# Response to Editor and Reviewer Issues

## Associate Editor

**Issue 1:** Please pay particular attention to Reviewer 1. We are much more interested in aspects of their manuscript related to the clinical clustering of Long COVID symptoms than of predictors of prolonged convalescence.

**Answer 1:** We agree with the opinion of Editor and Reviewer 1. Thorough characteristic of the symptom profiles in long COVID (≥ 28 days with symptoms) and PASC (post-acute sequelae of COVID-19, ≥ 90 days with symptoms in our cohort is urgently needed for optimal clinical management of COVID-19 convalescence and research.

The selection bias towards participants with protracted recovery present in our survey was pointed out as an obstacle to risk modeling by the Reviewers and stressed by us in Discussion. At the same time, the relatively large subgroup of long COVID and PASC individuals let us to define and characterize co-occurrence patterns of persistent symptoms and explore phenotypic diversity of the participants with perturbed recovery. Following the Reviewers’ and Editor suggestions, this is not the focus of the revised manuscript.

In the revised text, results of clustering of long COVID participants in respect to the numbers of persistent MOP (multi-organ phenotype), FAP (fatigue phenotype) and HAP (hyposmia/anosmia phenotype) symptoms were included in the main figure set (**Figures 6 – 8**). The long COVID phenotyping was appended with symptom profiling data (**Figure 6BC**, **Supplementary Figure S11**), comparison of demographic and clinical features (**Figure 7**, **Supplementary** **Table S5**) as well as assessment of physical impairment, quality of life, mental health, stress and self-perceived convalescence in the participant subsets (**Figure 8**, **Table S6**).

Additionally, we followed suggestion of the Reviewers and repeated the kinetic (**Supplementary Figure S1**), symptom frequency (**Figure 3**, **Supplementary Figure S3**), symptom clustering (**Supplementary Figure S7 - S8**). and participant association analysis (**Supplementary Figures S10 - S13**) for the subgroups of the study cohort with a minimal observation time (SARS-CoV-2 test – survey completion) of ≥ 90 days and suffering from PASC.

The results presented in the revised text indicate an approx. 2-times deceleration of the general symptom count resolution in both long COVID-19 and PASC as compared with non-affected individuals as well as analogical set of leading symptoms including smell/taste disorders, fatigue, tiredness, concentration and memory deficits and tachypnea. By PAM clustering, an analogical set of phenotypes was observed in long COVID and PASC including MOP, FAP and HAP. Finally, the DBSCAN clustering of long COVID and PASC individuals returned three distinct subsets differing primarily in the count of smell and taste disorders, i. e. HAP symptoms: HAP-negative, HAP-intermediate and HAP-high. Interestingly, the HAP-intermediate and HAP-high participant subsets demonstrated lower counts of multi-organ and fatigue complaints as well as better rating of physical recovery. Particularly in the South Tyrol cohort

symptom patterns/phenotypes of post-COVID-19 syndrome (Figure …). In addition, this participant sub-cohorts were subjected to DBSCAN clustering in respect to the numbers of multi-organ (MOP), fatigue (FAP) and hyposmia/anosmia phenotype (HAP) symptoms present in post-COVID-19 syndrome (Figure …), analogically to the long COVID participant clustering. The results indicate presence of isolated smell and taste disorders in approximately one-third of the post-COVID-19 participants not accompanied by other prolonged complaints such as memory and concentration deficits, fatigue, tiredness or pulmonary manifestations. In sum, the results of such association analyses of symptoms and participants indicate clearly, that long COVID and post-COVID-19 syndrome previously defined by the bare presence of any persistent symptoms are heterogeneous conditions with most likely different biological mechanisms, varying impact on quality of life and which supposedly require different treatment modalities. We discuss these findings in light of literature evidence in the discussion section.

In the revised manuscript version, figures coping with risk modeling were reduced to three main readouts (symptoms of acute COVID-19, long COVID and post-COVID-19 symptom risk) and presented in a compacted form in Supplementary Material (Figure … with univariate and … multivariate LASSO modeling). The results are discussed in the revised manuscript in a strongly abbreviated form.

