PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The MONITOR-IC study, a mixed methods prospective multicenter controlled cohort study assessing five-year outcomes of ICU survivors and related healthcare costs: a study protocol
AUTHORS	Geense, Wytske; Zegers, Marieke; Vermeulen, Hester; van den Boogaard, Mark; van der Hoeven, Hans

VERSION 1 - REVIEW

REVIEWER	J Hofhuis Gelre Hospital, Apeldoorn The Netherlands
REVIEW RETURNED	21-Jun-2017

GENERAL COMMENTS	This protocol is aiming to conduct a multicentre prospective cohort study of patients in ICU units in four Dutch hospitals. Patients will be followed for five years. They estimate to include 12.000 ICU patients between July 2016 and July 2021. Primary outcomes are HRQOL, and physical, cognitive and mental symptoms. Secondary outcomes: ICU survivors care and support needs, their health care use and related costs. A control cohort of otherwise seriously ill patients will be assembled to compare long-term patients reported outcomes. Although the long-term follow-up data are a strength of the study, I have concerns with the study design and would therefore challenge the authors' protocol. The authors are using a mixed methods design. Very importantly, this study assesses HRQOL before ICU admission, which most of the long-term studies of critically ill survivors don't. However it is not clear to me when the baseline measurement is performed and is this only by the patient (prospectively or retrospectively) or also when this is not possible by close family members? Lastly who will perform all these assessment? Consistently the same investigator? In person or by telephone? Furthermore is this study feasible? The authors are using many questionnaires at the same time and sometimes the questions of the different questionnaires looks like te same. This will possibly influence the loss to follow up rates and also influence the reliability of the study because the many questions can be boring for the patients and may influence the responses, patients will give a faster answer or no answer because they want to be ready and this may influence the reliability of the study. Further comments: 1. Abstract: contains no information regarding the statistical analysis

	2. Methods: It is not clear when the baseline questionnaire on the ICU (acute admissions) is taken, and if this questionnaire will be answered by patient or also when not possible by family members? and by personal interview of self-administered questionnaires ? 3. Inclusion criteria: Why 16 years or older and not 18 years and older?, 12 hours ICU admission for the control group is this not very short? In that way the control group will differ markedly from the study cohort. They estimate to include 12.000 ICU-patients in five years, is this feasible? 4. Questionnaires: The authors are using many questionnaires at the same time point during the five years in a patient group who is in majority older. This can be a limitation. 5. Why are they using the SF-36 and the EQ5-D?

REVIEWER	Leanne Aitken City, University of London, UK
REVIEW RETURNED	22-Jun-2017

GENERAL COMMENTS	The authors have presented a protocol for a large multi-centre cohort study with control group that is currently underway. The potential strengths of this study are significant, most notably the large sample size and potential to answer a myriad of questions regarding the relationship between pre-existing conditions, elements of care and different aspects of recovery. The use of mixed methods and recruitment across multiple sites are also important strengths. The protocol paper, in its current form, does not do justice to the researchers in terms of providing the detail that is required to clearly articulate what data will be collected, and what questions will be answered through the analysis phase. I have no doubt this information exists given the obviously significant funding that has been obtained, however it may be the case that too much detail has been removed in preparing this paper (no doubt with the aim of fitting within the word count). The authors have made a general assertion that there is an urgent need to shift our focus from short-term mortality to long-term outcomes of ICU survivors – while the principle that the focus needed to shift is accurate, I believe this refocusing has taken place extensively over the past 10 – 15 years. Further, we now have good descriptions of most aspects of outcome after critical illness, although this has primarily been in small cohorts and often for short to medium (up to approximately 1 year) timeframes after ICU. As a consequence the authors of the planned study are well placed to fill some of the gaps that exist in terms of long-term outcomes and being able to model the factors that affect the various elements of outcome. Specific comments and suggestions include: The introduction could be more specific in highlighting what it is we already know and what it is we do not know, for example a number of small studies examining mental health have been quoted, but 3 extensive systematic reviews dealing with anxiety (Nikayin et al, General Hospital Psychiatry 2016), depression (Rabiee et al, CCM 2016) and PTSD (Parker et al, CCM 2015) have been omitted. Various reviews of physical function after critical illness also exist and would be useful to incorporate. Similarly, there are various systematic reviews of the interventions that are mentioned (page 6, lines 3 – 4); while I note that one of these reviews examining

