Thank you for the opportunity to revise our work. We appreciate the thoughtful comments of the editorial committee and the reviewers. Please find our detailed response to all of the comments by the editorial committee and the reviewers below.

We are looking forward to your decision.

Sincerely,
Raphael Peter, Dietrich Rothenbacher, and Winfried Kern

Decision: Put points

Detailed comments from the meeting:

Point 1: First, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below.

Response: Please pass on our severe thanks to the reviewers. We are very grateful for their constructive comments and suggestions and think the paper improved considerably with this revision.

Point 2: Please also respond to these additional comments by the committee:

We are very interested in this topic because it has important implications for physicians, patients and society at large when making decisions about personal and societal risk regarding COVID. Nevertheless we recognize that this study has serious limitations related specific response and recall biases that might alter the interpretation of the estimates presented here. We suggest that instead of solely focusing on the proportion of those with COVID and long COVID, the focus shift slightly to concentrate on the types of symptoms present and how they cluster. Our reasoning is that while there are selection and recall biases that can bias the proportion of all COVID patients with long COVID, the response pattern and symptom clusters might be more informative overall. Presenting the estimates as you have them now is fine as long as limitations of prevalence are strongly worded and honest, and as long as the focus is shifted as suggested above. We have more specific comments below, and reviewers also raised many important points that need to be addressed.

Response: We agree with this recommendation and have now slightly shifted the focus of the paper to types of symptoms and symptom clusters. We are also presenting now more clearly the limitations of our study design and the potential and implications of selection and recall bias (see below).

Point 3: Please be direct about the potential for bias to affect the magnitude of the associations is substantial.

1) Selection bias
This is highlighted in the low response rate. People are invited to participate about symptoms and COVID, and it is possible that those with symptoms are more likely to participate than those without. Again, the low response rate means the magnitude of this bias could be large.

**Response:** Overall, we now appreciate the likelihood of selection bias much more in the current version of the paper (see also the response to reviewer 2 point 2 and reviewer 3 point 2). We are presenting now more clearly the limitations of our study design and the potential and implications of selection and recall bias. We, therefore, deleted to “estimate their prevalence” also from the objectives (see page 3, abstract, objectives and page 5 last para). In addition, we deleted it from the summary of the key results at the beginning of the discussion (see page 12). In figure 4 we now more clearly present the minimum possible prevalence assuming that all non-responders fully recovered and were free of symptoms and emphasised this also in the text (details see figure 4 and page 15, para 2). Considering the potential of selection bias and the extreme assumption that all non-responders were free of post-COVID sequelae, the unbiased (“true”) estimates should be in between these estimates and the estimates of the included study population. In addition, to reflect the shift of the focus, we would also suggest modifying the title accordingly to: "Post-acute sequelae of COVID-19 six to 12 months after infection: a population-based retrospective cohort study" (see new title on page 1).

2) Recall bias

Participants are asked about symptoms before, during and after COVID infection, all of which are at varying times in the past. Recall is likely to be affected by previous events and experience and the effects could be large.

**Response:** The potential for recall bias for existing symptoms 6-12 months post-COVID while filling out the questionnaire should be low and may be related to the severity of symptoms. Recall bias for symptoms existing before acute SARS-CoV-2 infection, and secondly, during acute SARS-CoV-2 infection may be of more relevance and may lead to an overestimation of the post-COVID symptom prevalence differences and prevalence ratios. This limitation is mentioned in the discussion section and is especially relevant in subjects with neurocognitive sequelae (details see page 15, para 2).

See also the response to reviewer 2, point 2 and reviewer 3, point 2.

3) Lack of a comparator

Only those with a positive test were invited. a) It is possible that tying symptoms to an event will heighten reporting. There should be a control group recruited around an alternative event. b) It is possible that societal changes at the same time were associated with symptoms and this seems plausible for psychological symptoms such as anxiety and mild fatigue. Again these are potentially large effects.

