PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Enhanced Recovery after Intensive Care (ERIC): study protocol for a stepped wedge cluster randomized controlled trial to evaluate the effectiveness of a critical care telehealth program on process quality and functional outcomes
AUTHORS	Adrion, Christine; Weiss, Bjoern; Paul, Nicolas; Berger, Elke; Busse, Reinhard; Marschall, Ursula; Caumanns, Jörg; Rosseau, Simone; Mansmann, Ulrich; Spies, Claudia

VERSION 1 – REVIEW

REVIEWER	Meeta Prasd Kerlin University of Pennsylvania United States of America
REVIEW RETURNED	31-Jan-2020

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. It presents the study protocol for "ERIC," a stepped wedge, cluster randomized trial of a tele-ICU program. In general, the protocol is clear and detailed. My suggestions for improvement are primarily for further clarifications. By section: Introduction: Page 7/87, lines 44-50 - The authors mention "evidence-based strategies to reduce the PICS/CCI burden." In the following sentence, several of these practices are mentioned. Could the authors provide some citations? METHODS: Page 9/87, lines 31-33 - I was unsure what is meant by "intrasectoral modification." Please clarify this phrase, or remove it if it is not necessary. In the description of the patient inclusion and exclusion criteria, the authors do not detail how they will handle patients who are admitted to an ICU on more than one occasion. Later in the statistical analysis plan, they describe a plan for excluding patients, for example, who are exposed to both control and intervention. However, given that they are evaluating long-term clinical outcomes, it would be reasonable to include only a single admission for patients with multiple ICU admissions, at least for those outcomes (if not the primary outcome, which is more of a process measure). Please clarify the plan for handling patients with multiple admissions, even if they are exposed to only one condition on all admission. Page 12/87, lines 20-22 - please provide the rationale for why data analysts will not be blinded? Can the data processing and management be separate from the analysis, in which case could the

	analyst be blinded to study group? It would be helpful to provide some more detail in the narrative portion of the main manuscript regarding the rounds in the intervention condition, perhaps in the description of the intervention conditions. What do these rounds entail? How is a determination made regarding adherence to the QI measures? They are listed in a table, but it is unclear to me how the determination of yes/no is made. The secondary outcomes may be presented in a table to summarize them more concisely. That would be visually easier than the long bulleted list. In the follow-up procedures, please specify whether the all the clinicians involved "on-demand counseling sessions between the treating GP or rehab facility and investigators of the Charite" will be blinded to the study groups. If the investigators are able to become aware of the study group, there may be a perceive risk of introducing systematic differences in these counseling sessions. In the Process Evaluation section (Page 22/87), could the authors provide more details about intervention monitoring, how they will ensure that rounding is occurring as expected, what standards will be used to assess the adherence to the rounding intervention, and what, if anything, will be done to improve rounding if it is not occurring as planned. Table 1 - it may be useful to incorporate into this table some details regarding what would qualify as adherence to each QI measure. Is it at the discretion of the rounding team, the tele-ICU clinicians, or are there pre-specified definitions for each? If these details are elsewhere in the narrative, they are buried, and it would be useful to be able to access this information quickly, such as from a table. Table 2: Please clarify what is meant by "(x)" in some of the cells of the table. Please note that I was unable to review anything in German, including the website that was referenced in a few placed, and the last document in the PDF file, which I believe was a study protocol.

REVIEWER	Sikandar Khan Indiana University School of Medicine
REVIEW RETURNED	12-Apr-2020

GENERAL COMMENTS	This is a very well written protocol. I would recommend the following clarifications: - The introduction focuses on PICS and ICU survivorship, but the primary outcome of the trial appears to be quality improvement in the ICU. More attention to ICU quality metrics and literature should be included in the introduction. - I would provide more detail on the 8 QI domains and how they are operationalized in German ICUs, this will be helpful for the international audience. - I would address contamination in a bit more detail; do physicians or other staff work at more than 1 of the participating hospitals?

	- Please provide additional details on how adherence of the individual components will be ascertained, and whether it is entirely based on review of the medical records? - How will missing data be handled? - The introduction could be strengthened by discussing how increased completion of the QI domains would affect the secondary outcomes of the study - Is the primary outcome measured at the patient level or the cluster/institution level? Would provide more detail on whether "beneficial effects" are simply increased completion rate of the domain? - Does randomization include use of block design? Also would explain whether randomization is at the cluster level, since the authors also include eligibility criteria at the patient level. - The authors note under trial status that recruitment will end this year (October 2020). It would be helpful to clarify if the protocol has changed during the trial.