	rehabilitation interventions has been referred to, there are various other reviews including some Cochrane Reviews that indicate inconclusive evidence in the areas of most of the interventions suggested – these could be made use of to help demonstrate the gaps in current evidence. The objectives and research questions (page 6, lines 37 – 50) are sometimes vague, and often multidimensional. Specifically: - What is meant by 'more insight'? - PICS is not a single concept, but an amalgamation of various concepts – for patients this consists of physical health, cognitive health and mental health; given one measure of PICS does not exist, but instead multiple measures of the various concepts I assume multiple prediction models are planned but this is not clear in either the research questions or the analysis (page 13) - It is not clear to me what you mean by determining 'the ratio between HRQoL and healthcare related costs' - 'care and support needs' is very vague and would benefit from increased specificity. Page 7, line 42 – do you mean 'ICU admissions' rather than 'hospital admissions? Page 7, lines 42 – 49 – the description of sample size is incomplete and lacks coherence. It is not clear to me why you have indicated 23,500 patients 'could be included' when the numbers of admissions that you indicate adds up to 45,500 over 5 years – this suggests that you are not planning to recruit for the full 5 years but I could not find that information anywhere. You have indicated an inclusion rate of 60% but no rationale for this figure has been provided. There is no justification for the planned sample size of 12,000 patients in the study cohort or 3000 patients in the control cohort – what analysis has this sample size been based on? There is no indication of what drop-out rate you anticipate (incorporating post recruitment mortality, withdrawal and loss to follow-up) other than that you have indicated an expectation that it will be substantial. Given the expectation that drop-out rate will be substantial, what strategies are in place to minimise drop-out, and how will you deal with drop-out in your analysis? It is unusual to identify multiple measures as the primary outcome – this links with the sample size question above to aid clarity; is it possible to identify a single primary outcome that the study has been powered to answer? Table 1 does not add significantly to the message conveyed in my opinion, consider combining tables 1 and 2. It is beneficial to make the questionnaires available in different formats as you have indicated (paper and online) however there is some evidence that format of completion may result in a bias in the results; will the method of completion be recorded and potentially included as a variable in the analysis? Questionnaires – as the authors have acknowledged, there is a large number of proposed questionnaires and ensuring completion of these while minimising participant drop-out will require a significant amount of effort. There has been no rationale or justification for selection of each of the questionnaires, particularly where there is overlap. For example, SF-36 and EQ-5D both measure HRQoL yet no justification for both has been provided. I suggest the detail of all the instruments could be summarised in a table (given they are generally well known and validated instruments), with the detail shifted to supplementary information. By removing some of this detail from the body of the paper, explanation of the rationale for each of the instruments could then be provided. Further detail about the variables to be collected from the medical

	record would be useful, for example what pre-existing comorbidities and will this be using something like the Charlson Comorbidity Index, how will disease severity be measured, how will delirium be assessed and how often, what medications will be recorded? There is an indication that the information gained from this study will enable development of interventions to prevent or mitigate long-term consequences as well as the detection of unnecessary ICU care, but there is currently insufficient detail about what ICU care will be measured, and how, to enable an appreciation of how this will be achieved. Given the sample size and length of follow-up this could be a significant outcome of the study and it would be useful to more clearly convey the extent of the potential. The mixed methods nature of the study is a strength, although the role of the interview data, and how this will be integrated with the information from the survey and modelling data, is not clear. Additional detail about how patients will be purposively sampled (based on what 'various experienced outcomes', page 11, line 51) would also be helpful. Interviews – given the plan for data saturation to be the point when interviews cease, I assume analysis will occur concurrently with the interviews, but this is not clear. Further, will it be the same person undertaking the interviews and completing the analysis or different people. Finally, you have indicated that 2 researchers will independently code the transcripts – some detail about how these 2 sets of analyses will be compared and combined would be useful. Subgroups – in the relevance of findings (page 13, line 55) subgroups of patients have been mentioned, but these have not been previously identified in either the research questions, recruitment methods or analysis – further detail would be useful. Page 14, lines 4 – 5: a number of the interventions that are listed here as being implemented do not yet have clear evidence of benefit – suggest this needs to be modified to reflect the current body of evidence. Page 14, lines 8 – 16: I find this section regarding the influence on policy confusing – are you able to clarify how this influence might occur. Similarly, I am having difficulty seeing the links between the work that you describe and the influence on healthcare professional decision making – this may be because of the lack of detail in the methods and once more detail is provide the potential will become clear, but I suggest you reconsider this section to make the potential links more obvious. The authors are correct in their identification that this study does not provide any information about PICS – Family but that is a different question – it is not reasonable for you to need to address all aspects of recovery in one sudy. Overall: The manuscript would benefit from editorial review to ensure the use of consistent tenses throughout.