**Response:** We only have a before-after comparison within infected subjects (which included, however, all infected subjects within a defined geographic region in a specified time period). On the one hand, this limits the interpretation of findings, as acknowledged in our discussion section. On the other hand, it minimises confounding as every subject is its own control (details see page 15, para 2). In an ideal world, we would prefer both types of controls: within-subject comparisons and uninfected, test-negative participants with a similar structural homogeneity. However, given the time pressure of capturing enough participants 6-12 months after the first wave and the impossibility of
recruiting test-negative subjects with similar structural characteristics (to minimise confounding and bias) 6-12 months after their negative test forced us to the presented approach. Unfortunately, in Germany, test-negative controls were (and are) not available within the sampling frame of our cases (as negative tests were not reported by name and address to the public health authorities due to the lack of a legal basis). We added this information to the discussion section of the paper (page 15, para 2). (See also response to reviewer 3, point 2 and reviewer 3, point 7).

Point 4: Please be more specific about missing data. Please highlight how that was evaluated and addressed.

Response: As listed in table 1, there were few missing data for most covariables (the highest number of missing data was observed for cancer as comorbidity with [only] 3.3% missing values). Because the proportions were so low, we did not attempt any imputation for missing data. We now added this information to the methods (page 7, para 2). We now also include the total Ns in the legends of figures 2&3.

Point 5: Can the authors offer a of sense of how representative the sample was.

Response: As outlined in our supplemental table S5, we had a limited response with some overrepresentation of older persons and female sex. Our study regions were located around medium-sized university cities, with respondents having higher education than the general population, which may limit generalizability. This is addressed as a limitation in our discussion section (page 15, para 2). (The reason why we placed the study region around all our four university cities in the State of Baden-Württemberg is that we are currently inviting 1500 subjects of this study for a thorough clinical examination in the nearest of the four southwestern university hospitals to further evaluate and validate the reported sequelae and also to shed light on the pathophysiology of the different post-COVID sequelae.)

Point 6: Cluster analysis is a valid technique, but it is unclear how clusters were chosen. More detail is needed. The linkage approach to clusters and general health and working capacity was also unclear. How was this done?

Response: We used a two-step approach to identify symptom clusters, which we now describe in detail (page 7, last para).

Firstly, strongly correlated current symptoms (not present before the acute SARS-CoV-2 infection) were identified using exploratory polychoric factor analysis based on symptom severity (not present, no impairment, light impairment, moderate impairment, or strong impairment). In technical terms, we used the ‘fa’ function of the R package ‘psych’ to find the minimum residual solution (‘fm=minres’). We applied the oblimin rotation as oblimin rotation usually leads to more interpretable factors compared to other possible rotation algorithms. To identify the ideal number of factors ‘parallel’ analysis was used.

Secondly, we included each symptom into the cluster (identified in the first step as factor) for which its factor loading was highest. However, for the definition of the symptom clusters, we just used the occurrence of symptoms (without any weighting by factor loadings).
Regarding the linkage approach of clusters and general health and working capacity: The association of each current symptom cluster with loss of general health and working capacity compared to pre-infection was estimated using a linear model (adjusted for other symptom clusters). For the attributable loss, we multiplied the estimated cluster prevalence by its’ associated loss. E.g., we estimated a prevalence of 37.2% for the fatigue cluster and an associated mean general health loss of 6.1% (independent of other clusters). Therefore, the population attributable loss is $6.1\% \times 0.372 = 2.27\%$. Corresponding 95% CIs for the attributable loss were estimated using a parametric bootstrap based on the robust variances for the associated loss and the log-prevalence estimated via a log-Poisson model. We now describe this approach in more detail (page 8, para 2).

**Point 7:** Only 3.5% were hospitalized. Does this indicate a group with predominately mild disease.

**Response:** According to data from our State Department of Health Protection, Infection Control and Epidemiology collected in the same time period from our four defined study regions from 51.433 subjects aged 18-65 years whose infection was notified to the local public health authorities, 1.808 were hospitalised. This is a proportion of 3.5% and reflects precisely the number in our study who reported in-patient treatment (see table 1). Therefore we can conclude that with respect to the course of the disease, we had a very representative sample of our source population.

As most chart-reviews or clinical studies mainly overrepresent severe cases with medical treatment needs, we consider this population-based approach a strength of the study. It also underlines that study participants responded when their course of the disease was mild, as we urged them to do so in the information material (see also response to reviewer 2, point 2).