REVIEWER	Sue Lasiter University Of Missouri, Kansas City
REVIEW RETURNED	21-Apr-2020

GENERAL COMMENTS	The study protocol is well written and complete.

VERSION 1 – AUTHOR RESPONSE

Reply to Reviewer 1

(Meeta Prasd Kerlin, University of Pennsylvania, U.S.)

Comment:

Introduction:

Page 7/87, lines 44-50 - The authors mention "evidence-based strategies to reduce the PICS/CCI burden." In the following sentence, several of these practices are mentioned. Could the authors provide some citations?

Author Response:

Thank you for this suggestion. We rearranged some existing references and text in the first paragraph of the introduction (line 9 to 13). As requested, we added reference numbers [4, 11-17] which primarily reflect how deep sedation and delirium are associated with unfavorable, long-term patient-centered outcomes and how monitoring of delirium can contribute to a reduction in mortality.

METHODS:

Comment:

Page 9/87, lines 31-33 - I was unsure what is meant by "intrasectoral modification." Please clarify this phrase, or remove it if it is not necessary.

Author Response:

As the phrase ″intrasectoral modification” is not essential to the corresponding sentence specifying the primary objective of the ERIC trial, we removed it. Besides, the word “intra-sectoral” was removed in the header of Table 1 (list of quality indicators). We also changed “intersectoral barriers” (page 7) to “the post-ICU period” to clarify this phrase.

Comment:

In the description of the patient inclusion and exclusion criteria, the authors do not detail how they will handle patients who are admitted to an ICU on more than one occasion. Later in the statistical analysis plan, they describe a plan for excluding patients, for example, who are exposed to both control and intervention. However, given that they are evaluating long-term clinical outcomes, it would be reasonable to include only a single admission for patients with multiple ICU admissions, at least for those outcomes (if not the primary outcome, which is more of a process measure). Please clarify the plan for handling patients with multiple admissions, even if they are exposed to only one condition on all admission.

Author Response:

The reviewer makes an important point. To clarify this issue, we included the following sentence at the end of the section ‘Patient population and eligibility criteria’:

> *“If a study patient is re-admitted to a participating ICU at a later date, the same eligibility criteria will be applied, and the patient can be enrolled on multiple occasions.”*

Patients admitted to ICU and enrolled on multiple occasions, and who are exposed to the same condition as during their index ICU stay are handled as follows:

We are interested in the whole QI trajectory of a patient and will include each episode (QI measurements) during each ICU stay of a patient (provided that a patient is readmitted at the same or another participating ICU while the study is still enrolling patients). To adjust for the dependency over time in the model-based principal analysis, each patient will have his/her own random effect, even if further ICU admissions took place in different randomization units. The day of ICU admission and treatment day will provide the unique indicator to which condition the patient was exposed to (i.e. control or intervention).

We will exclude an ICU period of a patient when it overlaps with the crossover time of the randomization unit, i.e. the patient is exposed to both conditions.

The time schedule of implementing the intervention in a cluster defines its control and intervention phase (switch-over date as defined by randomization of clusters performed prior to trial commencement).

Comment:

Page 12/87, lines 20-22 – please provide the rationale for why data analysts will not be blinded? Can the data processing and management be separate from the analysis, in which case could the analyst be blinded to study group?

Author Response:

Thank you for raising this issue. We agree that this point requires further clarification. As stated in the section ‘Randomization, allocation concealment and masking’, ICU personnel cannot be blinded due to the nature of the intervention. The structured round alongside the QIs is led by a critical care specialist that also evaluates whether or not a QI is fulfilled; in the control group, the same critical care team evaluates the QI adherence on the basis of electronic medical record (EMR) data. The limited availability of critical care specialists mandated that we could not implement two independent teams for the consultation and the evaluation (i.e. assessment of QI adherence). This might be a potential limitation.

Concerning trial statisticians, we added the following lines to the manuscript (page 11):

"*Independent data analysts are not involved in outcome assessment. They will be handed the final datasets for evaluation by the consortium leader in order to perform the pre-planned statistical analyses."*

In order to enable an objective analysis, the statistical analysis plan (SAP) will be written and published before the final analyses will be done. This was already mentioned at the end of the section 'Statistical methods – clinical effectiveness'. The University of Munich is the independent consortium partner responsible for the evaluation of clinical effectiveness and the statistical concept (see Supplementary File 2) and is not involved in the assessment or intervention delivery.

Comment:

It would be helpful to provide some more detail in the narrative portion of the main manuscript regarding the rounds in the intervention condition, perhaps in the description of the intervention conditions. What do these rounds entail?