VERSION 1 – AUTHOR **RESPONSE**

Reviewer #1 = R.1

Response of authors = A

This protocol is aiming to conduct a multicentre prospective cohort study of patients in ICU units in four Dutch hospitals. Patients will be followed for five years. They estimate to include 12.000 ICU patients between July 2016 and July 2021. Primary outcomes are HRQOL, and physical, cognitive

and mental symptoms. Secondary outcomes: ICU survivors care and support needs, their health care use and related costs. A control cohort of otherwise seriously ill patients will be assembled to compare long-term patients reported outcomes.

R.1.1

Although the long-term follow-up data are a strength of the study, I have concerns with the study design and would therefore challenge the authors' protocol. The authors are using a mixed methods design. Very importantly, this study assesses HRQOL before ICU admission, which most of the long-term studies of critically ill survivors don't. However it is not clear to me when the baseline measurement is performed and is this only by the patient (prospectively or retrospectively) or also when this is not possible by close family members?

A.1.1

The baseline measurements could be performed in two different ways:

1. Before the ICU admission: for example the cardiothoracic surgery patients can fill in the baseline questionnaire preoperatively, before their ICU admission.

2. During the ICU admission: patients have to fill in the questions regarding their situation two weeks before the ICU admission. If patients are not able to fill in the questionnaire by themselves, their relatives will be asked.

The following sentence is added in the method section (page 8):

All patients, or their relatives in case patients are not able to fill in the questionnaire themselves, will be approached to fill in the self-administered paper based or online questionnaire (depending on their preferences) in total eight times: at ICU admission (T0), at hospital discharge (T1), after 3 months (T2), 12 months (T3), 24 months (T4), 36 months (T5), 48 months (T6) and 60 months after ICU admission (T7). To get insight into the situation before the ICU admission, the baseline questionnaire (T0) is provided when the patients is asked for informed consent.

This could be preoperatively for the planned admissions or after admission at the ICU. Then, patients are asked to rate their situation before the ICU admission.

R.1.2.

Lastly who will perform all these assessment? Consistently the same investigator? In person or by telephone?

A.1.2.

The questionnaires are self- administered. When patients are not able to fill in the questionnaires by themselves, their relatives can do it for them. Patients can specify how they want to receive the follow-up questionnaires: online or paper based.

On page 8 we added the following sentence:

The investigators keep track on when patients should receive the next questionnaire or the postal or telephone reminders after four and six weeks.

R.1.3.

Furthermore is this study feasible? The authors are using many questionnaires at the same time and sometimes the questions of the different questionnaires looks like the same. This will possibly influence the loss to follow up rates and also influence the reliability of the study because the many questions can be boring for the patients and may influence the responses, patients will give a faster answer or no answer because they want to be ready and this may influence the reliability of the study.

A.1.3.

We agree with the reviewer that there are many questionnaires at the same time and that different questionnaires look similar. However, the questionnaires are established in close collaboration with worldwide experts in the field of long-term outcome, and the FCIC (Family and Patient Centered Intensive Care); the national foundation for ICU survivors and their family members. Importantly, we pilot tested the final versions of the questionnaires with the FCIC members. They fully supported the content and they felt that regarding the length of the questionnaire this would be feasible for ICU survivors to fill out. Moreover, we want to use validated questionnaires to measure several symptoms, such as cognitive impairment and post traumatic stress syndrome. Because of the construct validity of these questionnaires, we are not able to remove items from the specific questionnaires. A benefit of the overlap between the different validated questionnaires is the ability to check the internal consistency.

We added a sentence on page 10:

Although we are aware of the overlap between the used questionnaires, it will allows us to check the reliability.

R.1.4

Abstract: contains no information regarding the statistical analysis.

A.1.4.

We have considered to add information regarding the statistical analysis in the abstract. However, due to the many different research questions, related extent of different analyses and the maximum amount of words of 300 we decided not to do.

