PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Life after COVID-19 the road from intensive care back to living: a prospective cohort study
AUTHORS	Wiertz, Carolina; Hemmen, Bena; Sep, Simone JS; van Santen, Susanne; van Horn, Yvette; van Kuijk, Sander; Verbunt, Jeanine

VERSION 1 – REVIEW

REVIEWER	Timothy Walsh Edinburgh University, Critical Care Medicine; Centre for Inflammation Research
REVIEW RETURNED	10-Apr-2022

GENERAL COMMENTS	This is a relevant and topical paper exploring disability after severe COVID19 disease. It is a cohort study with a 'convenience' sample population in Holland, from a regional rehabilitation institution. Evidence is emerging rapidly in this area, and research has been planned and undertaken under challenging circumstances during the pandemic. The focus on the WHO classification of disapbility and attempts to measure participation is of interest. I commend the authors for their efforts. I have the following comments: Methods 1. The cohort classification was referral criteria from regional ICUs. However, it is unclear what the referral criteria were which is critical to placing the population in context of all ICU survivors from COVID19. Were there any exclusions? What were the referral criteria and how was patient selection done. 2. Are there any data on key co-interventions such as steroids or other anti-inflammatory agents? Any data of determinants of neuromuscular function after critical illness generally such as steroids and neuromuscular blockade (paralysis) 3. It would be extremely useful to understand the cohort in the context of all ICU survivors from the referring ICUs. Can any data be provided on this, ideally with some demographic data for the non-referred cohort? 4. The definition of lack of functional impairment prior to COVID19 admission was rather unclear but clearly important to long term outcomes. Can anything be provided or clarified?? 5. Was there any data on critical illness polyneuropathy, which is likely to be prevalent, and is known to predict recovery after critical illness? 6. It is unclear what the rehabilitation programme was for this cohort. Some information would seem relevant. Results 7. What were the characteristics of the 'missed' patients? Were they similar or a potential source of inclusion bias? Can demographics for this group be provided for comparison and to demonstrate inclusion bias is unlikely?

	8. For the primary outcome, are there any population normal values for age/sex matched population that could be used to compare with the cohort to give an indication of the severity of participation restriction. Alternatively can the authors provide data on non COVID ICU patients (perhaps historical controls) to act as a comparator? 9. Similarly, there are population normal comparator data published for some of the outcomes such as the 6 minute walk test. It would be more meaningful to have a 'case-matched' indication of percent of 'normal' age/sex matched values, especially given the small descriptive cohort. 10. The exploration of symptoms associated with participation recovery is interesting, but with these numbers of cases will be inherently underpowered risking type II error. This could be highlighted more clearly (in the discussion). 11. The relationship with non-physical predictors is interesting. There has been some interest in 'resilience' and recovery from critical illness although this is still under-researched. Can the authors comment on the relationship between 'coping style' and resilience as concepts? Coping is discussed in the discussion (see below) but I think it would be useful to specifically clarify the relationship to and relevance of resilience (there is some consideration in the discussion but not clarification for the reader of the different concepts and what they mean for rehabilitation. 12. In terms of interpreting the relationship between breathlessness, fatigue, restrictions, coping style and recovery in participation, the authors state stronger positive associations with those with poorer baseline health status (page 10 paragraph 2). Can the authors comment whether this may simply be a reflection of baseline status or severity of disability, in other words less need or capacity for improvement? Presumably the values of baseline status would enable comment on this? This seems quite likely based on the data in figure 4? 13. In general there are a large number of statistical comparisons undertaken with no correction for multiple testing, in a small cohort. The risk of chance and type 1 error is likely high and needs some specific acknowledgement. This is included in the weaknesses section, but could highlight the need to be cautious interpreting the P values <0.05. 14. In table 1: how was level of education determined? 15. In table 1: there seems to be quite a high prevalence of potentially relevant co-morbidities, especially cardiovascular disease, diabetes, and OSA. Were these explored as associated with outcome, as other ICU recovery cohort studies have demonstrated a strong predictive value from pre-existing health and long term HRQOL and physical function? 16. In table 1 it is unclear why duration of inpatient rehabilitation and coping style data are recorded under the 12 month column? 17. In figure 3, the authors report mean with 95% CIs. While this is reasonable especially to describe populations, given this is a descriptive cohort study in a small population I wonder whether a box and whisker plot would provide more useful information, ie the medians, quartiles and ranges? Discussion 18. There is an emerging literature on factors associated with greater risk of Long COVID, mainly in non ICU patients. Although not directly comparable, it would be relevant to cite the up to date emerging data emerging from research. There are also emerging theories for ongoing Long COVID such as micro-emboli, persistent viral presence, and inflammation. These might be 'effect modifiers' to rehabilitation and worth mentioning.

	For example: https://www.nature.com/articles/s41591-021-01292-y and https://www.dovepress.com/risk-predictors-and-symptom-features-of-long-covid-within-a-broad-prim-peer-reviewed-fulltext-article-POR 19. In the discussion the authors state that it was 'remarkable' (page 13 final paragraph) that markers of acute illness severity were not predictors. However, this was a selected cohort of referrals for rehabilitation, and the number of variables available to describe the acute illness severity and pre-illness health were limited. This highlights the importance of placing this rehabilitation cohort in the context of all ICU survivors from COVID, as noted above. This could be more clearly acknowledged as a weakness in this study. However, I do agree that it is important, as noted in the subsequent sentences, that consideration of mental health and coping need greater consideration in ICU patients. 20. In relation to measuring coping (or by inference resilience), as highlighted on page 14 paragraph 2, it would be useful to comment on the timing of measuring coping as this is unclear in the post ICU patient. For example prior to hospital discharge versus some weeks after discharge home etc.