**Point 8:** Can the authors offer any clear time frame for "pre-Covid" period?

**Response:** We evaluated the list of symptoms at three points in time: firstly, symptoms already existing before acute SARS-CoV-2 infection, secondly during acute SARS-CoV-2 infection, and thirdly at the time point the standardised questionnaire was filled out, which resulted in a mean of 8.5 months after the positive test result of the acute infection. The time period before acute SARS-CoV-2 was not further specified. (see also response to Reviewer 2, point 3)

**Point 9:** Can you provide more information about how clusters were developed? It is not quite clear why some symptoms were lumped to form a cluster while others were not.

**Response:** Please see the detailed response to point no. 6.

**Point 10:** We are not sure table S2 shows "determinants" (indicating causal link) of the 13 symptom clusters, as other confounders might exist. Can the authors clarify this?

**Response:** We changed the terminology in the text accordingly to avoid causal thinking and now speak of “characteristics associated with” (see page 10, last para).

**Point 11:** There is no information of variants and vaccination status, and their impact on the outcomes. Can the authors provide this linkage?
Response: As the first variants of concern (VOC) appeared in January 2021 in Germany, our study cohort was infected mainly with the wild type of SARS-CoV-2. Based on national data on the spread of VOC, we estimated that less than 15% of the cohort had been infected with B.1.1.7 (Alpha) and less than 1% with B.1.351 (Beta) (see table below). We now added this information to the text (page 12, last para).

<table>
<thead>
<tr>
<th>PCR test dates of the study cohort</th>
<th>Share of VOC based on national data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alpha (B.1.1.7)</td>
</tr>
<tr>
<td>2020 Oct. to Dec.</td>
<td>65.0%</td>
</tr>
<tr>
<td>2021 Jan.</td>
<td>14.5%</td>
</tr>
<tr>
<td>2021 Feb.</td>
<td>6.3%</td>
</tr>
<tr>
<td>2021 Mar.</td>
<td>14.2%</td>
</tr>
<tr>
<td>Estimated total in the study cohort</td>
<td>100%</td>
</tr>
</tbody>
</table>

The first vaccine doses became available in Germany in late December 2020. However, access was initially restricted to frontline workers and the elderly (and others at especially high risk). Therefore, very few study participants in our cohort already had a chance to receive their first vaccine shot. Only 1.9% of our study participants received their first shot within or prior to the month of infection (we now added these data to table 1).

The association of post-acute vaccination with potential sequelae is currently analysed in more detail in a separate work. However, first data do not suggest a relevant association so far.

Point 12: Please make sure to include a required dissemination statement in the end matter. This statement should describe any plans to promote your research and findings. This can include posts on social media. You can even offer information on the pre-print.

Response: After the embargo, we will distribute the publication and main results via different social media and via the press offices of the contributing universities. We also plan a central press conference with the Federal State Health Office to present data and allow press inquiries officially. In addition, we will present the data at national epidemiological congresses, i.e. at the German Epidemiological Association (DGEpi) congress in Greifswald (Germany) end of September 2022. In addition, we will contact patient advocacy groups and inform them about our results in lay terms. This and further plans are no included in the required dissemination statement (see page 17).

Point 13: In your response please provide, point by point, your replies to the comments made by the reviewers and the editors, explaining how you have dealt with them in the paper.

Response: Please find a detailed point-to-point response to the points of the editorial team and the reviewers.
Comments from Reviewers

Reviewer: 1

Comments:

Point 1: It is an interesting and well designed work that deserves publication, in my opinion, both because it is based on a large number of cases and because it introduces some new elements, such as symptom clusters.

Response: Thank you very much for the thorough and very positive evaluation of this work and the clear recommendation that it deserves publication.

Point 2: It may be useful to ask the authors for two further insights:

1) highlight any treatment performed, provided that it has been included in the questionnaire (see page. 10, third line);

2) given that the study covers the period in which anti-covid vaccination was offered regularly, the percentage of vaccinated patients and any differences between the two groups (vaccinated vs unvaccinated).

Response: Ad 1) In this population-based study conducted in four administrative and geographically defined regions and including all subjects aged 18-65 years who were tested positive in a SARS-CoV-2 PCR between Oct. 1st, 2020 and April 1st, 2021 and had to be reported by law to the local public health authorities, we had only self-reported data regarding treatment needs during acute infection. This information is reported in table 1 (see table 1 rows “Treatment of acute SARS-CoV-2 infections”).

