Implications for post critical illness trial design: sub-phenotyping trajectories of functional recovery among sepsis survivors

Zudin A. Puthucheary[1,2], Jochen S. Gensichen[3,4,11], Aylin S. Cakiroglu[5], Richard Cashmore[2], Lara Edbrooke[6,10], Christoph Heintze[7], Konrad Neumann[8], Tobias Wollersheim[9], Linda Denehy[6,10] and Konrad F.R. Schmidt[3,7,11]

Corresponding author

Dr Zudin Puthucheary, Critical Care and Perioperative Medicine Research Group,

Adult Critical Care Unit, Royal London Hospital, London, E1 1BB, United Kingdom

Email: z.puthucheary@qmul.ac.uk

Tel: +44(0) 203 594 0351

Abstract

Background

Patients who survive critical illness suffer from significant physical disability. The impact of rehabilitation strategies on Health-Related Quality of Life (HRQoL) is inconsistent, with population heterogeneity cited as one potential confounder. This secondary analysis aimed to examine trajectories of functional recovery in critically ill patients to delineate sub-phenotypes; examine the distinguishing clinical characteristics between these cohorts and assess differences in clinimetric properties of assessment tools of physical function between cohorts.

Methods

291 adult sepsis survivors were followed up for 24 months by telephone interviews. Physical function was assessed using the Physical Component Score (PCS) of the Short Form-36 Questionnaire (SF-36), Activities of Daily Living (ADL) and the Extra Short Musculoskeletal Function Assessment regarding physical function and disability (XSFMA-F/B). Longitudinal trajectories were clustered by factor analysis. Logistical regression analyses were applied to patient characteristics potentially determining cluster allocation. Responsiveness, floor and ceiling effects and concurrent validity were assessed within clusters.

Results

159 patients completed 24 months follow-up, presenting overall low PCS-scores. Two distinct sub-cohorts were identified, exhibiting complete recovery or persistent impairment. A third sub-cohort could not be classified into either trajectory. Age, education level and

number of co-morbidities were independent determinants of poor recovery (AUROC 0.743 ((95%CI 0.659-0.826); p<0.001). Those with complete recovery trajectories demonstrated high levels of ceiling effects in Physical Function (15%), Role Physical (45%) and Body Pain (57%) domains. Those with persistent impairment demonstrated high levels of floor effects in the same domains: Physical Function (21%), Role Physical (71%) and Bodily Pain (12%). The Physical Function domain of the SF-36 demonstrated high responsiveness between ICU discharge and at 6 months was predictive of a trajectory of persistent impairment (AUROC 0.859 (95%CI 0.804-0.914); p<0.001).

Conclusions

Within sepsis survivors, two distinct recovery trajectories of physical recovery were demonstrated. Older patient with more co-morbidities and lower educational achievements were more likely to have a persistent physical impairment trajectory.

In regard to trajectory prediction, the Physical Function score of the SF-36 was more responsive than the Physical Component Score and could be considered for primary outcomes. Future trials should consider adaptive trial designs that can deal with non-responders or sub-cohort specific outcome measures more effectively.

Keywords

Sepsis, Post intensive care Syndrome (PICS), physical function, Health-Related Quality of Life (HRQoL), Patient reported outcome measures (PROMS), co-morbidity.

Background

Increasing numbers of patients are successfully surviving critical illness. Unfortunately, residual functional and/or mental disabilities affect many critical care survivors after hospital discharge [1, 2]. Despite extensive research into rehabilitation strategies, few studies have been able to demonstrate a positive effect on this ensuing dysfunction or improve Health-Related Quality of Life (HRQoL) [3-6]. Given that rehabilitation strategies have a strong evidence base in other patient populations [7], trial-related methodological issues have been proposed as a source of influence in this area and examined [8, 9].

Population heterogeneity within the critically ill cohort is one area that may hinder current outcome analysis. Certain specific patient characteristics have already been identified as influential in regards to an individuals' subsequent HRQoL outcome. To date, these include, age [10], pre-critical illness comorbidity [11], and socioeconomic-status [12]. Severity of critical illness, Intensive Care Unit (ICU) Length of stay and the effect of within-ICU physiology remain unclear influences, as does sex [10, 11, 13-16]. If these factors are not accounted for in trial design, patient stratification, or analysis, outcome data may be unintentionally skewed. Many of the current outcome assessments for trials in critical care fail to account for these confounders [15, 17]. Patient reported outcome measures are increasingly prioritised as endpoints [18-20]. The Physical Component Score (PCS) of the Short Form-36 Questionnaire (SF-36) is used to demonstrate the physical disability of critical care survivors [21], and is widely reported in rehabilitation trials.