REVIEWER	Martine Nurek Imperial College London, Surgery and Cancer
REVIEW RETURNED	21-Apr-2022

GENERAL COMMENTS	Thank you for the opportunity to review this important article. It reports the results of a prospective cohort study that aimed to 1) measure recovery of participation amongst post covid-19 patients in the 12 months following discharge from ICU and 2) identify factors that might predict the speed/course of recovery. It finds that 1 year after discharge, many patients remain restricted in their ability to work, exercise, keep house and/or engage in leisure activities. Factors influencing the speed/course of recovery were coping style, anxiety, depression, and self-reported breathlessness and fatigue. I consider this vital research and I have no major concerns regarding its conduct or reporting. The article is comprehensible to a wide audience. I have some minor comments/suggestions, which I have outlined below. ABSTRACT 1. P3, lines 32-37 ("Main outcome measures…mixed-effects model"): the statistical analysis is somewhat unclear. I think it would help if the authors provided a clear statement of the dependent and independent variables in the linear mixed effects model. 2. P3, lines 40-44 ("After one year…recovery of participation"): It would be great to see coefficients and p values for the associations mentioned here. I understand that words are limited in the abstract, so a catch-all phrase (e.g., "all coefficients>x, all p<y") would suffice. INTRODUCTION 3. Well-motivated and very clear. METHODS 4. P8, lines 144-145 ("Calculating T-scores…physical function"): this does not appear to be a full sentence. 5. P8, line 162 ("T0"): presumably "T0" refers to the baseline characteristics (collected from medical transfer letters). If so, I'd

	remind the reader of this here, as it's the first time we've encountered the label "T0". 6. Was an a priori sample size calculation performed? I understand that the sample size is largely out of the authors' control (i.e., a function of the number of eligible patients); still, it would be useful to know whether the study had sufficient power to identify the associations of interest. (If it did not, then this will affect our interpretation of non-significant findings.) RESULTS 7. P9, line 180 ("6.1 95%CI"): I'm unfamiliar with this style of reporting confidence intervals and therefore unsure how to interpret them. I am familiar with the styles used on the next page (e.g., line 189: "95%CI 0.23-0.97" and line 193: "95%CI -0.12 to -0.05"). I suggest that the authors use a consistent method throughout, preferably the last one ("95%CI X to Y") as it is the clearest, in my view. 8. P10, lines 205-207 ("While participation…anxiety complaints"): grammar needs some work here (this is not a full sentence). 9. P25, Figure 3: It seems to me that the 95%CIs for T2 (3 months) and T3 (12 months) overlap, suggesting no significant difference between the two? This is not my area of expertise, so I may be wrong; if so, please educate me because I'd like to learn! 10. P26, Figure 4: To aid interpretation, I suggest that the authors reconsider their use of colour coding here. Currently, blue lines indicate low scores and orange lines indicate high scores. But a low score means different things on the different scales; on some scales, lower scores are "better" (e.g., breathlessness, HADS) whereas on others, they're "worse" (e.g., PROMIS 8b, UPCC). I would invert the colour coding for PROMIS and UPCC, so that blue lines always represent "better" scores and orange "worse" scores. DISCUSSION 11. P14, line 263 ("mental impairments (such as anxiety and depressive symptoms)"): some readers might take issue with this phrase. The authors might consider alternative phrases such as "mental health issues/conditions/challenges". 12. P14 lines 264-268 ("Previous literature…optimal recovery"): again, there are some grammatical problems here that make the sentences difficult to read. I also suggest that the authors explain what they mean by "post-ICU impairments", perhaps by providing some examples in brackets. 13. P14, lines 276-277 ("…or patients…mild infection"): I agree, though I would change "mild infection" to "milder infections", to include the many individuals that had something between a "severe" and a "mild" infection. 14. P15, lines 287-290 ("First, although…within different domains"): I think that the authors need to elaborate: what are the implications of a small sample size? To my eyes, a small sample size means that the study may be underpowered to detect the effects of interest, and therefore non-significant results should be interpreted with caution. Specifically, in an underpowered study, a non-significant result could mean 1) lack of association or 2) insufficient power to detect an association. These are two quite different things, with different implications. If the authors performed a sample size calculation, then it would be very useful to see this, so that we can gauge just how underpowered this study is, if at all (see point 6 above).

	Thanks to the authors for an interesting read.