## Reviewer 1

**Issue 1:** The authors used a large survey of adults in bordering regions of Austria and Italy to determine if there were clusters of acute and persistent symptoms consistent between the two regions. Machine learning methods were used to identify clustering in relation to acuity of symptoms. There are interesting findings about the clustering of symptoms, but the analysis and description of these clusters are underdeveloped with room taken by analyses vulnerable to confounding in a cross-sectional study (i.e., identifying factors associated with relapse of persistent symptoms or major physical impairment).

**Answer 1:** We thank Reviewer 1 for thorough review of the text and constructive feedback.

We fully agree that a more detailed analysis of the symptom and participant clusters could improve our manuscript and such data are included in the revised manuscript (see: response to Associate Editor and below for details, Figures ...). At the same time, we are aware of the limitations to the risk modeling based on our cross-sectional survey, which were extensively discussed in the study team and were partially yet imperfectly addressed by weighting and confounder inclusion. Hence, the modeling results were constrained to three main responses (acute COVID-19 symptom number, long COVID and post-COVID-19 symptom risk) and presented in an abbreviated graphical form in Supplementary Material (Figures … with univariate modeling and … with multivariate modeling results).

**Issue 2:** There seems to be a missed opportunity to focus and thoroughly evaluate clusters of persistent symptoms (in Figure S3). There should be an increased emphasis on describing "the three phenotypes of persistent COVID-19 manifestations". Were there demographic differences between these phenotypes?

**Answer 2:** We appreciate this important point.

In the current manuscript version, we present a thorough comparison of demographic, socioeconomic and clinical (e.g., comorbidities, medication intake) characteristic of the long COVID (≥ 28 days with symptoms, Figures ...) and post-COVID-19 syndrome participant subsets (≥ 90 days with symptoms, Figures …). In addition, we compare aspects of recovery between those participant clusters such as self-reported quality of life, physical recovery, frequency of relapse and subjective complete recovery (Figure ...). In summary, we are confident that the three phenotypes of persistent COVID-19-manifestation are now clearly described.

**Issue 3:** There also seems to be a missed opportunity to develop a model for the symptom clusters in one region and validate those clusters in the other. This was somewhat done by running LASSO regression separately for each region and looking at overlap of covariates between the two models but is less robust.

**Answer 3:** This is an important question raised by Reviewer 1 and was also a matter of discussions in the study team. Treating the Tyrol cohort as a training data set and the South Tyrol cohort as a test one is, from the statistical point of view the ‘clearest’ solution and potentially more robust and comprehensive than analysis the overlap between the cluster members of significant correlates done in the previous version of the manuscript. In the revised manuscript we now follow the argumentation of Reviewer 1 and use the Tyrol/TY cohort for cluster and model development and, in case of multi-parameter LASSO, for cross-validation (see: Methods and Supplementary Methods, Figures ...). The South Tyrol/STY cohort is used solely for validation of the clustering and LASSO results. Specifically for clustering analyses, assignment of the symptoms and participants from the STY cohort to the clusters trained in the TY collective is accomplished by k-nearest neighbor classifier (DOI: 10.1007/978-3-319-40367-0\_33, DOI: 10.1109/ICDM.2009.143 and DOI: 10.3923/jse.2014.14.22). The quality of such semi-supervised cluster assignment has been assessed by comparing within-cluster, between-clusters and total sum of squares.

**Issue 4:** The combination of a long list of covariates such as number of symptoms make clinical translation of the results challenging. Clinical relevance of some of the symptoms should be determined prior to being included as a candidate covariate.

**Answer t:** We concur with this opinion. Since we present only abridged modeling results of modeling of acute COVID-19 symptom number, risk of long COVID-19 and post-COVID-19 development, only the 10 strongest favorable and unfavorable correlates validated in the STY cohort are presented (Figure ...)