In addition, in figure S3 we report the frequency of each evaluated symptom 6-12 months after acute infection (given symptom was not present before infection), stratified according to sex and medical treatment (see figure S3 bottom panel).

Ad 2) The first vaccine doses became available in Germany in late December 2020. However, access was initially restricted to frontline workers and the elderly (and others at especially high risk). Therefore, very few individuals in our study cohort already had a chance to receive their first shot. Only 1.9% of our study participants received their first shot within or prior to the month of infection (we now added these data to table 1). The association of post-acute vaccination with potential sequelae is currently analysed in more detail in a separate work. However, first data do not suggest a relevant association so far.
Reviewer: 2

Comments:

Point 1: Dr. Peter and colleagues report about a large retrospective cohort study conducted in Southern Germany that assessed symptoms and symptom clusters of Long Covid. The study was carefully designed, conducted and reported. It provides timely evidence on Long Covid when the wild type of SARS-CoV-2 was still prevalent. Particular strengths are the population-based setup, the use of standardized questionnaires, the large sample size (important to have enough power to assess specific risk factors like specific pre-existing conditions), the detailness of symptoms assessment and the follow-up of up to 12 months. The following comments may be considered to further improve the paper and facilitate interpretation of its results.

Response: Thank you very much for the thorough and very positive evaluation of this work. We are very grateful for the constructive comments and think the paper improved considerably with this revision.

Point 2: It was difficult to find the study on the DRKS study registry and it would be useful to have access to the protocol for some details. For example: How were participants informed about the purpose of the study? Was there specific reference to Long Covid or was it frame more generally as a cohort study that aims to address several questions around SARS-CoV-2? This is quite key since self selection into a study may have substantial impact on the results. Generally, studies that explicitly refer to Long Covid (or similar terms) have higher prevalence and also sometimes different associations because there is some self-selection of persons into a study who may be more affected by Long Covid as compared to population based studies that also had different aims and do not focus entirely on Long Covid.

Response: The study participants were informed about the study with the following wording included in a two-sided participant information leaflet (which was positively evaluated by our ethical review committees and informed about the purpose of the study. In Germany, it is required by ethical committees to fully disclose the purpose of the study):

“Dear Sir or Madam,

The University Hospitals of Baden-Württemberg, on behalf of the state government and in cooperation with the State Health Office and local health authorities, are planning a study on the topic of "Long Covid", i.e. illnesses and complaints as a result of a SARS-CoV-2 infection (Corona). [...] In this study, adults who have been reported (to the health authorities) as having a corona infection 6 to 12 months ago will be interviewed. [...] Important: please participate in the current survey even if you never have had or no longer have symptoms. This is the only way we can determine how common these complaints are overall. [...]”

The latter statement is emphasised in many places in the information and promotion material of the study (the study was also promoted on local tv, radio and many local newspapers).

Overall, we now appreciate the likelihood of selection bias much more in the current study (see also the response to point 2 of the editorial team).

We are also presenting now more clearly the limitations of our study design and the potential and implications of selection and recall bias. We, therefore, deleted to “estimate their prevalence” also from the objectives (see page 3, abstract, objectives and page 5 last para). In addition, we deleted it
from the summary of the key results at the beginning of the discussion (see page 12). In figure 4, we now more clearly present the minimum possible prevalence assuming that all non-responders fully recovered and were free of symptoms and emphasised this also in the text (details see figure 4 and page 15, para 2). Considering the potential of selection bias and the extreme assumption that all non-responders were free of post-COVID sequelae, the unbiased (“true”) estimates should be in between these estimates and the estimates of the included study population. In addition, to reflect the shift of the focus, we would also suggest modifying the title accordingly to: "Post-acute sequelae of COVID-19 six to 12 months after infection: a population-based retrospective cohort study" (see new title on page 1).

The potential for recall bias for existing symptoms 6-12 months post-COVID while filling out the questionnaire should be low and may be related to the severity of symptoms. Recall bias for symptoms existing before acute SARS-CoV-2 infection, and secondly, during acute SARS-CoV-2 infection may be of more relevance and may lead to an overestimation of the post-COVID symptom prevalence differences and prevalence ratios. This limitation is mentioned in the discussion section and is especially relevant in subjects with neurocognitive sequelae (details see page 15, para 2).

The study protocol is currently only available in German language.