VERSION 1 – AUTHOR RESPONSE

Comments reviewer 1:

METHODS

1. The cohort classification was referral criteria from regional ICUs. However, it is unclear what the referral criteria were which is critical to placing the population in context of all ICU survivors from COVID19. Were there any exclusions? What were the referral criteria and how was patient selection done.
    a. We would like to thank the reviewer for this valuable comment. A reference and the following text has been added to clarify the process of transferring from IC to rehabilitation centre: '*Patients with an indication for inpatient multidisciplinary rehabilitation were transferred to the rehabilitation centre. The indication was determined in the hospital by a consultant in rehabilitation medicine, based on their clinical judgement of the severity of physical, mental and/or cognitive impairments. (13,14)*. See track changes (Line 95-98, page 5).
2. Are there any data on key co-interventions such as steroids or other anti-inflammatory agents? Any data of determinants of neuromuscular function after critical illness generally such as steroids and neuromuscular blockade (paralysis)
    a. We agree with the reviewer that this is relevant information. However, for information during ICU admission we were dependent on data stated in the medical transfer letters. Information about co-interventions such as steroids, other anti-inflammatory agents or neuromuscular blockade was most of the time not available. Since patients were transferred from 7 different ICU's, it was not feasible to retrieve all this information, especially not during the first COVID-19 wave. Generally, corticosteroids and other anti- inflammatory agents were not standard of care during the first wave of the COVID-19 pandemic in our region. Ventilator strategies included lung-protective ventilation (VT < 6 ml/kg), prone positioning and neuromuscular blockage when sedation was not sufficient. During ICU admission, early physiotherapy with daily passive/active mobilization sessions with or without in-bed cycle ergometry was part of routine care.

3. It would be extremely useful to understand the cohort in the context of all ICU survivors from the referring ICUs. Can any data be provided on this, ideally with some demographic data for the non-referred cohort?
    a. We agree with the reviewer that this is relevant information. Unfortunately, we have no information available on these variables for the non-referred cohort. In the discussion we refer to an article of Heesakker et al (2022), assessing the occurrence of physical, mental and cognitive symptoms among patients with COVID-19 at 1 year after ICU treatment. Heesakker et al included 246 patients, from 11 Dutch hospitals, mean age 61.2 years and 71.5% men. With a median length of stay of 18.5 days in the ICU and 81.5% received invasive mechanical ventilation. Furthermore, van Gassel et al (2021) includes 48 COVID-19 mechanically ventilated ICU patients (in the first COVID-19 wave), mean age 63 year and 69% men. With a median length of stay of 20 days in the ICU and a median of 18.5 days requiring invasive mechanical ventilation. Of these 48 patients, 83% were transferred for inpatient multidisciplinary rehabilitation. Demographic data from Dutch ICU cohorts during the first COVID wave therefore seems to correspond reasonably well with the demographic data from our cohort. Certainly, if we take into account that the majority of the patients eventually is transferred to a rehabilitation centre.
    b. Based on this reviewer's comment, the text in the discussion section has been adjusted to: '*In previous Dutch studies focusing on overall post-ICU COVID-19 survivors, an average age of 61-63 was found, 69-72% men, with a median length of stay of 18-20 days in the ICU (33,34). These demographic data seem to correspond with findings of the current study, taking into account that in the first Covid wave, 83% of all post-ICU COVID-19 patients were eventually transferred to a rehabilitation centre (34)*'. See track changes (Line 248-252, page 13).
4. The definition of lack of functional impairment prior to COVID19 admission was rather unclear but clearly important to long term outcomes. Can anything be provided or clarified?
    a. The definition 'functioning independently before their COVID-19 infection' refers to the indication for inpatient multidisciplinary rehabilitation. This is defined as a premorbid functional level that indicates that he/she is not dependent on others for daily

activities. This information was obtained from the patient or through hetero-anamnesis by caregiver. See also text changes at comment 1 and reference 13. We clarified this in the manuscript text accordingly, as already described in our response to comment 1 (line 95-98, page 5). Furthermore, we made the following adjustments in text for clarification: : '*All patients (aged 18 or older) referred for inpatient rehabilitation after ICU discharge for COVID-19 were eligible to participate in the study. The exclusion criterion was not speaking or reading the Dutch language fluently*'. See track changes (Line 98-102, page 5).

5. Was there any data on critical illness polyneuropathy, which is likely to be prevalent, and is known to predict recovery after critical illness?
    a. Data on critical illness polyneuropathy on the intensive care is not available and measurement of critical illness polyneuropathy is not a routine assessment during ICU admission. Wiertz et al (2021) showed an ICU-acquired weakness in 72.7% of the post-ICU COVID-19 patients in the rehabilitation centre (8). Data were retrieved based on muscle strength testing without additional electromyographic examination. We agree with the reviewer that this is relevant information. Therefore, we added the prevalence of the ICU-acquired weakness of the current population at admission to the rehabilitation centre in table 1. To measure muscle strength, a handheld dynamometer (HHD) was used. HHD values were measured in Newtons and percentages of the norm compared with healthy persons of the same sex, age, and weight. Severe muscle weakness was defined as <80% of the norm score. There is no data available about additional electromyography. Furthermore, we now added the presence of ICU-acquired weakness as a covariate in table 2. See track changes in table 1 and 2.
6. It is unclear what the rehabilitation programme was for this cohort. Some information would seem relevant.
    a. We would like to thank the reviewer for this valuable comment. We have added references regarding the Dutch protocol for multidisciplinary rehabilitation of post-ICU COVID-19 patients. The text has been adjusted to: '*All patients received inpatient multidisciplinary rehabilitation treatment including physiotherapy, occupational*

therapy, speech therapy and psychology personalized to the patient's limitations and needs according to the Dutch guideline for post-COVID ICU rehabilitation (13, 15)'. See track changes (Line 102-105, page 5).