Several re-analyses have demonstrated sub-phenotypes based on recovery trajectories [9, 15, 22]. How these sub-phenotypes respond to the variety of assessments that measure HRQoL currently in use is not yet defined. It may be that these assessments, often applied as outcome measures, have different clinimetric properties within patient sub-populations.

Understanding this aspect of measurement in addition to recovery trajectories will be important to future trial design and outcome interpretation.

We performed a secondary analysis of a critical care trial of sepsis survivors using two-year follow-up data [23]. The aim of this was to i) examine the trajectories of functional recovery in critically ill patients using an agnostic approach to delineate patient sub-phenotypes; ii) examine the distinguishing clinical characteristics between these cohorts and iii) assess the differences in clinimetric properties of assessment tools of physical function between cohorts.

Methods

The patient cohort comprised of those recruited to a randomised control trial conducted between February 2011 and December 2015 evaluating a primary care-based sepsis aftercare intervention [23] [24]. Two hundred and ninety-one adult survivors of sepsis were recruited from nine centres across Germany. Trial design, methodology and outcomes are described in detail in the original manuscript [23, 25]. Briefly, trained study nurses collected baseline data at in-person interviews while participants were still hospitalized. Follow-up data pertaining to HRQoL and physical function were collected at 6 months, 12 months and 24 months by telephone interviews. Those instruments specific to this analysis were the Physical Component Score (PCS) of the SF-36 [26], three of its four subdomains (Physical Function, Role Physical and Body Pain), activities of daily living (ADL) and the Extra Short Musculoskeletal Function Assessment regarding physical function and disability (XSFMA-F/B) [27]. This extra short questionnaire is derived from the 101-item Musculoskeletal Function Assessment (MFA) by Engelberg and al. to assess functional status from the patient's perspective [28]. It has been mainly used in Germany for patients following

orthopaedic surgery [27]. Functional outcome data were also analysed for sub-phenotype concurrent validity and clinimetric properties. Both randomisation groups were included into analyses, as no effects of the intervention were shown regarding functional or HRQoL outcomes [23]. Only those with complete data sets (all four time points) were used in this analysis.

Education and Family status classifications are shown in Additional Table 1 and addressed domains of instruments used in Additional Table 1.1.

Trajectory Projection cluster analysis

Groups of longitudinal trajectories of Physical Component Scores of the SF-36 (the most commonly reported 6-month HRQoL outcome measure [3, 6, 29-34]) were clustered using the R-package TRAJ [35-37] and applied. Briefly, this package implements a 3-step procedure [36]. Firstly, 24 summary measures (available in Additional Table 2) are calculated that measure the features of trajectories. These measures were then analysed using factor analysis to select those that best describe the main features of trajectories. Lastly, using these factors the trajectories were clustered.

General statistical analysis

Continuous data were assessed for normality using D'Agostino and Pearson omnibus normality tests and analysed using paired two-tailed Student's t-test or Mann Whitney U test as appropriate. Normally distributed data were described using mean (95% Confidence Interval) and non-normally distributed data as median (interquartile range). Categorical variables were analysed by $\chi 2$ testing. Multivariable and univariable logistic regression analyses were applied to variables potentially determining cluster allocation (dependent

variable) Unclustered participants were not used in the logistical analysis, and a multinomial regression performed as a sensitivity analysis. Independent variables were determined as characteristics (Table 1), with a univariable screening threshold set at $p<0.10$. Significance for all other tests was set at $p<0.05$. Area under the Receiver-Operator-Curve was used to test the predictive capacity of early ICU discharge and 6 months assessments for persistent functional impairment.