**Point 3:** - When asking about symptoms, were participants specifically asked if they related the symptoms to the SARS-CoV-2 infection? Given the rather unspecific nature of most symptoms of Long Covid specifying whether the symptoms relate to the infection or not (and participants can usually distinguish quite well) is useful to narrow down the symptoms that are a consequence of an infection or were there before or occurred after (e.g. due to some other illness). The authors may have dealt with the issue by comparing pre- and post infection symptoms but it should be more clearly stated how specific symptoms were assessed.

**Response:** We evaluated the list of symptoms at three points in time: firstly, symptoms already existing before acute SARS-CoV-2 infection; secondly, during acute SARS-CoV-2 infection (the question was formulated the following way: “Did you have any complaints or symptoms as part of the acute corona infection at that time?”), and thirdly at the time point, the standardised questionnaire was filled out (“still existing or new onset?”), which resulted in a mean of 8.5 months after the positive test result of the acute infection. So the symptoms during acute infection were clearly related to it. We added this information into the methods section (page 6, last para). We then analysed the difference post-pre infection to focus on symptoms not present before the acute infection. The potential for recall bias for existing symptoms while filling out the questionnaire should be low and may be related to the severity of symptoms and neurocognitive sequelae (page 15, para 2).

**Point 4:** From the description of the statistical analysis it is not entirely clear if time since infection was considered in the models. Since symptoms do decline substantially over time (see also https://doi.org/10.1101/2022.05.26.22275532) it is important to adjust for time since infection because the follow-up differs across participants.

**Response:** We fully agree with this statement. Time since the infection was included in table S2 as a covariate (mutually adjusted). However, we did not find a substantial change over time (except for hair loss). An explanation could be that most of the change occurs during the first three (to six)
months after infection and not during the six to 12 months we studied. This hypothesis is also supported by the data of other studies included in the above cited review.

**Point 5:** The authors acknowledge the limitation of recall bias but also refer to self reports and no validation as limitation. I disagree that the latter two are a limitation since this is the only way to get the information. Self report per se is by no means a limitation and often, medical validation is a pseudovalidation based on the idea that such tests are more valid and reliable. But this is simply to based on evidence. Thus I suggest to remove the reference to self reports and no validation as limitation.

**Response:** We thank you for pointing to this important aspect, fully agree with the reviewer’s assessment and changed the wording of the para accordingly to indicate that the lack of medical validation of self-reports is a limitation, but not the self-reports themselves (see page 15, para 2).

There is a phase 2 of the EPILOC study that addresses medical validation after inviting a sample of suspected Long Covid cases and controls from the participants of EPILOC phase 1.

The potential for recall bias for existing symptoms 6-12 months post-COVID while filling out the questionnaire should be low and may be related to the severity of symptoms. Recall bias for symptoms existing before acute SARS-CoV-2 infection, and secondly, during acute SARS-CoV-2 infection may be of more relevance and may lead to an overestimation of the symptoms prevalence difference and prevalence ratio. This limitation is mentioned in the discussion section (details see page 15, para 2).

**Point 6:** Finally, the authors may want to point out the unique points of the study a bit more. Most results are not surprising and are in line with previous, but often smaller and less valid studies. Thus study substantially strengthens the evidence base which would be good to point out more.

**Response:** We revised and strengthened the unique points of this study based on the very helpful and constructive comments of the editorial board and the reviewers. Besides revision of the manuscript text this is now also reflected in a revision of the section “What this study adds” (see page 3).
Reviewer: 3

Point 1: Comments: Interesting work from a German group on PASC/Long Covid.

Response: Thank you very much for the thorough and very positive evaluation of this work. We are very grateful for the constructive comments and think the paper improved considerably with this revision.

I have the following comments.

Point 2: a. The paper would be strengthened considerably by the inclusion of a control group. The lack of control (e.g. people with no known infection) limits the interpretation of the findings significantly.