**Issue 5:** There is discussion about prediction or identification of people at risk of long COVID from their acute symptoms throughout the manuscript. There seems to be evidence of selection bias in that people with persistent symptoms were more likely to respond to the survey further out from their symptoms. There is also great potential for recall bias that was not clearly evaluated. So, given the cross-sectional and retrospective nature of the survey, I would caution against using the results for prediction. However, the clustering techniques of symptoms could be of great use and is a relevant long COVID research gap. This study should focus on describing these symptom clusters more rather than prediction of persistent symptoms.

**Answer 5:** We agree with Reviewer 1 and are aware of the limitations to drawing predictive conclusions from the survey data. As outlined above, we are now focusing on describing detailed characteristics and symptoms in participant clusters during acute disease and recovery (Figures ...).

**Issue 6:** There is a lack of epidemiologic rigor in this study. In the results, the authors mention people >65 years were under-represented but it is not clear what reference they are using. Additionally, those further out from symptom onset to survey were more likely to have symptoms which suggests more of a selection bias. This is described in the discussion but was not fully explored and seems assumed by the authors. If there is census data available, the authors could use that determine if the results are skewed in the survey or use it for weighting the results in the survey. However, I would assume that this is a skewed population due to selection bias until then and focus on the symptom clusters. The symptom clusters should still be discernible with a skewed population given the sample size, but the prevalence estimates may not be generalizable.

**Answer 6:** This is an important point raised by Reviewer 1. We are now presenting measures of demographic skewing of the study cohort compared with the respective regional population of COVID-19 convalescents in Results and as a Supplementary Table …. Univariate modeling results were additionally weighted for age and sex as described in Methods and Supplementary Table As noted by Reviewer 1, age distribution but also sex and education are highly suggestive for selection bias and which is unfortunately discernible in most of similar studies tackling COVID-19 recovery in non-hospitalized/cross-sectional cohorts. We are also fully aware, that the prevalence of long COVID-19, post-COVID-19 syndrome, relapse or physical performance loss hast to be extrapolated on the general convalescent population with the greatest caution and we outline this in the revised discussion section. However, the selection bias resulted in an ‘enrichment’ of the study cohorts with participants with persistent manifestations and enabled us to explore co-occurrence of protracted symptoms and clustering of individuals. We included this point in the strength/limitation section of Discussion.

**Issue 7:** Abstract: Methods: "International, multi-center," - while this seems potentially true it seems a little misleading. I'm not sure multi-center would apply and I would add that they are bordering regions. I would also add at least a sentence about the analytical methods and the time period of data collection. The phrasing in methods is more appropriate "neighboring European regions Tyrol (Austria) and South Tyrol (Italy)". Results – the sentence about the symptom burden is not clear. I would make sure it is clear that this is the number of symptoms present. A dash is used for an IQR but then used to denote the two regions, which is confusing. I would indicate "ST: XX% and STY: XX%", for example, rather than use a dash throughout the paper, as it is not a range.

**Answer 7:** we appreciate these suggestions. We changed the Abstract accordingly:

*Methods: A two cohort, online survey study was conducted between Setptember 2020 and July 2021 in the neighboring European regions Tyrol (Austria, n=1157) and South Tyrol (Italy, n= 893). Data on demographics, comorbidities, COVID-19 symptoms and recovery of SARS-CoV-2 adult outpatients were collected. Phenotypes of acute COVID-19 and post-acute sequelae and risk of protracted recovery were explored by semi-supervised clustering and multi-parameter LASSO modeling.*

Answer 8???: we avoid the imprecise ‘symptom burden’ wording and use ‘symptom count/number’ instead both in Abstract and the main text. Statistical number such as percentages or medians in the text follow now consistently the ‘TY: …, STY …’ pattern to improve clarity.

**Issue 8:** Introduction: It would be helpful to have more background about the online survey.

**Answer 8:** Although we are constrained by the Journal word limit, we provided few details on the survey study population, type of collected data and analysis techniques in Introduction.