RESULTS

7. What were the characteristics of the 'missed' patients? Were they similar or a potential source of inclusion bias? Can demographics for this group be provided for comparison and to demonstrate inclusion bias is unlikely?
    a. We agree with this omission. In total, 103 post-ICU patients were admitted for inpatient rehabilitation during the inclusion period. Due to the overwhelming number and time pressure on the resource on the COVID-19, ward 23 patients were not approached to participate in this study. To our knowledge, this selection was related to work load on the ward and was not patient related. Of these 23, no informed consent was obtained and therefore no data were collected. Therefore, it is impossible to provide demographics for this group for comparison. In the discussion, this has been pointed out as one of the limitations of the study as '*Third, due to the high workload on the ward, 23 patients were not approached in time for consent to participate. In our opinion, this is unlikely to have led to selection bias, but this cannot be excluded*'. See line 335-337, page 16.
8. For the primary outcome, are there any population normal values for age/sex matched population that could be used to compare with the cohort to give an indication of the severity of participation restriction. Alternatively, can the authors provide data on non COVID ICU patients (perhaps historical controls) to act as a comparator?
    a. We agree with the reviewer that this would be relevant information. The primary outcome variable is participation in society, assessed by the 'Utrecht Scale for Evaluation of Rehabilitation-Participation (USER-P) restriction subscale'. There are no norm values available for healthy persons of the same sex and age, which has been confirmed by the developers of the USER-P scale after we asked for this information prior to the submission of this manuscript (15). We adapted the text in the discussion section to clarify the comparison with other non-COVID patients: '*Mean participation levels increased to 86.1 one year after ICU discharge. As a reference, in*

other non-COVID patients (i.e. stroke, acquired brain injury, progressive neurologic diseases, spinal cord injury and acute coronary syndrome), participation levels between 70.6-83.5 have been observed (35-37). In all non-COVID patient groups, patients mainly reported restrictions in work/education, housekeeping, physical exercise and performing leisure activities, which is in accordance with restrictions in participation reported in current study (36,37)'. See track changes (Line 256-264, page 13).

9. Similarly, there are population normal comparator data published for some of the outcomes such as the 6 minute walk test. It would be more meaningful to have a 'case-matched' indication of percent of 'normal' age/sex matched values, especially given the small descriptive cohort.
    a. We would like to thank the reviewer for this comment. The percentages of predicted of the 6 minute walk test are now added in table 1 (69.5%, 76.9% and 79.2% at one, three and 12 months respectively), see track changes (22). The percentages of predicted for muscle strength and pulmonary function were already shown in table 1.
10. The exploration of symptoms associated with participation recovery is interesting, but with these numbers of cases will be inherently underpowered risking type II error. This could be highlighted more clearly (in the discussion).
    a. We would like to thank the reviewer for this valuable comment, in line with comment number 13. Therefore, this is highlighted more clearly in the strengths and limitations part in the discussion. The text has been adjusted to: '*First, sample size is limited and a number of factors were studied for their effect on the course of participation recovery. The limited sample size contributed to relatively wide confidence intervals. The risk of type II error should therefore be considered while interpreting the data. […]. Still, the p-values of the multiple tests of association should be interpreted cautiously, because we cannot exclude erroneous interpretations of statistical significant findings (i.e. type I error). However, since our results support a certain pattern, we believe that the main conclusions of this study are solid*'. See track changes (Line 323-333, page 16).

11. The relationship with non-physical predictors is interesting. There has been some interest in 'resilience' and recovery from critical illness although this is still under-researched. Can the authors comment on the relationship between 'coping style' and resilience as concepts? Coping is discussed in the discussion (see below) but I think it would be useful to specifically clarify the relationship to and relevance of resilience (there is some consideration in the discussion but not clarification for the reader of the different concepts and what they mean for rehabilitation.

    a. We would like to thank the reviewer for this valuable comment. Coping style and resilience are related, however different. Coping style refers to cognitive and behavioural strategies to handle and manage stressful events or negative psychological and physical outcome (1). While resilience refers to the ability to face the challenges and difficulties of life in a positive and adaptive manner, as well as the capacity to recover from an adverse event (42). Higher levels of resilience have been consistently linked to improved mental and physical health in the general population (43). Furthermore, higher levels of resilience may be a protective factor against the development of disability, chronic illness, depression, or low health-related quality of life (2,3). It is possible to improve the level of resilience of a person, which implies that resilience can be used to improve (emotional) well-being (4). Resilience can be presented as a personal factor in the ICF-model. Therefore, it is an important variable to consider as it will influence the level of social participation. Strengthening resilience may lead to strengthening personal factors, with the possible consequence of improving participation levels. This stresses the need to use an integral vision of health when considering long-term consequences of (COVID-19) ICU survivors. The text in the discussion section has been adjusted to: '*Previous literature on post-ICU patients indicated that critical care recovery has largely focused on post-ICU impairments experienced by patients. Whereas important aspects of recovery within the rehabilitation phase, including coping style and resilience, seem to be ignored (40,41). Resilience refers to the ability to face the challenges and difficulties of life in a positive and adaptive manner, as well as the capacity to recover from an adverse event (42) Higher levels of resilience have been linked to improved mental and*

physical health (43). It is possible to improve the level of resilience of a person. Which implies that resilience can be used to improve (emotional) well-being, with the possible consequence of improving participation levels'. See track changes (Line 288-296, page 14).

