shown to be in cells within the granule cell layer without any Purkinje cell staining, while Figure 3a-d shows significant staining in Purkinje cells, molecular layer, white matter, as well as some staining in the granule cell layer, further emphasizing the need for appropriate controls to determine whether this staining is real versus artifact.

Per the reviewer's helpful suggestion, we evaluated our IF assay in pre-pandemic CNS control tissues, including cerebellum, and observed background autofluorescence. When we modified this assay to eliminate autofluorescence, we no longer observed IF staining in our COVID-19 cerebellar tissues, including those displayed in Fig. 3 from our prior submission. Images of P38 cerebellum show SARS-CoV-2 positivity by ISH using our S and N probes (Fig. 3k and l, respectively) and by chromogenic IHC (Fig. 3m) with predominant positive staining in the granular cell layer observed across these assays. As mentioned previously, we did not have adequate pre-pandemic cerebellum tissue available for an ISH control. However, our ISH positive control human house-keeping gene probe was positive, our ISH negative control *E. coli* probe was negative, and our Rabbit IgG isotype control for our anti-SARS-CoV-2 N Ab, was negative on P38 cerebellum, supporting the specificity of our findings (Supplementary Data 3, panel aaa). The pattern of predominant SARS-CoV-2 positive staining within the cerebellar granular cell layer was also observed in P42 cerebellum by ISH using our S probe (Supplementary Data 3, panel uu). Our pre-pandemic cerebellum IHC control using our anti-SARS-CoV-2 N Ab, was negative except for Purkinje cells, which demonstrated pigment with both our Rabbit IgG isotype control and anti-SARS-CoV-2 N Ab, and so are interpreted as indeterminant. Consultation with an experienced neuropathology researcher confirmed this as a recognized challenge with this cell type, hence Purkinje cells have been avoided in all photomicrographs of the cerebellum and we no longer call these cells out as SARS-CoV-2

positive.

6) It's curious that cerebellum appears to have the greatest amount of staining compared to other brain regions in Figures 2 and 3, despite having lower viral loads than many other areas. Could this be due to higher amount of RNA from higher cellularity resulting in lower normalized numbers? Would normalizing to tissue mass used for viral quantification better align with staining?

The reviewer is correct that cerebellum is highly cellular and thus may account for relatively lower normalized SARS-CoV-2 copy numbers detected by ddPCR in cerebellum compared with other tissues. As mentioned in our prior rebuttal, alternative approaches to normalizing SARS-CoV-2 copy number for comparison of SARS-CoV-2 burden across tissues also have limitations. Normalizing SARS-CoV-2 copy number to tissue mass, when the concentration of cells susceptible to infection differs across tissues, also has inherent limitations. We employed our microscopy-based assays primarily to validate our ddPCR detection SARS-CoV-2 in tissues and when possible, to allow us to identify specific cell types infected. Given the inherent differences between our ddPCR and microscopy based-assays, variability in findings across these assays is expected. We cannot exclude other factors including sampling and the heterogeneity of presences of virus as playing some role in these differences. Although the ddPCR: ISH correlation is very strong, the ISH is rarely homogenous in pattern across any organ.

7) To avoid over-interpretation of these data, it should be specified in the abstract that only 11 cases included brain examination (line 74), and that the conclusion that virus could be detected in brain at 230 days was based on a single case (only other case greater than 100 days did not include brain exam). In addition, it should be better emphasized that while viral RNA was detected in 10/11 brains, sgRNA was not detected past 18 days, and no cultures were positive, and as such the persistence of virus in the brain after 18 days is best characterized as low level RNA without evidence of replication (and with possible antigen if the issues with staining controls are sorted out).

Per the reviewer's suggestion, we have now specified in the abstract that we examined brain in 11 of 44 autopsy cases (Lines 72-73) and that detection of viral RNA in brain as late as 230 days was based on a single case (Line 80). We now also emphasize that we successfully isolated SARS-CoV-2 from thalamus of P38 13 days after illness onset using a modified Vero cell line expressing human ACE2 and TMPRSS2 receptors, providing definitive evidence of replication-competent SARS-CoV-2 in human brain.

Referee #2 (Remarks to the Author):

The authors provide extensive statistical analysis of their earlier data that

strengthens the paper significantly. In addition they provide new figures showing IHC stains of viral antigens in a range of tissue samples and a summary of their attempts to isolate and culture virus from an array of clinical samples.