Floor and Ceiling Effects

Scores at their lowest point are defined as 'floor effects' and a 'ceiling effect' occurs where patients 'may show no improvement in function if a functional scale is not able to assess high level instrumental ADLs (a ceiling effect) [38, 39]. Floor and ceiling effects render a measure unable to discriminate between participants at either extreme of the scale. This negatively affects measurement properties, including sample size requirements. Reducing these effects by choice of the right measure can therefore improve study efficiency [40].Floor effects were calculated as the percentage of participants scoring the worst possible score for the measure. Ceiling effects were calculated as the percentage of participants scoring the best possible score for the measure. Components of the SF-36 were examined at the differing time points for floor and ceiling effects, for the cohort as a whole and for the individual clusters. Floor and ceiling effects were considered relevant if >15% of the participants had the highest or lowest score respectively [41].

Concurrent validity

Concurrent validity is a measure of how well a test compares to a gold standard (such as the PCS) [38] and its substitutability. Therefore, it is a component of criterion validity, an

156 estimate of accuracy based on an external criterion [42]. Coefficient of Determination from
157 regression between parameters was used to measure concurrent validity (the degree to
158 which a test can be used as a substitute measure for the gold standard) between the PCS
159 and PF of the SF-36, ADLs and XSFMA-F/B. All coefficients were interpreted as: little (0.00-
160 0.25), fair (0.25-0.50), moderate (0.50-0.75) and excellent association (0.75-1.0) [43].
161
162 *Responsiveness*
163 Responsiveness is a measure of sensitivity to change and discriminatory properties (the
164 ability to detect clinically relevant change in health status over time), and part of the
165 COSMIN checklist (COnsensus-based Standards for the selection of health
166 Measurement)[42, 44, 45]. Change in scores from hospital discharge to 24 months were
167 assessed using paired t-tests and data represented as mean difference and 95% CI [43].
168 Responsiveness of each test to time/recovery post critical illness was calculated using the
169 effect size index, calculated as the mean change score divided by the baseline pooled
170 standard deviation [38, 46]. Changes were interpreted according to Cohen's d effect size as
171 small (0.2 to 0.49), moderate (0.5 to 0.79) and large (>0.80) [47, 48].
172
173 **Results**
174 Of the original 291 participants recruited, 24-month follow-up data was collected on 186
175 participants (41 lost to follow-up, 64 died <24 months). Complete data was available on 159
176 participants who were included in the final analyses. Those with incomplete follow-up were
177 not included. When compared, those who died were older, had a longer length of stay and
178 more co-morbidities, all of which is not unexpected (see Additional Table 3).

PCS of the SF-36 for critically ill participants were reduced relative to population norms at ICU discharge and remained low at 24 months (Figure 1A).

(insert Figure 1)

Trajectory Clustering

Trajectory projection analysis identified two distinct sub-cohorts: one cohort exhibited a faster and more complete recovery trajectory defined as within one standard deviation of population norms (n=61). A second cohort exhibited more persistent functional impairment (n=76) (Figure 1B). The remaining 22 participants were not classified into either cohort, as no clear trajectory was seen (Additional Figure 2). The differing characteristics of the cohorts are shown in Table 1.

(insert Table 1)

The complete recovery cohort, were on average younger (56 years (IQR 43-70) vs. 65 years (IQR 54-72), P=0.002, Figure 2A), with higher education levels (5(4-8) vs. 5(3-5), P= 0.039, Figure 2B), more likely to be unmarried (Figure 2D) and had a lower BMI (25.8(22-29) vs. 27.8(24-32), P=0.006.

(insert Figure 2)

A multivariable logistic regression analysis demonstrated age, education level and number of co-morbidities as independent determinants of poor recovery (Additional Table 4). A

model with these factors had a predictive capacity with an AUROC of 0.743 ((95%CI 0.659-0.826); p<0.001; Additional Figure 1) for cohort membership and was not over-fitted (Hosmer-Lemeshow statistic 8.456, p=0.390). Neither Body Mass Index nor Family Status at discharge were significant within this analysis. In a multinomial analysis, age and education remained independent determinants of recovery with the addition of Body Mass Index (Additional table 4.1) but not number of co-morbidities (p=0.051). No determinants were independently associated with the unclustered trajectory (see Additional Table 4.2).