12. In terms of interpreting the relationship between breathlessness, fatigue, restrictions, coping style and recovery in participation, the authors state stronger positive associations with those with poorer baseline health status (page 10 paragraph 2). Can the authors comment whether this may simply be a reflection of baseline status or severity of disability, in other words less need or capacity for improvement? Presumably the values of baseline status would enable comment on this? This seems quite likely based on the data in figure 4?
    a. Higher levels of self-experienced breathlessness and fatigue complaints, more perceived limitations in daily life as well as personal factors (having a passive coping style, anxiety complaints or depression complaints) in the first month after ICU discharge were found as early determinants for lower participation levels in the first month after ICU discharge. However, it is noted that despite a different baseline a considerable degree of recovery of participation levels is possible for these determinants. Based on our results we cannot say this is a reflection of severity of disability. Because other early determinants related to the severity of the COVID-19 infection (such as age, sex, Number of comorbidities, Length of ICU stay, duration of mechanical ventilation) did not show significant difference in the recovery of participation over time. Furthermore, no correlation was analysed between self-experienced complaints and severity of illness. These findings may only indicate that non-physical factors, subjectively experienced physical impairments and mental impairments seem more important to determine progress in recovering the level of participation.
13. In general there are a large number of statistical comparisons undertaken with no correction for multiple testing, in a small cohort. The risk of chance and type 1 error is likely high and needs some specific acknowledgement. This is included in the weaknesses section, but could highlight the need to be cautious interpreting the P values <0.05.

a. We would like to thank the reviewer for this valuable comment and refer to the clarifications described in our response to comment number 10.

14. In table 1: how was level of education determined?
    a. We would like to thank the reviewer for this valuable comment. We changed the definition of the level of education according to the International Standard Classification of Education (ISCED) instead of the Dutch classification used in the prior version. See track changes table 1.
        i. Low educational level was determined as ‘primary and secondary education and post-secondary school’.
        ii. High educational level was determined as ‘bachelor’s degree, master’s degree or doctorate or equivalent’.
15. In table 1: there seems to be quite a high prevalence of potentially relevant co-morbidities, especially cardiovascular disease, diabetes, and OSA. Were these explored as associated with outcome, as other ICU recovery cohort studies have demonstrated a strong predictive value from pre-existing health and long term HRQOL and physical function?
    a. We would like to thank you for this valuable comment. It was examined whether the number of co-morbidities were associated with recovery of participation in the first year after ICU discharge, see table 2 (regression coefficient: 0.04 (95%CI -0.34-0.42; p-value 0.83)). The individual co-morbidities were not added as separate covariates because of the relatively small numbers of some of the co-morbidities.
16. In table 1 it is unclear why duration of inpatient rehabilitation and coping style data are recorded under the 12 month column?
    a. We would like to thank you for this valuable comment. The duration of inpatient rehabilitation as well as the coping style data has been moved to the T0 column. The coping style data was collected 12 months after ICU discharge (T3), as stated in the methods section: ‘Proactive coping skills were assessed at T3 with the Utrecht Proactive Coping Scale (UPCC), which is a 21-item self-assessment tool scored on a 4-point scale ranging from ‘not competent at all’ to ‘competent’. See track changes in table 1.

17. In figure 3, the authors report mean with 95% CIs. While this is reasonable especially to describe populations, given this is a descriptive cohort study in a small population I wonder whether a box and whisker plot would provide more useful information, ie the medians, quartiles and ranges?
    a. We would like to thank you for this comment. We have chosen to report confidence intervals as our goal was inference, not to provide a summary of the distribution. Using the confidence interval, the figure reflects precision in our point estimates.

DISCUSSION

18. There is an emerging literature on factors associated with greater risk of Long COVID, mainly in non ICU patients. Although not directly comparable, it would be relevant to cite the up to date emerging data emerging from research. There are also emerging theories for ongoing Long COVID such as micro-emboli, persistent viral presence, and inflammation. These might be 'effect modifiers' to rehabilitation and worth mentioning.
    For example: https://www.nature.com/articles/s41591-021-01292-y
    and https://www.dovepress.com/risk-predictors-and-symptom-features-of-long-covid-within-a-broad-prim-peer-reviewed-fulltext-article-POR
    a. We would like to thank the reviewer for this comment. Based on our research question and findings, our focus is about the long term consequences in post intensive care COVID-19 patients. Based on our result, we cannot directly extrapolate our conclusions to the situation of long COVID-19 patients that were not admitted to an ICU during the acute infection. Therefore, this is highlighted more clearly in the implications for clinical practice and further research part in the discussion. The text has been adjusted to: '*Whether identical variables can be used to identify a delay in recovery in patients who had a milder infection is currently still unclear. In this study, conclusions can be made for a selected group (with ICU admission) of patients. Extrapolation to other populations needs to be done with caution*'. See track changes (Line 303-306, page 15).
19. In the discussion the authors state that it was 'remarkable' (page 13 final paragraph) that markers of acute illness severity were not predictors. However, this was a selected cohort of referrals for rehabilitation, and the number of variables available to describe the acute illness

severity and pre-illness health were limited. This highlights the importance of placing this rehabilitation cohort in the context of all ICU survivors from COVID, as noted above. This could be more clearly acknowledged as a weakness in this study. However, I do agree that it is important, as noted in the subsequent sentences, that consideration of mental health and coping need greater consideration in ICU patients.

a. We would like to thank the reviewer for this comment. For information during ICU admission (such as co-interventions or markers of acute illness such as Apache score, SOFA score or CRP) we were dependent on data stated in the medical transfer letters. For this reason, a distinct set of parameters (length of ICU stay, invasive mechanical ventilation and duration of invasive IMV) were considered as parameters related to severity of the critical illness. Based on these characteristics, our population is largely comparable to other Dutch ICU cohorts as described above. Especially, as we take into account that the majority of the patients (83%) eventually was transferred to a rehabilitation centre, see comment 3. We agree with the reviewer that the number of variables available to describe the acute illness severity were limited. But we also think it is remarkable, that in our set (with medical, psychological and social variables available), the medical variables related to the acute illness severity did not show any association with a delay in recovery while other (psychological) variables as part of the same set did. However, the limited number of variables addressing acute illness severity is now highlighted more clearly in the strengths and limitations part in the discussion. The text has been adjusted to: '*Second, number of variables available to describe the acute illness severity were limited. Further research is needed to investigate this in a larger cohort (all ICU patients, not just rehabilitation patients) to confirm this finding*'. See track changes (Line 333-335, page 16).