How is a determination made regarding adherence to the QI measures? They are listed in a table, but it is unclear to me how the determination of yes/no is made.

Author Response:

Thank you for pointing this out. We now provide additional information on the tele-medical rounds in the intervention condition and on determination of QI adherence on the patient level.

In the description of the intervention condition (first item), we amended the text as follows (page 12):

"*Each round entails the discussion of all QIs (Table 1) in a step-by-step approach discussing, for example, sedation, ventilator settings and antibiotics. QI-related criteria are subsequently assessed based on information obtained and documented by the tele-ICU consultant during daily tele-medical ward rounds and on medical records at the remote-ICU. Beyond mere QI assessment, the tele-ICU consultant can provide additional advice concerning treatment plans to remote-ICU physicians.*

Further, in Table 1, we added a description of criteria for each QI to be fulfilled (QI adherence yes vs. no).

Comment:

The secondary outcomes may be presented in a table to summarize them more concisely. That would be visually easier than the long bulleted list.

Author Response:

As suggested, the long bulleted list of secondary outcomes (several of these core outcomes of PICS) together with time points, measurement instruments and corresponding references was transferred to a new Table 3.

Comment:

In the follow-up procedures, please specify whether all the clinicians involved "on-demand counseling sessions between the treating GP or rehab facility and investigators of the Charité" will be blinded to the study groups. If the investigators are able to become aware of the study group, there may be a perceive risk of introducing systematic differences in these counseling sessions.

Author Response:

We agree with the reviewer that the follow-up procedures and the "on-demand counseling sessions between the treating GP and the investigators" might lead to systematic differences. However, the blinding potential is limited as the patient can tell the counseling physician about the telemedical round, and the round itself is documented in the EMR of the patient. This might be a potential risk of

bias and will discuss this in the discussion of the results after analyzing the data. We added the following sentence to the paragraph 'Follow-up procedures and post-acute care':

> "*Investigators of Charité cannot be blinded regarding the treatment condition, as they need to assess the electronic medical record data, which might be a potential source of bias*."

Comment:

In the Process Evaluation section (Page 22/87), could the authors provide more details about intervention monitoring, how they will ensure that rounding is occurring as expected, what standards will be used to assess the adherence to the rounding intervention, and what, if anything, will be done to improve rounding if it is not occurring as planned.

Author Response:

The reviewer makes an important point. We are aware that we use a rather pragmatic approach monitoring selected aspects of the intervention and no in-depth implementation monitoring. To give more details on measures ensuring that the intervention is delivered as planned, we added the following sentences to the subsection on process evaluation:

> *"For tele-medical ward rounds (QI visits), data on the connection failure rate and overall connection quality between the tele-ICU and remote-ICUs are closely monitored to avoid inadequate implementation of the intervention."*

Related information is already included in the subsection 'Adherence to the intervention' (page 13).

Comment:

Table 1 - it may be useful to incorporate into this table some details regarding what would qualify as adherence to each QI measure. Is it at the discretion of the rounding team, the tele-ICU clinicians, or are there pre-specified definitions for each? If these details are elsewhere in the narrative, they are buried, and it would be useful to be able to access this information quickly, such as from a table.

Author Response:

We thank the reviewer for this important concern and agree that the manuscript requires a description of criteria used to assess the adherence of a single QI at a more prominent position. In line with your comment no. 5, we added some details of these QI-related criteria in an additional column of Table 1 (primary outcomes).

Comment:

Table 2: Please clarify what is meant by "(x)" in some of the cells of the table.

Author Response:

Depending on the time of day of a patient's ICU admission or discharge, his/her outcomes might not be documented at this day. This has been clarified by adding an additional footnote to the SPIRIT-Table 2:

> (X) depicts an optional measurement.

Reply to Reviewer 2:

(Sikandar Khan, Indiana University School of Medicine, U.S.)

Comment:

This is a very well written protocol. I would recommend the following clarifications:

- The introduction focuses on PICS and ICU survivorship, but the primary outcome of the trial appears to be quality improvement in the ICU. More attention to ICU quality metrics and literature should be included in the introduction.