Response: We only have a before-after comparison within infected subjects (which included, however, all infected subjects within a defined geographic region in a specified time period). On the one hand, this limits the interpretation of findings, as acknowledged in our discussion section (details see page 15, para 2). On the other hand, it minimises confounding as every subject is its own control. In an ideal world, we would prefer both types of controls: within-subject comparisons and uninfected participants with a similar structural homogeneity. However, given the time pressure of capturing enough participants 6-12 months after the first wave and the impossibility of recruiting test-negative subjects with similar structural characteristics (to minimise confounding and bias) 6-12 months after their negative test forced us to the presented approach. (See also reviewer 1, point 2). Unfortunately, in Germany, test-negative controls were not available within the sampling frame of our cases (as negative tests were not reported by name and address to the public health authorities due to lack of a legal basis).

Overall, we now appreciate much more the likelihood of selection bias (and other types of biases) in the current study (see also the response to editorial team point 2). We are also presenting now more clearly the limitations of our study design and the potential and implications of selection and recall bias. In figure 4 we now more clearly present the minimum possible prevalence assuming that all non-responders fully recovered and were free of symptoms and emphasised this also in the text (details see figure 4 and page 15, para 2). Considering the potential of selection bias and the extreme assumption that all non-responders were free of post-COVID sequelae, the unbiased (true) estimates should be between these estimates and the estimates of the included study population. In addition, to reflect the shift of the focus, we would also suggest modifying the title accordingly to: "Post-acute sequelae of COVID-19 six to 12 months after infection: a population-based retrospective cohort study" (see new title on page 1).

The potential for recall bias for existing symptoms 6-12 months post-COVID while filling out the questionnaire should be low and may be related to the severity of symptoms. Recall bias for symptoms existing before acute SARS-CoV-2 infection, and secondly, during acute SARS-CoV-2 infection may be of more relevance and may lead to an overestimation of the symptoms prevalence difference and prevalence ratio. This limitation is mentioned in the discussion section (details see page 15, para 2).

Point 3: b. The methodology of polychoric factor analysis and co-occurrence network analyses are not well described. I found it hard to evaluate given paucity of description in methods. Also how and why is this meaningful? The authors should provide discussion on why this analysis is meaningful and how does it advance our understanding of Long Covid.
Response: We agree. We now provide more details about the factor analysis and definition of symptom clusters. Please also see our detailed response to the editorial committee’s point no. 6.

The primary motivation was to reduce the complexity of the data while preserving the relevant information. Also, we think symptoms in a cluster may be related through a common mechanism, and identified clusters may be helpful for risk factor identification and future clinical studies. In addition, rehabilitation measures could be targeted better by describing prevalent cluster types.

Besides the single symptoms, the evaluation of symptom clusters is highly recommended. NHS England and the Chief Medical Officer of the Scottish Government have formulated this recommendation also in their guideline “COVID-19 guideline scope: management of the long-term effects of COVID-19”. They offer the following definition of post-COVID-19 syndrome: “Signs and symptoms that develop during or following an infection consistent with COVID-19, continue for more than 12 weeks and are not explained by an alternative diagnosis. It usually presents with clusters of symptoms, often overlapping, which can fluctuate and change over time and can affect any system in the body.” They further formulate some key questions. One of them is: “What is the prevalence of symptoms or clusters of symptoms (physical and mental health) and problems carrying out usual activities, including work, education and leisure, among people who have symptoms of COVID-19 beyond 12 weeks?”

(Further details: https://www.nice.org.uk/guidance/ng188/documents/final-scope (last accessed July 1, 2022))

Point 4: c. it is also unclear how "loss of general health" and working capacity are derived and estimated from these questionnaires? these are self-reported? how? what are the questions or sets of questions? have the questions been validated? how do we know that these measure what we think they measure? Again, there is very little detail on this.

Response: To assess working capacity, we adapted questions from the short form of the work ability index as suggested by Tuomi K, Huuhtanen P, Nykyri E, Ilmarinen J. Promotion of work ability, the quality of work and retirement. Occup Med Oxf Engl 2001; 51:318-24.

Work ability (WA) is an important concept in occupational health research and, for over 30 years, assessed worldwide with the Work Ability Index (WAI). We used a modified version of the WAI, which in general evaluates current work ability compared to the best ever achieved working capacity (i.e. in our study: current work ability compared to work ability directly before the acute SARS-CoV-2 infection). Several short measures for work ability have been developed and used over the years.