20. In relation to measuring coping (or by inference resilience), as highlighted on page 14 paragraph 2, it would be useful to comment on the timing of measuring coping as this is unclear in the post ICU patient. For example prior to hospital discharge versus some weeks after discharge home etc.

a. We would like to thank the reviewer for this valuable suggestion. The implications for clinical practice and the further research part of the discussion was consequently adjusted to: '*Early detection of a passive coping style or mental impairments seems important and should therefore be included in screening during early multidisciplinary rehabilitation. Further research is needed to study the effect of early screening of a patients' level of coping/ resilience during the first months after ICU discharge. As a consequence, an early intervention to increase resilience/strengthen coping on indication could be promising to further strengthen social participation, but needs to be further studied*. See track changes (Line 309-314, page 15).

Comments reviewer 2:

ABSTRACT

2. P3, lines 32-37 (“Main outcome measures…mixed-effects model”): the statistical analysis is somewhat unclear. I think it would help if the authors provided a clear statement of the dependent and independent variables in the linear mixed effects model.
    a. We would like to thank the reviewer for this valuable comment. Text has been adjusted to ‘*Statistical analyses: linear mixed-effects model, with recovery of participation levels as dependent variable. Patient characteristics in the domains of body function, activity limitations, personal and social factors were added as independent variables*’ See track changes (Line 37-39, page 2).
3. P3, lines 40-44 (“After one year…recovery of participation”): It would be great to see coefficients and p values for the associations mentioned here. I understand that words are limited in the abstract, so a catch-all phrase (e.g., “all coefficients>x, all p<y”) would suffice.
    a. We would like to thank the reviewer for this valuable comment. Given the high range of the regression coefficients with positive (more complaints of breathlessness, fatigue, anxiety or depression) and negative (less pro-active coping style and more perceived limitations in daily life) estimates, it is difficult to describe the regression coefficient as ‘all coefficients>x’. Instead, we have clarified in the abstract the direction of the association. The text has been adjusted to ‘*Self-reported complaints of breathlessness and fatigue, more perceived limitations in daily life, as well as personal factors (less pro-active coping style and anxiety/depression complaints) were associated with delayed recovery of participation (all p-value < 0.05)*’. See track changes (Line 46, page 2).

INTRODUCTION

4. Well-motivated and very clear.
    a. Thank you very much for this positive feedback

METHODS

5. P8, lines 144-145 (“Calculating T-scores…physical function”): this does not appear to be a full sentence.
    a. Text has been adjusted to ‘*A web-based scoring service was used to calculate T-scores (maximum score 60.1 and 50.0, corresponding to the mean in the general population of the USA), whereas a higher score indicates better physical function*’. See track changes (Line 151-154, page 7).
6. P8, line 162 (“T0”): presumably “T0” refers to the baseline characteristics (collected from medical transfer letters). If so, I’d remind the reader of this here, as it’s the first time we’ve encountered the label “T0”.
    a. We agree with the reviewer that it is the first time the label ‘T0’ is used. Therefore, the text at the beginning of the methods has been adjusted to ‘*Data was collected in the form of baseline information at admission to the rehabilitation centre (T0), through physical examination and self-administered questionnaires after one (T1), three (T2) and twelve months (T3)*’. See track changes (Line 113-115, page 5). Furthermore the following text has been adjusted to ‘*Data on age, sex, comorbidities and parameters related to critical illness was collected from the medical transfer letters (T0)*’. See track changes (Line 126-127, page 6).
7. Was an a priori sample size calculation performed? I understand that the sample size is largely out of the authors’ control (i.e., a function of the number of eligible patients); still, it would be useful to know whether the study had sufficient power to identify the associations of interest. (If it did not, then this will affect our interpretation of non-significant findings.)
    a. We would like to thank the reviewer for this valuable comment. Since the study population is a clinical cohort, a priori sample size calculation was not performed. Therefore, it is important to interpret the results cautiously. This is highlighted more clearly now in the strengths and limitations part in the discussion, as already pointed out in the responses above: ‘*First, the sample size is limited and a number of factors were studied for their effect on the course of participation recovery. The limited sample size contributed to relatively wide confidence intervals. The risk of type II error should therefore be considered while interpreting the data. A post-hoc power*

calculation revealed however that the study had 90% power to detect an effect size of 0.2 for changes in participation levels over time (alpha = 0.05, mean correlation between repeated measures = 0.53). Still, the p-values of the multiple tests of association should be interpreted cautiously, because we cannot exclude erroneous interpretations of statistical significant findings (i.e. type I error). However, since our results support a certain pattern, we believe that the main conclusions of this study are solid'. See track changes (Line 323-333, page 16).