R.1.5.

Methods: It is not clear when the baseline questionnaire on the ICU (acute admissions) is taken, and if this questionnaire will be answered by patient or also when not possible by family members? and by personal interview of self-administered questionnaires ?

A.1.5.

See our comment on point R.1.1 and 1.2

Questionnaires filled in by proxies is a reliable option. Hofhuis et al., (2003) concluded for example in their study, that the SF-36 questionnaire completed by a proxy can reliable assesses the quality of life of the critically ill patient on admission to the ICU.

Hofhuis, J., Hautvast, J., Schrijvers, A., Bakker, J. (2003) Quality of life on admission to the intensive care: can we query the relatives? Intensive Care Medicine 29:974-979

R.1.6.

Inclusion criteria: Why 16 years or older and not 18 years and older?

A.1.6.

In the Netherlands, it is generally accepted that patients of 16 and 17 years are treated in the adult ICU instead of the Pediatric Intensive Care Unit (PACU). Additionally, patients of 16 years and older are able to give their informed consent for research themselves without the permission of their parents.

We consider to exclude the data of 16 and 17 years old patients, when comparing our results with other international studies.

R.1.7.

12 hours ICU admission for the control group is this not very short? In that way the control group will differ markedly from the study cohort.

A.1.7.

To compare the outcomes such as the quality of life, and symptoms of ICU patients with non-ICU patients we will set up a control group as well. The following adults (≥16 years) patients will be included in the control group:

• admitted to the ICU less than 12 hours. So not a 'true' ICU patient, but a patient admitted for instance for monitoring during a short intervention like bronchoalveolar lavage, or insertion of central venous catheter.

• admitted to the Medium Care Unit or high dependency unit

We changed this in the method section on page 6:

Patients are eligible for the control cohort when they are 16 years or older and admitted either to the ICU for less than 12 hours, or to the Post Anaesthesia Care Unit (PACU), the Medium Care or high dependency unit for instance for monitoring during short interventions, such as bronchoalveolar lavage or insertion of a central venous catheter.

R.1.8.

They estimate to include 12.000 ICU-patients in five years, is this feasible?

A.1.8.

We understand the concern of the reviewer, but the estimation of 12,000 patients is based on: 1) the total initial ICU admissions in the four participating hospitals during five years, and 2) an estimated response rate of 60%, based on previous conducted ICU studies. For example Wolters et al., (2014) reported a response rate of 64% and Van den Boogaard et al., (2012) a response rate of 71%.

Wolters, E., van Dijk, D., Pasma, W., Cremer, O., Looije, M., de Lange, D., Veldhuijzen, D., Slooter, A. (2014) Long-term outcome of delirium during intensive care unit stay in survivors of critical illness; a prospective cohort study. Critical Care.44:1267-1277.

Van den Boogaard, M., Schoonhoven, L., Evers, A., van der Hoeven, J., van Achterberg, T., Pickkers, P. (2012) Delirium in critically ill patients: impact on long-term health- related quality of life and cognitive functioning. Critical Care Medicine. 40:112-118

We changed the inclusion section on page 6:

For the MONITOR-IC study we estimated to include 12.000 patients. This estimation is based on: 1) the initial ICU admissions in the academic hospital and the three other participating hospitals together (2,500 and 2,200 respectively per year), and 2) an estimated response rate of 60%, which is based on previous conducted ICU studies37 38

R.1.9.

Questionnaires: The authors are using many questionnaires at the same time point during the five years in a patient group who is in majority older. This can be a limitation

A.1.9.

We agree with the reviewer that patients have to fill in many questionnaires at the same time point. This could be indeed a reason for loss to follow up. However, the questionnaires are established in close collaboration with worldwide experts in the field and the FCIC (Family and Patient Centered Intensive Care); the national foundation for ICU survivors and their family members. Importantly, we pilot tested the final versions of the questionnaires with the FCIC members. They fully supported the content and they felt that regarding the length of the questionnaire this would be feasible for ICU survivors to fill out; see also point R.1.3.

R.1.10.

Why are they using the SF-36 and the EQ5-D?

A.1.10.

For the assessment of the quality of life, both the SF-36 and the EQ-5D could be used. They are both standard, well validated instrument, straightforward and applicable in different countries and languages. Since 1999, 63% of the articles measuring the quality of life in intensive care patients, use the SF-36 and 19% use the EQ-5D.