**Issue 9:** Methods: The authors should describe the type of media that was used for recruitment.

**Answer 9:** We are grateful for raising this important issue and outline the media type in Methods (local broadcasters, such as ORF Tirol and RAI Südtirol and newspapers)

**Issue 10:** It seems like the survey had not been previously validated which should be listed as a limitation in the discussion. The survey should also be included in the supplemental methods.

**Answer 10:** Even though designed by an expert panel and including parts based on established tools (e. g. mental health assessment as described in a sub-project pre-print DOI: 10.1101/2021.09.22.21263949), the survey has not been validated before. We agree with Reviewer 1 that this may pose a limitation and stress this point in Discussion. The original German and Italian questionnaires as well as the English translation are provided as Supplementary Files accompanying the revised manuscript.

**Issue 11:** The phrase "i. e. the duration of the official quarantine." Seems not directly relevant and I would remove it.

**Answer 11:** We concur with the suggestion. In particular, the quarantine duration varies with country and momentarily testing strategy. We removed the phrase from the revised manuscript.

**Issue 12:** For odds ratios described in results, please provide 95% confidence interval.

**Answer 12:** For the univariate modeling results cited in the text, 95% confidence intervals were provided. However, we support the opinion, that inference statistics for LASSO (but also for related elastic net and ridge algorithms) estimates implementing the ‘shrinkage’ λ factor are not 100% proper (see Dezeure et al. for discussion and solutions to the problem, DOI: 10.1214/15-STS52). Similarly, because of the strongly data-driven character of λ, calculation of LASSO confidence intervals by bootstrap as suggested in the seminal paper by Robert Tibshirani is also a matter of lively discussion (DOI: 10.1111/j.2517-6161.1996.tb02080.x). For these reasons, we provide LASSO estimates (labeled consistently ORLASSO) without confidence intervals. Instead we provide more relevant global error estimates for LASSO models: redistribution (TY cohort), cross-validation (TY) and validation (STY) mean absolute errors (MAE) in the text and/or figures.

**Issue 13:** As above, I would move Figures S4 and S4 out of the supplement to be major figures.

**Answer 13:** we agree with the suggestion and moved the Supplementary Figures with participant cluster characteristic to the main manuscript.

## Reviewer 2

**Issue 1:** Sahanic et al have prepared a manuscript reporting findings from a large online cross-sectional survey that includes laboratory confirmed convalescent COVID patients from two neighboring in Austria/Italy. They used cluster analysis to define two phenotypes of long-COVID- (non-specific infection vs multi-organ) and further identify 3 phenotypes of persistent symptoms (hyposmia/anosmia vs fatigue vs multi-organ). Major strengths of this manuscript are the focus on outpatients, which remains a gap in the long-COVID phenotyping literature. This report is well written and data presented are thorough and clear. This would add to the current literature on long COVID, as its attempt to further identify phenotypes would add to the current literature that mostly describes overall prevalence employing vague case definitions.

**Answer 1:** We thank Reviewer 2 for the careful study of our text and improvement suggestions.

**Issue 2:** The authors note in the introduction that current case definitions typically use a 12-week cutoff to differentiate between ongoing COVID and post-COVID-19 syndrome or PASC; however, this cut off was not employed in the methods. Was a sub analysis attempted to look at this later convalescent period? Specifically, did the authors attempt modeling using a 4th group that included persistent symptoms beyond 12 weeks or was this not performed due to the burden of missing data beyond 12 weeks/right censoring? The median days from symptom onset to survey completion was 96 days in South Tyrol, so may be feasible in this cohort at least. This would be quite informative to report, especially if the cluster analysis still identifies the same sub-phenotypes or if a new late convalescent phenotype emerges.