The authors extend the discussion of their MS to highlight their cohort of unvaccinated elderly subjects and how their results may translate to more diverse populations and the impact of newer VOCs.

We thank the reviewer for these supportive comments.

Referee #3 (Remarks to the Author):

The revised manuscript by Stein et al put forth a lot of effort in addressing all of the referee's comments to significantly strengthen the evidence for their claims. Working with human samples especially makes it difficult to perform new experiments, so we would like to commend the authors for their great effort in trying to answer many of the critiques with new samples and analysis.

With the new experimental evidence and the rephrasing of their original claims to be toned down/more focused, we believe the manuscript is suitable for publication in Nature.

We thank the reviewer for these comments.

Minor fixes
The authors refer to Post Acute Sequelae of COVID-19 as PACS (not PASC) in several paragraphs, please fix this typo.

We thank the reviewer for pointing this out. We have made this correction.

Referee #4 (Remarks to the Author):

Overview

Overall the authors have done extensive additional work in an effort to respond thoughtfully to reviewer critiques, and the paper is strengthened as a result. A few minor issues remain. This work provides an important contribution to our understanding of SARS-CoV-2 infection at extrapulmonary sites.

Major points

None

Minor points

1. The authors should amend some passages that may still give the impression that the study supports a role for disseminated infection in most cases. For example, the last sentence of the abstract should include a statement briefly clarifying that the findings of systemic infection and persistence are most clearly supported for older and/or hospitalized patients with multiple comorbid conditions.

We appreciate the reviewer's input. In the abstract we do state that our observations are "predominantly among patients who died with severe COVID-19" (Lines 76-77). Per the reviewer's suggestion we edited the final sentence of the abstract to read "Our data indicate that **within some patients** SARS-CoV-2 **can** cause systemic infection and persist in the body for months" (Lines 82-83).

2. Lines 113-114: Here the authors could clarify that virus isolation was attempted for selected tissues and not all tissues.

As per the reviewer's suggestion we have clarified that virus isolation was attempted for selected and not all tissues (Line 115).

3. The authors should provide some more context for the viral sequencing results presented in Extended Data Figure 4 and discussed starting on line 196. In order to interpret findings of apparent compartmentalization, it would be very helpful to specify here at what timepoint after symptom onset each patient's virus was sequenced.

Similarly, although it is tempting to speculate that nonsynonymous mutations apparently restricted to specific tissues might be under tissue-specific selective pressure, this conclusion cannot be supported with the existing data. It is also possible that tissue-specific differences in viral sequences represent results of stochastic seeding of infection of different sites at different times. In this context it would also be helpful for the authors to clarify what they view as the distinction between "independent" replication vs. "spillover" from one anatomical site to another (lines 321-22).

Per the reviewer's suggestion, we now include the illness day (D) for each of the patients described in this section (Lines 209-210, 224). We also agree with the reviewer that compartmentalization of S sequences between CNS and pulmonary tissues does not prove

differential selective pressure and that stochastic seeding of the CNS is an alternative explanation. We have therefore moderated the statement in the Discussion that deals with this issue, as follows: "Further studies will be needed to understand whether these cases might represent stochastic seeding of the CNS or differential selective pressure on spike by antiviral antibodies within the CNS, as others have suggested" (Lines 349-351). We have removed the language discussing "spillover" for clarity (Line 344).

s3. Lines 280ff: This main conclusion statement should be qualified to note that these findings apply to patients with severe SARS-CoV-2 disease and a high burden of comorbidities. It is difficult to extrapolate from these findings that disseminated infection is common in all SARS-CoV-2-infected individuals. The authors note this in lines 337ff, but go on to argue from some unusual cases that disseminated infection might nonetheless be common.

Per the reviewer's suggestion we have modified the language in discussion to read: "We show SARS-CoV-2 can disseminates across the human body and brain early in infection in some patients…" (Lines 296-297). We have further softened a later conclusion to read: "Additionally, P36 was a juvenile with an underlying neurological condition, but without evidence of multisystem inflammatory syndrome in children (MIS-C), suggesting that children without severe COVID-19 may develop systemic infection with SARS-CoV-2." (Lines 362-364)

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

In this revised manuscript, the authors have done a commendable job identifying appropriate controls to support the detection of viral RNA/antigen in various tissues. The isolation of replicating virus from brain tissue is particularly noteworthy, as this has not been previously reported.

All of my prior comments have been reasonably addressed, and I have no further comments at this time.