We decided to use both the SF-36 and the EQ-5D after an expert meeting with Dale Needham (professor of medicine at the Johns Hopkins University School of Medicine, with an expertise in critical care medicine) and Dr. E. Adang (expert in health technology assessment). They advised to add the EQ-5D to the SF-36, because it is a measure of utility, being necessary for the calculation of quality adjusted survival; a key measure of health effect for cost effectiveness assessments.

Reviewer #2 = R.2

Response of authors = A

The authors have presented a protocol for a large multi-centre cohort study with control group that is currently underway. The potential strengths of this study are significant, most notably the large sample size and potential to answer a myriad of questions regarding the relationship between pre-existing conditions, elements of care and different aspects of recovery. The use of mixed methods and recruitment across multiple sites are also important strengths.

The protocol paper, in its current form, does not do justice to the researchers in terms of providing the detail that is required to clearly articulate what data will be collected, and what questions will be answered through the analysis phase. I have no doubt this information exists given the obviously significant funding that has been obtained, however it may be the case that too much detail has been removed in preparing this paper (no doubt with the aim of fitting within the word count).

The authors have made a general assertion that there is an urgent need to shift our focus from short-term mortality to long-term outcomes of ICU survivors – while the principle that the focus needed to shift is accurate, I believe this refocusing has taken place extensively over the past 10 – 15 years. Further, we now have good descriptions of most aspects of outcome after critical illness, although this has primarily been in small cohorts and often for short to medium (up to approximately 1 year) timeframes after ICU. As a consequence the authors of the planned study are well placed to fill some of the gaps that exist in terms of long-term outcomes and being able to model the factors that affect the various elements of outcome.

Specific comments and suggestions include:

R.2.1.

The introduction could be more specific in highlighting what it is we already know and what it is we do not know, for example a number of small studies examining mental health have been quoted, but 3 extensive systematic reviews dealing with anxiety (Nikayin et al, General Hospital Psychiatry 2016), depression (Rabiee et al, CCM 2016) and PTSD (Parker et al, CCM 2015) have been omitted. Various reviews of physical function after critical illness also exist and would be useful to incorporate. Similarly, there are various systematic reviews of the interventions that are mentioned (page 6, lines 3 – 4); while I note that one of these reviews examining rehabilitation interventions has been referred to, there are various other reviews including some Cochrane Reviews that indicate inconclusive evidence in the areas of most of the interventions suggested – these could be made use of to help demonstrate the gaps in current evidence.

A.2.1.

We thank the reviewer for the literature suggestions. First of all we added the three suggested reviews regarding anxiety (Nikayin et al., 2016), depression (Rabiee et al., 2016) and PTSD (Parker et al., 2015) and adjusted the sentence in the introduction (see page 4) and the related references: In addition, mental impairment, such as depression,13 anxiety,14 and sleep disturbances are common.1 2 In 25% of the ICU survivors, posttraumatic stress disorder symptoms (PTSD) occur at 1 year follow-up.15 These PTSD symptoms can persist for 8 years.2

We also added two reviews regarding patients' physical symptoms and function (see page 4): Desai and Needham (2011) and Granja et al (2012): Examples of these physical impairments are pain, breathing difficulties, fatigue and loss of bodyweight resulting in physical weakness and problems in daily functioning and activities.1 8-10 11

Additionally, we replaced the two RCT studies of Jensen et al., (2016) and Cuthberson et al., (2009) with four systematic reviews, regarding diaries (Aitken et al., 2013 and Ullman et al., 2014), active mobilization (Tipping et al., 2016) and follow-up consultations (Jensen et al., 2015), see page 5.

R.2.2.

The objectives and research questions (page 6, lines 37 – 50) are sometimes vague, and often multidimensional.

Specifically:

• R.2.2.1.

What is meant by 'more insight'?

A.2.2.1.

We agree that this is somewhat vague. We changed the objective on page 5 into: To quantify and describe the extent of the physical, mental and cognitive long-term outcomes and HRQoL of ICU survivors during five years following ICU admission, in order to ultimately improve care for ICU patients.

• R.2.2.2.