Author Response:

We thank the reviewer for this comment. We acknowledge that the introduction had an emphasis on PICS and functional ICU outcomes. We strongly believe that these outcomes are of utmost importance to patients, but admit that they are secondary and not the primary outcomes of our study. Taking your comment into consideration, we shortened the section on ICU survivorship for the benefit of additional details on quality indicators and quality metrics. More specifically, we added the following text:

> "*In QI development, areas of clinical practice need to be identified where evidence-based, best practice diverts from what is actually delivered to patients. Then, measurable key performance indicators are defined, serving as surrogates for the level of implementation of the respective single QI. These must be accepted by the clinical team and, in the next step, implemented in daily practice to ultimately improve patient outcomes.[18] As such, QIs can be used to align allocation of healthcare resources for the improvement of patient outcomes within acute care as well as in the post-ICU period."*

Comment:

- I would provide more detail on the 8 QI domains and how they are operationalized in German ICUs, this will be helpful for the international audience.

Author Response:

Thank you for this suggestion which is in line with comments no. 5 and 9 of Reviewer 1. For further clarification, we now provide a description of the operationalization for the assessed QIs in Table 1 (new third column included). Besides, we added further the text to the section 'Intervention condition: telemedicine' (page 12) to explain what the ward rounds entail and related processes concerning QI assessments.

For a detailed description on the raw QI-related parameters we have to refer to the original article [Kumpf O, et al. 2017] mentioned several times in the manuscript as it is not possible to list and explain all of them in Table 1 and Table 2 or the manuscript text.

Comment:

- I would address contamination in a bit more detail; do physicians or other staff work at more than 1 of the participating hospitals?

Author Response:

We agree with the reviewer's comment that this issue requires further clarification. To allow determination of the risk of within cluster contamination in our trial, we included the following sentence in the subsection on 'Site selection' (page 9/10):

> "*The defined randomization units are geographically and organisationally separated which prevents workforce movements between clusters, and thus, intervention contamination*."

Furthermore, we monitor and ensure that none of the staff members involved in the intervention works at more than one cluster which minimizes the risk for any contamination between clusters caused by staff turnover.

Comment:

- Please provide additional details on how adherence of the individual components will be ascertained, and whether it is entirely based on review of the medical records?

Author Response:

We agree that this point requires further clarification.

For every single QI, the associated criteria, i.e. QI-related performance parameters to be assessed to derive the binary outcomes (the patient's QIs are fulfilled yes or no, at a certain day), are routinely documented in the patient's medical records. Whether a single QI for patient *i* on day *t* is fulfilled or not is subsequently assessed by an intensive care specialist at the tele-ICU cockpit rating the underlying electronicI documentation recorded by the remote-site physician. This central endpoint adjudication process is utilized irrespective of whether the ICU delivers care on control or intervention condition.

In the subsection on the intervention condition (page 12), we now provide more information on data sources for determination of QI adherence (also see our response to your second comment):

> "QI-related criteria are subsequently assessed based on information obtained and documented by the *tele*-ICU consultant during daily tele-medical ward rounds and on medical records at the *remote*-ICU.

Similarly, we give more details on data sources in the subsection about the control condition:

> "*… assessment of QI adherence in the control condition is based on examination of medical records regarding QI-related parameters."*

Comment:

- How will missing data be handled?

Author Response:

In the ERIC trial, missing outcome data means missing QI documentation or evaluation at specific ICU days of a patient. We assume missingness at random (MAR) for the principal analyses of binary primary efficacy outcomes. A longitudinal random effects model incorporates this approach. In this case, multiple imputation techniques are not needed [White IR et al. 2012; O'Kelly M 2017].

In sensitivity analyses we will also apply approaches to study informative missingness of QIs (depending on disease severity and length of ICU stay).

As mentioned in a response to reviewer 1: In order to provide objective and pre-planned statistical analyses (clinical effectiveness), the statistical analysis plan (SAP) will be written and published by the trial statistician before the final analyses will be conducted.

White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials 2012;9: 396-407.

O'Kelly M, Ratitch B. Clinical trials with missing data: a guide for practitioners. John Wiley & Sons 2014.

Comment:

- The introduction could be strengthened by discussing how increased completion of the QI domains would affect the secondary outcomes of the study.

Author Response:

We thank the reviewer for this comment. To give more insights on the causal link between QI adherence and the secondary, patient-centered functional outcomes, the introduction was modified. Further, we added the following paragraph to the introduction (page 6):

> "*This set of ten QIs reflects consensus-based strategies to improve quality of acute ICU care and, thus, is supposed to reduce PICS/CCI burden. For instance, if patients are frequently assessed and consequently treated for delirium, cognitive outcomes might improve.[4]…*".

While PICS is a patient-centered outcome (based on multiple core domains), the ERIC trial is motivated by the idea that a better quality at ICU improves PICS-related patient-

centered outcomes defined as secondary outcomes. It is not the goal of the trial to prove that a quality improvement intervention (QII) improves the PICS-related outcomes, but rather that QII results in measurable improvements of established quality indicators. This is considered as an essential intermediate step on the pathway between QII and improvement in PICS-related patient outcomes.