Most prominent in occupational health research is the Work Ability Score (WAS), which also comprises our used question WAI 1, the single item measuring work ability in relation to lifetime’s best - in our version before acute SARS-CoV-2 infection. Participants in our study assessed their current working capacity (6-12 months after acute infection) compared with the situation before the acute SARS-CoV-2 infection on a 10-point Likert scale (10% steps from 0% to 100%). The wording of the question was “What percentage of your original work capacity (before your positive Corona Test) have you regained today?”

The use of the WAI 1 single question has shown similar relations to sick leave and health-related quality of life in occupational studies (details see Ebner & Hasselhorn. Int J Environ Res Public Health 2019, 16, 3386).

In a similar manner, loss of general health is based on self-assessment by our study participants in analogy to the general question used in many quality of life assessments (e.g. the EuroQol Visual Analogue Scale (EQ-VAS)) but also modified accordingly. Participants assessed their current general
health condition (6-12 months after acute infection) compared with the situation before the acute SARS-CoV-2 infection. They answered the question “What percentage of your general health (before the acute SARS-Cov-2 infection) have you reached again today?” – This was also assessed on a 10-point Likert scale (10% steps from 0% to 100% - 100% means complete recovery) in analogy to the question on working capacity. We added some of the details cited here and now explicitly stated the wording of the two questions in the Methods section of the paper (details see page 7, first para).

Point 5: d. percent missing is not provided for the main data points.

Response: As listed in table 1, there were few missing data for most covariables (the highest number of missing data was observed for cancer as comorbidity with 3.3%). We now added this information to the methods section (page 7, para 2). Please also see our response to the editorial committee’s point no 4.

Point 6: e. the response rate is quite low (24%) and it is unclear how this affects the results? the authors mention older adults and females had a higher response, and refer to table s4 (i looked for table s4, and it is not provided anywhere). will this over or under represent people with LC? This should be discussed in greater detail, and if anything is known about the people who did not respond, it should be included. By the way (minor comment, the tables/figures are not ordered correctly, and it was very confusing for me to go through the supplemental file).

Response: We are sorry for the typo; table S8 in the supplement should have been named S4 (now, after revision S5, as another table has been included).

As outlined in our supplemental table S5 we had a limited response with some overrepresentation of older persons and female sex (appendix, table S5). Our study regions were located around medium-sized university cities, with responders having higher education than the general population, which may limit generalizability. This is addressed as a limitation in our discussion section (page 15, para 2). (The reason why we placed the study region around all our four university cities in the State of Baden-Württemberg is that we are currently inviting 1500 subjects of this study for a thorough clinical examination to evaluate further and validate the reported sequelae and also shed light on the pathophysiology of the different post-COVID sequelae.) (See also response to editorial committee’s point 5).

Point 7: f. It is not clear to me how the symptom clusters in figure 3 were developed? How was this validated? Ideally, you would develop them in a dataset (or half the dataset) and validate them in a second dataset. Also co-occurrence is a function of underlying frequency. I think the approach should take into consideration natural frequency of 2 things happening together, and test whether covid result in excess co-occurrence or in some cases less co-occurrence. i think the lack of control limits your ability to develop this more.

Response: We clarified the definition of the symptom clusters now - please, see the previous detailed response to your point 3.

We only have a before-after comparison within infected subjects (which included all infected subjects within a defined geographic region in a specified time period). On the one hand, this limits the interpretation of findings, as acknowledged in our discussion section (details see page 15, para 2). On the other hand, it minimises confounding as every subject is its own control. In an ideal world, we would prefer both types of controls: within-subject comparisons and uninfected participants with a similar structural homogeneity. However, given the time pressure of capturing enough participants
6-12 months after the first wave and the impossibility of recruiting test-negative subjects with similar structural characteristics (to minimise confounding and bias) 6-12 months after their negative test forced us to the presented approach. (See also reviewer 1, point 2). Unfortunately, in Germany, test-negative controls were not available within the sampling frame of our cases (as negative tests were not reported by name and address to the public health authorities due to lack of a legal basis).

**Point 8:** g. the data visualization is not that clear, it should be optimized, and perhaps decluttered a bit. There is just too much data. Focus the story on the most relevant message and develop the visuals around it.