PICS is not a single concept, but an amalgamation of various concepts – for patients this consists of physical health, cognitive health and mental health; given one measure of PICS does not exist, but instead multiple measures of the various concepts I assume multiple prediction models are planned but this is not clear in either the research questions or the analysis (page 13)

A.2.2.2.

We changed the second research question on page 5 into: What are important predictors for the various physical, mental and cognitive long-term outcomes?

On page 13, we changed the sentence in the method section: In order to predict the various physical, cognitive and mental long-term outcomes, multiple prediction models will be developed.

• R.2.2.3.

It is not clear to me what you mean by determining ‘the ratio between HRQoL and healthcare related costs’

A.2.2.3.

The third research question refers to an economic evaluation of the costs-effectiveness of ICU care. What are healthcare costs related to (poor) patient outcomes? In our study, we want to calculate quality adjusted life years (QALYs), a measure of the state of health of a person or group in which the benefit, in terms of length of life, adjusted for quality of life. One QALY is equal to one year of life in perfect health, zero QALY is associated with death. QALY is used in economic evaluations, to assess the value for money for example medical treatments and interventions.

We decided not to change the research question, but we added the above mentioned information in the method section page 13.

• R.2.2.4.

Care and support needs’ is very vague and would benefit from increased specificity.

A.2.2.4.

We agree, however, we decided not to change the research question. In this study we want to explore the specific care and support needs of ICU survivors (additional to the usual care), based on the reported outcomes regarding the physical, mental and cognitive problems. Possibly, this will be a need for more information, emotional support and practical support, but in this study we want to find this out

• R.2.2.5.

Page 7, line 42 – do you mean ‘ICU admissions’ rather than ‘hospital admissions?

A.2.2.5.

The reviewer is correct. We changed the sentence on page 6: For the MONITOR-IC study we estimated to include 12.000 patients. This estimation is based on: 1) the initial ICU admissions in the academic hospital and the three other participating hospitals together (2,500 and 2,200 respectively

per year), and 2) an estimated response rate of 60%, which is based on previous conducted ICU studies.37 38

R.2.3.

Page 7, lines 42 – 49 – the description of sample size is incomplete and lacks coherence. It is not clear to me why you have indicated 23,500 patients 'could be included' when the numbers of admissions that you indicate adds up to 45,500 over 5 years – this suggests that you are not planning to recruit for the full 5 years but I could not find that information anywhere. You have indicated an inclusion rate of 60% but no rationale for this figure has been provided.

A.2.3.

The MONITOR- IC study is a prospective cohort study, in which we aim to include all ICU patients admitted to the four participation ICUs. A sample size calculation is therefore unnecessary. The number of patients is an estimation based on 1) the initial ICU admissions in five years in the academic hospital (5 years X 2500 patients) and in four years in the three other centers (4 years X 2200 patients), and 2) the response rate of 60% (see point 1.8).

However, we agree we could have been more specific. We therefore changed the sentences, see point A.2.2.5.

For the MONITOR-IC study we estimated to include 12.000 patients. This estimation is based on: 1) the initial ICU admissions in the academic hospital and the three other participating hospitals together (2,500 and 2,200 respectively per year), and 2) an estimated response rate of 60%, which is based on previous conducted ICU studies37 38

R.2.4.

There is no justification for the planned sample size of 12,000 patients in the study cohort or 3000 patients in the control cohort – what analysis has this sample size been based on? There is no indication of what drop-out rate you anticipate (incorporating post recruitment mortality, withdrawal and loss to follow-up) other than that you have indicated an expectation that it will be substantial. Given the expectation that drop-out rate will be substantial, what strategies are in place to minimize drop-out, and how will you deal with drop-out in your analysis?

A.2.4.

To minimize the drop-out in the study, survey reminders will be carried out. Four weeks after sending the T0 questionnaires, the first postal reminder will be send. Only for the T0 questionnaires this will be followed by a telephone call in case of no response after two weeks. For the other questionnaires, two postal reminders will be send. Moreover, when patients are included in the study, they receive quarterly a MONITOR-IC newsletter about for example the progress of the study.

The number of 3000 patients which we will include in the control group, is an estimation based on case-control studies with an optimal ratio of 1:4. We reversed this in a 4:1 ratio.

R.2.5.

It is unusual to identify multiple measures as the primary outcome – this links with the sample size question above to aid clarity; is it possible to identify a single primary outcome that the study has been powered to answer?

A.2.5.