Comment:
- Is the primary outcome measured at the patient level or the cluster/institution level? Would provide more detail on whether "beneficial effects" are simply increased completion rate of the domain?
Author Response:
All eight primary efficacy outcomes are measured on the patient level and on a daily basis. It is hypothesized that a quality improvement intervention (QII) is beneficial for the patient. The trial tries to prove that a beneficial effect of the QII can be measured. This is the primary goal of ERIC. The corresponding effects on the PICS-related domains are secondary outcomes. The resulting data will help to design a future trial with the PICS as primary efficacy endpoint (binary outcome or composite outcome/scale for PICS domains) which will be a very complex and methodologically challenging undertaking.
The modified Table 1 gives an overview of all 10 consensus-based German QIs, eight of these were defined as primary outcomes for the ERIC study. QI V and QI X (numbering according to the original article) were not specified as primary efficacy outcomes since they are assessed on the institutional level and also not on a daily basis; therefore QI V and QI X would not enable estimation of an interpretable effect of the intervention (see footnote of Table 1).
We think this issue is further clarified by the revised introduction section.

Comment:
- Does randomization include use of block design? Also would explain whether randomization is at the cluster level, since the authors also include eligibility criteria at the patient level.
Author Response:
ERIC is a pragmatic, cluster-randomized study with a stepped wedge design. Hence, randomization to one of three possible sequence groups (defining the crossover date for switching to the intervention condition) was conducted at the hospital site (cluster) level, and this design choice is clearly justified. Randomization units were 12 eligible clusters, i.e. medical facilities, which provided a letter-of-intent prior to trial commencement. To generate the randomization list to randomly allocate clusters to the three sequence groups, a block design was applied (block size = 3). No stratified or covariate constrained allocation method was used.
The former aspect was included in the subsection on ‘randomization, allocation concealment and masking’. We believe that design issues for cluster-randomized trials are sufficiently explained in the subsection on randomization. In particular by the sentence “*Due to the stepped wedge design, hospitals defined as clusters are randomized….*”
More details on the randomization method (in particular, the block size) will be reported in the final publication following the CONSORT SW-CRT reporting guideline.

Broad eligibility criteria on the patient-level were defined to recruit participants. Every adult admitted to the participating randomization unit (i.e. ICU of a participating medical facility) had to fulfil only very few inclusion criteria: coverage by a German statutory health insurance; expected to be at ICU >= 48 h; written informed consent.

Comment:

- The authors note under trial status that recruitment will end this year (October 2020). It would be helpful to clarify if the protocol has changed during the trial.

Author Response:

As stated in the Declaration section 'Ethics approval' (and in the Supplementary File 1), this protocol manuscript presents the protocol version 1.1, dated 27.05.2019.

One amendment was submitted to the ethics committee in May 2019:

The principle structure of the trial was not changed (cluster randomization, population, interventions, endpoints). There have been sponsor-related issues. At the beginning, only patients inscribed in a specific healthcare plan have been recruited and enrolled into the trial. This restriction was removed in order to enhance recruitment. The trial amendment addressed organizational issues related to this adaption. We decided to mention this aspect on the first page of the supplementary file 2 [Study Group] while explaining the ERIC consortium.

Reply to Reviewer 3:

(Sue Lasiter, University of Missouri, Kansas City, U.S.)

Comment:

The study protocol is well written and complete.

Author Response:

Thank you very much for your extremely positive appraisal of our manuscript.

In the revised version of our manuscript a small number of additional minor edits were made, all indicated in track changes.

Concerning the formatting of amendments, the original Supplementary file 4 [SPIRIT checklist] which we missed to cite in the main text is now referred right at the beginning of the Methods Section as Supplementary file 1. Therefore, the numbering of all the other Supplementary files (including the file names) was changed accordingly.

Due to the inclusion of an additional Table 3 for secondary outcomes positioned at the end of the document, associated references changed accordingly.

We are grateful to the editor and the three referees for their careful evaluation of our submission, their valuable comments and their helpful suggestions, which we believe contributed to a significant improvement of the quality of our manuscript.

VERSION 2 – REVIEW

REVIEWER	Meeta Prasad Kerlin University of Pennsylvania, U.S.A.
REVIEW RETURNED	18-Jun-2020

GENERAL COMMENTS	Thank you for the opportunity to review a revised version of this manuscript. The authors have been highly responsive to my and the other reviewers' suggestions and I have no further comments at this time.