**Response:** We improved the readability of tables and figures. In figure 4 we now more clearly present the minimum possible prevalence assuming that all non-responders fully recovered and are free of symptoms and emphasise this also in the text (details see figure 4 and page 15, para 2). Considering the potential of selection bias and the extreme assumption that all non-respondents were free of post-COVID sequelae, the unbiased (true) estimates should be in between these estimates and the estimates of the included study population. In addition, to reflect the shift of the focus, we would also suggest modifying the title accordingly to: "Post-acute sequelae of COVID-19 six to 12 months after infection: a population-based retrospective cohort study" (see the new title on page 1).
Reviewer: 4

Comments:

The authors performed a retrospective study aimed at investigating the prevalence of sequelae at 6 to 12 months following SARS-COV2 infection in Germany. They conclude that the burden of self-reported post-acute symptoms and possible sequelae, notably fatigue and neurocognitive impairment, remains considerable after acute infection

The issue is relevant and contemporary. The sample size is large. The study has limitations inherent to the retrospective nature.

Response: Thank you very much for the thorough and very positive evaluation of this work. We are very grateful for the constructive comments and think the paper improved considerably with this revision.

Comments:

Point 1: A Table reporting the features of patients with >= 2 symptoms vs those with 1 symptom/no symptom could be added.

Response: We added the suggested table to the supplement (see table S1). Participants with two or more symptoms were, on average, slightly older (45.4 vs 42.5 years), more often female (64.5 vs 52.6%), more often obese (21.9 vs 14.9%), and required medical care during the acute phase of the infection more often (32.6 vs 11.4%) compared to participants reporting one or no symptoms still present (see page 10, para 2).

Point 2: The limitation that over time changes of symptoms during follow-up were not captured should be included

Response: Indeed, we did not evaluate changes (i.e. trajectories) of symptom development over time after acute SARS-CoV-2 infection, but status at 6-12 months after acute infection (mean was 8.5 months). We further plan to follow up on this cohort 1 year after the last evaluation.

Point 3: The risk of inclusion bias should be specifically added in the limitations

Response: Please see the response to the editorial committee’s point 2.

Point 4: Previous studies reporting the high prevalence of sequelae even at 1 year after COVID-19 (Bellan M, et al. Sci Rep 2021) could be included.

Response: Thank you for pointing this out – we included the suggested study in the revised version (reference 10, cited on page 5, para 2).
Reviewer: 5

Comments:

I found this article well written, concise, and easy to follow. I believe it achieved its stated goals.

Response: Thank you very much for the thorough and very positive evaluation of this work. We are very grateful for the constructive comments and think the paper improved considerably with this revision.

My only comments and questions are these:

Point 1: Did the present study control well enough for pre-existing conditions and comorbidity? Perhaps the survey could be repeated by canvassing medical professionals dealing with such cases, which would be a good way to re-test the results.

Response: We adjusted for pre-existing conditions when analysing factors associated with the identified symptom clusters. For the within-person pre-post comparisons, confounding by pre-existing conditions should not be an issue (see response to reviewer 3, point 2). We are currently inviting 1500 subjects of this cohort for a thorough clinical examination to evaluate further and validate the reported sequelae and shed light on the pathophysiology of the different post-COVID sequelae.

Point 2: The study is slanted to delineating differences between age and sex. However, it might be enlightening to include socio-economic and ethnicity as determinants. Depending on prevalence patterns of COVID-19 in Germany, this might be relevant.

Response: In table S2 and table S3, we included a measure of socioeconomic status. However, we only asked for the nationality of the participants, and 94.2% were of German nationality (see table 1). Ethnicity was not evaluated in the study. Ethnicity is a complex concept, and in Germany, the majority of the residents of the investigated regions are of Caucasian origin. However, in a small proportion (n=2794 of 50,457 or 5.5%), the postal invitations could not be delivered, probably because people were seasonal workers and/or recent immigrants that had already moved to another place without a postal forwarding request. This small group might have been enriched by other ethnicities, which are now not represented in our data. In our invitation letter, help was offered in translating the German questionnaire. However, we cannot exclude that non-German speaking people, some of these of non-Caucasian origin that recently migrated to Germany, are underrepresented in our study. This information is now included as a limitation in the discussion section (page 15, para 2).

Point 3: The study does limit age of participants to 65 years. Since the thrust of the study is the ability to work, this is a valid restriction. However, it does limit the predictive value of the study to the population as a whole.

Response: We agree with the statement. We appreciate this also when talking about the generalizability of the study results to the general population (see page 15, para 2).