We agree with the reviewer that it is unusual to have multiple measures as a primary outcome in intervention studies. We are, however, setting up a cohort study in which a construct (PICS) is explored. To avoid confusion, we removed the term primary and secondary outcomes in the abstract (page 2) and method section (page 8): The outcomes of the MONITOR-IC study are the HRQoL among ICU survivors and their physical (fatigue, vulnerability and frailty), cognitive and mental (anxiety, depression and stress) impairments. Additional outcomes are the patients' care and support needs, their healthcare use, and related costs.

R.2.6.

Table 1 does not add significantly to the message conveyed in my opinion, consider combining tables 1 and 2.

A.2.6.

We considered to merge table 1 and 2. However, we decided to keep them separate, because table 1 gives an overview of all used methods related to the different research questions, where table 2

provides an overview of the different questionnaires used in the survey. We think the separated table would be more beneficial for the readers.

R.2.7.

It is beneficial to make the questionnaires available in different formats as you have indicated (paper and online) however there is some evidence that format of completion may result in a bias in the results; will the method of completion be recorded and potentially included as a variable in the analysis?

A.2.7.

We agree that there is some evidence that the format of completion may result in a bias in the results. The method of completion will be recorded, however we will not include this as a variable in the analysis. Patients can choose the format of the survey they prefer; paper or online based. We assume this will lead to a higher response rate and therefore bias is less likely.

R.2.8.

Questionnaires – as the authors have acknowledged, there is a large number of proposed questionnaires and ensuring completion of these while minimising participant drop-out will require a significant amount of effort. There has been no rationale or justification for selection of each of the questionnaires, particularly where there is overlap. For example, SF-36 and EQ-5D both measure HRQoL yet no justification for both has been provided. I suggest the detail of all the instruments could be summarised in a table (given they are generally well known and validated instruments), with the detail shifted to supplementary information. By removing some of this detail from the body of the paper, explanation of the rationale for each of the instruments could then be provided.

A.2.8.

We added an overview of the used validated questionnaires in Supplementary 1, see page 21, describing the amount of questions, domains and scoring options per questionnaire. Additionally, on page 9 we added our rationally for the selection of the questionnaires.

R.2.9.

Further detail about the variables to be collected from the medical record would be useful, for example what pre-existing comorbidities and will this be using something like the Charlson Comorbidity Index, how will disease severity be measured, how will delirium be assessed and how often, what medications will be recorded? There is an indication that the information gained from this study will enable development of interventions to prevent or mitigate long-term consequences as well

as the detection of unnecessary ICU care, but there is currently insufficient detail about what ICU care will be measured, and how, to enable an appreciation of how this will be achieved. Given the sample size and length of follow-up this could be a significant outcome of the study and it would be useful to more clearly convey the extent of the potential.

A.2.9.

We agree that we could be more specific. Data regarding the ICU treatment will be extracted from the medical health record of the patients and the NICE database; the Dutch National Intensive Care Evaluation registry. All participating ICUs register a variety of patient and treatment data of all admitted ICU patients in this registry.

Data is registered regarding, for example, length of ICU and in-hospital stay, length of mechanical ventilation, associated hospital admissions, expected mortality(based on the APACHE II, SAPS II, MPM II and APACHE III-IV models), co-morbidities and the severity of illness (based on the APACHE III). With the registered co-morbidities we are able to calculate the Charlson Comorbidity Index.

Additionally, all participating centers are using the validated delirium assessment tool, CAM-ICU, and the pain assessment tool instrument, the CPOT. Data regarding these assessment tools will be extract from the medical health record as well.

We made the following changes, see page 10:

Patients' demographics and information regarding their diagnosis and treatment, such as primary conditions, pre-existing co-morbidity, disease severity, sepsis, (re)admission, length of mechanical ventilation, length of ICU stay, delirium (CAM-ICU), pain (CPOT), expected mortality (based on the APACHE II, SAPS II, MPM II and APACHE III-IV models and medication, will be extracted from their medical record and the Dutch National Intensive Care Evaluation (NICE) registry.52

R.2.10.

The mixed methods nature of the study is a strength, although the role of the interview data, and how this will be integrated with the information from the survey and modelling data, is not clear. Additional detail about how patients will be purposively sampled (based on what 'various experienced outcomes', page 11, line 51) would also be helpful.

A.2.10.