**Answer 2:** We appreciate this point a lot. At the time the manuscript was drafted, the mainstream literature focus was on the day 28 as the long COVID cutoff and we intended to investigate the recovery beyond the usual quarantine length. We agree with Reviewer 2 that a longer time frame is of emerging interest of epidemiologists, clinicians and authorities. Hence we repeated the symptom frequency characteristic (Figure ...), clustering of the post-COVID-19 syndrome manifestations (Figure …) and clustering analysis of the syndrome-affected participants (Figures … and ...) in both study cohorts. In such analyses, subsets of the initial cohorts with a minimal observation time of ≥ 90 days (TY: n = …, STY: n = ...) were investigated. The results underline the multi-facet character of the long-term recovery; in particular, they confirm that a substantial subset (approx. 30%) of post-COVID-19 syndrome individuals suffers from isolated smell and taste disorders not accompanied by fatigue, cognitive problems or pulmonary symptoms (Figure ...). In addition, we delineate differences in demographic and recovery status variables such as age, sex, comorbidity, self-assessment of quality of life and physical performance between the clusters of long COVID and post-COVID-19 individuals (Figure …). See also Issue 1 in the response to Associate Editor.

**Issue 3:** 21-32% relapse rate was reported, but it unclear how relapse was defined and assessed? Both true relapse after resolution and exacerbation of ongoing symptoms are characteristic of certain PASC phenotypes, specifically, neurocognitive dysfunction and fatigue. Did relapse differ across the three different persistent COVID-19 phenotypes (HAP, FAP, MOP)? A 6-fold increase was reported in those with acute MOP and fatigue, but association with persistent phenotypes is not reported and would be of interest.

**Answer 3:** Remitting/relapsing character of long COVID and post-COVID-19 syndrome is now reported clear in light of literature and knowledge acquired in the growing population of convalescents but were ill characterized at the study initiation in September 2020. We have to admit that this point was only insufficiently covered by the survey; the occurrence of relapse was gauged by a single yes/no question without specifying the recurring symptoms or their intensity (see the text attached as Supplementary File to the revised submission). Furthermore, we believe that characteristic of relapse events may be investigated by longitudinal studies with a far better resolution. A more detailed relapse assessment panel is however planned in the follow-up survey.

In the current manuscript version, the results of relapse risk modeling were removed as a response to bias issues raised by Reviewer 1. Instead, we decided to compare frequencies of relapse queried by the single yes/no item in the clusters of long-COVID and post-COVID-19 syndrome individuals (Figure ...)

**Issue 4:** "Psychosomatic disorders" was referred to when describing the MOP phenotype. Please clarify what this disease category refers to. Many long COVID patients have been stigmatized by the medical community, being told that their symptoms are psychosomatic when physical exam and laboratory assessments have been unremarkable. The authors should be cautious when using this phrase without appropriate context.

**Answer 4:** We agree with Reviewer 2 that the term ‘psychosomatic’ is used in a stigmatizing, pejorative way by parts of the medical community. We removed hence the phrase from the revised text. While the current manuscript focuses on epidemiology and phenotyping of the acute disease and recovery and only touches upon mental health aspects, our current sub-projects tackle the interplay between the social situation, disease symptoms, depression, anxiety and mental health in the ‘Health after COVID-19 in Tyrol’ study populations (DOI: 10.1101/2021.09.22.21263949).

**Issue 5:** Was vaccination status assessed, especially in South Tyrol where 30% of cases were in the winter/spring 2021. While not an emphasis of this report, the role of vaccine in preventing or attenuating Long-COVID is another topic of interest.

**Answer 5:** At the time of study initiation and design (Summer 2020), anti-SARS-CoV-2 vaccines were only at their early development stage. For this reason, no item on vaccination status was included in the survey. Instead, the vaccination interest was queried (Supplementary File ...) and is being analyzed as a part of study sub-project. The assessment of vaccination status and type of received vaccine and number of doses will be included in the follow-up survey.

**Issue 6:** Table 3: what does "Subjective complete convalescence" refer to?

Answer 6: We apologize for this unclarity. The ‘Subjective complete convalescence’ was measured by a single yes/no item. We explain this in Supplementary Methods and the corresponding table legend.