RESULTS

8. P9, line 180 ("6.1 95%CI"): I'm unfamiliar with this style of reporting confidence intervals and therefore unsure how to interpret them. I am familiar with the styles used on the next page (e.g., line 189: "95%CI 0.23-0.97" and line 193: "95%CI -0.12 to -0.05"). I suggest that the authors use a consistent method throughout, preferably the last one ("95%CI X to Y") as it is the clearest, in my view.
    a. We would like to thank you for signalling the use of the different ways to describe the 95% CI. The text has been adjusted to '*Mean participation levels increased from 62.0 (95%CI 55.9-68.1), 76.5 (95%CI 71.9-81.1) to 86.1 (95%CI 80.6-91.6) at one, three and 12 months respectively*'. See track changes (Line 189-190, page 8).
9. P10, lines 205-207 ("While participation…anxiety complaints"): grammar needs some work here (this is not a full sentence).
    a. We would like to thank you for signaling the incomplete sentence. The text has been adjusted to: '*For patients with a HADS anxiety score ≥8, no differences were found in participation levels in the first 3 months, while there was a significant difference in recovery of participation levels between 3 and 12 months. However, for patients with fewer anxiety complaints (HADS anxiety score ≤8) participation levels significantly improved between 1 and 3 months and between 3 and 12'*. See track changes (Line 212-217, page 9).
10. P25, Figure 3: It seems to me that the 95%CIs for T2 (3 months) and T3 (12 months) overlap, suggesting no significant difference between the two? This is not my area of expertise, so I may be wrong; if so, please educate me because I'd like to learn!

a. We would like to thank you for this valuable comment. Although a difference between groups will always be significant if the 95% confidence intervals of both groups do not overlap, it is not the case that the difference is not significant in all cases where they do overlap. The null-hypothesis that is tested is that the difference between groups is 0, which, in this situation, was rejected at the 0.01 level. Please mind that both confidence intervals exclude the point-estimate of the other group (76.5 (95%CI 71.9-81.1) and 86.1 (95%CI 80.6-91.6) at three and 12 months respectively).

11. P26, Figure 4: To aid interpretation, I suggest that the authors reconsider their use of colour coding here. Currently, blue lines indicate low scores and orange lines indicate high scores. But a low score means different things on the different scales; on some scales, lower scores are "better" (e.g., breathlessness, HADS) whereas on others, they're "worse" (e.g., PROMIS 8b, UPCC). I would invert the colour coding for PROMIS and UPCC, so that blue lines always represent "better" scores and orange "worse" scores.
    a. We would like to thank the reviewer for this suggestion, the colours have been adjusted accordingly. See changes in figure 4.

DISCUSSION

12. P14, line 263 ("mental impairments (such as anxiety and depressive symptoms)"): some readers might take issue with this phrase. The authors might consider alternative phrases such as "mental health issues/conditions/challenges".
    a. The text has been adjusted to: '*Contrary to expectations, this may indicate that non-physical factors such as coping style, subjectively experienced physical impairments (including fatigue and breathlessness) and mental health issues (such as anxiety and depressive symptoms) seem more important to determine the participation recovery levels*'. See track changes (Line 285-288, page 14).
13. P14 lines 264-268 ("Previous literature…optimal recovery"): again, there are some grammatical problems here that make the sentences difficult to read. I also suggest that the authors explain what they mean by "post-ICU impairments", perhaps by providing some examples in brackets.

a. We would like to thank you for signaling the grammatical problems. See the adjustments in discussion in response to comment 11 of the other reviewer. Track changes (Line 288-296, page 14).

14. P14, lines 276-277 ("…or patients…mild infection"): I agree, though I would change "mild infection" to "milder infections", to include the many individuals that had something between a "severe" and a "mild" infection.

    a. The text has been adjusted to: '*Whether identical variables can be used to identify a delay in recovery in patients who had a milder infection is currently still unclear*'. See track changes (Line 303-306, page 15).

15. P15, lines 287-290 ("First, although…within different domains"): I think that the authors need to elaborate: what are the implications of a small sample size? To my eyes, a small sample size means that the study may be underpowered to detect the effects of interest, and therefore non-significant results should be interpreted with caution. Specifically, in an underpowered study, a non-significant result could mean 1) lack of association or 2) insufficient power to detect an association. These are two quite different things, with different implications. If the authors performed a sample size calculation, then it would be very useful to see this, so that we can gauge just how underpowered this study is, if at all (see point 6 above).

    a. We thank the reviewer for the concerns raised. We fully agree that it is important to elaborate on potential power issues, even more so in the context of ad hoc analyses in COVID-19 cohorts. To assist the reader in interpreting the results, we have highlighted the risk of type II error in the discussion section, and added a post-hoc power calculation (see line 323-333, page 16). We hope that the implications of a small sample size are more clear now. Also, we adapted the text in the conclusion section accordingly, to highlight the indicative nature of the findings: "*Our results indicate that progress of recovery in participation in the first year after discharge is associated with early determinants in coping style, subjectively experienced physical impairments (breathlessness and fatigue) and mental impairments (anxiety and depression) rather than medical variables*." See track changes (Line 347-351, page 17).

Thanks to the authors for an interesting read.