In the first phase, the quantitative data collection (questionnaires) will start. Using the data from the first questionnaires we will be able to purposively sample patients for the interviews, based on the experienced symptoms or needs. We will conduct the interviews to get more insight into the consequences of their critical illness and its treatment for patients in their daily life. How do patients cope for example with these consequences? What does it mean when they are not able to work anymore? What were consequences for the relationship with their partner and other family members? Which care and support do they need?

With the qualitative data from the interviews, we want to illustrate the quantitative data about HRQoL and symptoms from the questionnaires.

To describe the purposive sampling more specific we added the following words on page 11: Interviews will be conducted until data saturation is reached.

Patients will be purposively sampled based on various experienced outcomes, such as the quality of life, daily functioning, anxiety, depression, and their experienced needs for more information or emotional support.

R.2.11

Interviews – given the plan for data saturation to be the point when interviews cease, I assume analysis will occur concurrently with the interviews, but this is not clear. Further, will it be the same person undertaking the interviews and completing the analysis or different people. Finally, you have indicated that 2 researchers will independently code the transcripts – some detail about how these 2 sets of analyses will be compared and combined would be useful.

A.2.11.

The analysis of the transcripts of the interviews will occur concurrently with the interviews (iterative process). One researcher (WG) shall conducted the interviews. After the first interviews, WG and another independent researcher (MZ, MvdB) will start with the first coding of the transcript. The differences and similarities between the codes will be discussed together and in case of disagreement a third researcher will be involved. In the meetings with the team, the codebook will be refined and emerging categories and themes will be discussed.

We add the following sentences in the method section, page 13:

The differences and similarities between the codes will be discussed together, and in case of disagreement, a third researcher will be involved. In the meetings with the team, the codebook will be refined and emerging categories and themes will be discussed.

R.2.12.

Subgroups – in the relevance of findings (page 13, line 55) subgroups of patients have been mentioned, but these have not been previously identified in either the research questions, recruitment methods or analysis – further detail would be useful.

A.2.12.

We agree that we could be more specific. We added this information in the method section page 13:

Subgroups will be identified based on their illness and condition (for example sepsis, delirium, co morbidities, ARDS), treatment (for example length of ICU stay, duration of mechanical ventilation, dialysis) and social demographics (age, gender, education, family setting etc).

R.2.13.

Page 14, lines 4 – 5: a number of the interventions that are listed here as being implemented do not yet have clear evidence of benefit – suggest this needs to be modified to reflect the current body of evidence.

A.2.13

Based on our study (results from the questionnaires and interviews) we will get a clear overview of all the different problems patients experience and which needs for treatment and support they have. Furthermore, at the moment we are performing a systematic review which will support us later on with more evidence regarding all these interventions. Based on this information we can tailor effective interventions to patient's needs or develop and evaluate new interventions to prevent or mitigate to the identified long term consequences.

R.2.14.

Page 14, lines 8 – 16: I find this section regarding the influence on policy confusing – are you able to clarify how this influence might occur. Similarly, I am having difficulty seeing the links between the work that you describe and the influence on healthcare professional decision making – this may be

because of the lack of detail in the methods and once more detail is provide the potential will become clear, but I suggest you reconsider this section to make the potential links more obvious.

A.2.14.

We agree. The results of our study will be disseminated through international and national publications and presentations, and the results can be helpful to stimulate, or support in order to change policy and protocols regarding early detection of patients who are prone for long term problems, for treatment to prevent or mitigate long term consequences and for follow-up care of ICU survivors. Ultimately, resulting into better care for individual patients on the ICU and ICU survivors.

R.2.15.

The authors are correct in their identification that this study does not provide any information about PICS – Family but that is a different question – it is not reasonable for you to need to address all aspects of recovery in one study.

A.2.15

We agree but we still think it is important to mention. However, we decided to shorten the relevant paragraph, see page 15: Moreover, PICS does not only occur among ICU survivors, but also among their family members and relatives, also called PICS-Family (PICS-F).57 These long-term consequences in families of survivors and non-survivors consist of psychological, physical and social consequences as well.58-60 Although it is important to increase awareness of these possible long-term consequences on family members,2 we decided to focus only on the ICU survivors. In the future extension of this study, family members might be included as well.

Overall: The manuscript would benefit from editorial review to ensure the use of consistent tenses throughout.