VERSION 2 – REVIEW

REVIEWER	Timothy Walsh Edinburgh University, Critical Care Medicine; Centre for Inflammation Research
REVIEW RETURNED	09-Aug-2022

GENERAL COMMENTS	Thank you for the opportunity to review this revised manuscript. I have restricted myself to responses to my previous comments. Comment 1, 2 Comment 3 This has been well answered. My only minor point would be to acknowledge in the methods that details of non-referred patients were not available, such that this was a selected cohort. This might also be noted as a limitation in terms of overall generalisability to post COVID patients. However, the reference to other studies showing similar characteristics is useful. Comment 4, 5, 6, 7, 8, 9, 10, 11 This has been adequately addressed Comment 12 The authors provide a detailed response to the comment. However, I am still a little unclear whether the observed greater association with poorer baseline function might in part simply reflect the great capacity to improve from a poorer baseline? This might generate correlations due to a 'ceiling effect' with smaller changes possible for those starting from a 'better' baseline? Comment 13, 14, 15, 16, 17, 18, 19, 20 Addressed The authors have done a thorough job of reviewing the manuscript.

REVIEWER	Timothy Walsh Edinburgh University, Critical Care Medicine; Centre for Inflammation Research
REVIEW RETURNED	09-Aug-2022

GENERAL COMMENTS	Thank you for the opportunity to review this revised manuscript. I have restricted myself to responses to my previous comments. Comment 1, 2 Comment 3 This has been well answered. My only minor point would be to acknowledge in the methods that details of non-referred patients were not available, such that this was a selected cohort. This might also be noted as a limitation in terms of overall generalisability to post COVID patients. However, the reference to other studies showing similar characteristics is useful. Comment 4, 5, 6, 7, 8, 9, 10, 11 This has been adequately addressed Comment 12 The authors provide a detailed response to the comment. However, I am still a little unclear whether the observed greater association with poorer baseline function might in part simply reflect the great capacity to improve from a poorer baseline? This might generate

	correlations due to a 'ceiling effect' with smaller changes possible for those starting from a 'better' baseline? Comment 13, 14, 15, 16, 17, 18, 19, 20 Addressed The authors have done a thorough job of reviewing the manuscript.

VERSION 2 – AUTHOR RESPONSE

Comments reviewer 1:

Comment 1, 2: Addressed

Comment 3: It would be extremely useful to understand the cohort in the context of all ICU survivors from the referring ICUs. Can any data be provided on this, ideally with some demographic data for the non-referred cohort?

- This has been well answered. My only minor point would be to acknowledge in the methods that details of non-referred patients were not available, such that this was a selected cohort. This might also be noted as a limitation in terms of overall generalisability to post COVID patients. However, the reference to other studies showing similar characteristics is useful.
    b. We would like to thank the reviewer for this comment. We disagree that there is a selected cohort. As we only include post-ICU COVID-19 patients with an indication for inpatient multidisciplinary rehabilitation. However, we agree with the reviewer that we cannot make a statement about the overall post-ICU COVID-19 population. Based on this comment, the aim of the study has been better clarified to: '*Consequently, the aim of this study is to evaluate the* recovery of *participation of COVID-19 patients in the first year after ICU discharge followed by inpatient rehabilitation*'. See track changes (Line 86-88, page 4). Furthermore, the text in the strengths and limitations part in the discussion has been adjusted to: '*Second, number of variables available to describe the acute illness severity were limited. Patients referred for inpatient multidisciplinary rehabilitation were included in this study. Generalisation of the results to all ICU survivors needs to be performed with caution, and needs further study*'. See track changes (Line 335-338, page 16).

Comment 4, 5, 6, 7, 8, 9, 10, 11: This has been adequately addressed

Comment 12: In terms of interpreting the relationship between breathlessness, fatigue, restrictions, coping style and recovery in participation, the authors state stronger positive associations with those with poorer baseline health status (page 10 paragraph 2). Can the authors comment whether this may simply be a reflection of baseline status or severity of disability, in other words less need or capacity for improvement? Presumably the values of baseline status would enable comment on this? This seems quite likely based on the data in figure 4?

- The authors provide a detailed response to the comment. However, I am still a little unclear whether the observed greater association with poorer baseline function might in part simply reflect the great capacity to improve from a poorer baseline? This might generate correlations due to a 'ceiling effect' with smaller changes possible for those starting from a 'better' baseline?

    c. We would like to thank the reviewer for this valuable comment. Higher levels of self-experienced breathlessness and fatigue complaints, more perceived limitations in daily life as well as personal factors (having a passive coping style, anxiety complaints or depression complaints) in the first month after ICU discharge were found as early determinants for lower participation levels in the first month after ICU discharge. For illustrative purposes only, subgroups have been created based on cut-off value (HADS) or the median if no specific cut-off value is known (figure 4). Where, it is noted that despite a different baseline a considerable degree of recovery of participation levels is possible for these determinants. We agree with the reviewer that a poorer baseline function may have better capacity to improve. A ceiling effect for participation levels cannot be completely ruled out. However, the maximum score of the USER-P restriction subscale is 100. One year after ICU discharge a mean participation restriction level of 86.1 was found. The 'ceiling score' of the USER-P restriction subscale is not reached. Based on this comment we added the following text in the discussion part: '*Poor baseline situation may also have provided more opportunity to improve. However, with a mean participation restriction level of 86.1 one year after ICU discharge, the maximum score of 100 of the USER-P restriction subscale has not been reached'.* See track changes (Line 276-279, page 14).

Comment 13, 14, 15, 16, 17, 18, 19, 20: Addressed.

Comments reviewer 2:

The authors have addressed my concerns. I wish them all the best with their future research.