Floor and Ceiling effects

At 24-month follow up, participants in the completed recovery cohort demonstrated relevant ceiling effects within the Physical Function (15%), Role Physical (45%) and Body Pain (57%) domains of the SF-36. In contrast, those participants with persistent functional disability demonstrated the reverse, with relevant floor effects within Physical Function (21%), Role Physical (71%) but not Bodily Pain (12%), see Table 2 and Figure 3. These results were relatively consistent over the preceding 24 months (Additional Tables 5A and B). Floor scores at ICU discharge were only moderately associated with a persistent functional impairment trajectory (PF (AUROC 0.609 (95%CI 0.537-0.681); p=0.002) and RP (AUROC 0.653 (95%CI 0.584-0.721); p<0.001)). However, floor scores at 6 months were good predictors of a trajectory of persistent functional impairment (RP (AUROC 0.586 (95%CI 0.513-0.658); p=0.014)), and PF (AUROC 0.938 (95%CI 0.901- 0.974); p<0.001)).

(insert Table 2)

Concurrent validity

Those participants with complete recovery demonstrated moderate to excellent concurrent validity between SF-36 PCS and both XSFMA-B AND XSFMA-F, and fair validity with ADL scores. Those participants with persistent disability demonstrated moderate concurrent validity between SF-36 PCS and both XSFMA-B AND XSFMA-F, and fair validity with ADL scores (Table 3).

(insert Table 3)

Responsiveness

High responsiveness was seen in the complete recovery group at all time points in the Physical Component Score (>1.0) and most notably in the Physical Function domain (>1.6), with a similar pattern seen in Role Physical. However, this was not seen in the persistent impairment cohort, where Physical Function and Role Physical achieved only moderate responsiveness at 6 months (>0.7). All other scores and time points demonstrated at best limited responsiveness (Table 4). PF responsiveness between ICU discharge and 6 months was predictive of a trajectory of persistent impairment (AUROC 0.859 (95%CI 0.804-0.914); $p<0.001$).

(insert Table 4)

Discussion

This post-hoc study examines the trajectories of functional impairment in cohorts of sepsis survivors regarding sub-phenotypes and specific clinical characteristics.

Two distinct sub-cohorts were identified: one of faster and more complete recovery and the other of slower recovery with more persistent functional impairment. A third sub-cohort could not be classified into either trajectory. This study also demonstrates that the older patient with more co-morbidities and with lower educational achievements is more likely to have a trajectory associated with persistent functional impairment. Importantly the measures used exhibit very different clinimetric properties when HRQoL is measured longitudinally in different sub-cohorts. Those with good recovery have significant ceiling effects with the physical components of the SF-36 questionnaire and demonstrate high responsiveness over time. The reverse is seen in those with persistent impaired HRQoL, where significant floor effects are seen and limited responsiveness. Moderate to excellent concurrent validity was obtained across tests of HRQoL and physical function. The Physical Function (PF) score had the highest degrees of responsiveness across sub-cohorts and time and was predictive of a trajectory of persistent impairment when measured up to 6 months. Scoring the lowest value of PF at 6 months also was predictive of poorer outcomes at 24 months, which might be an indicator for the necessity to develop individualized rehabilitation programs for every patient.

Individual Patient Characteristics

These data reiterate the role that age and multiple chronic diseases have on recovery of physical HRQoL post critical illness. Interestingly, the individual odds ratios for these factors are lower than that of educational status. This may be because educational status is

reflective of poorly quantified and measured socioeconomic factors as well as individual coping abilities that are essential for the rehabilitation process [49]. However, chronological age is increasingly recognised as less accurate in terms of function relative to physiological age in the elderly [50], and the Charlston Co-morbidity Index was not designed or validated for the critical care survivor population. Ultimately these data demonstrate that stratification (or population enrichment strategies) on one or two of these variables are unlikely to be sufficient. We have begun to understand how frailty, cognitive deficits [51], comorbidities [9], age and ICU length of stay [22, 52] interact to result in post-critical illness disability, and our data confirm these findings but also suggest that these factors need to be integrated with socioeconomic data for improved identification of sub-phenotypes. The impact of social isolation is reported in other chronic diseases and needs more attention in critical illness populations [12].

Physical Function and Health Related Quality of Life outcome measures

The use of HRQoL and patient reported outcome measures are important and increasingly mandated, and the data reported here may help to focus the field on the appropriateness of the specific domains of the SF-36 to measure HRQoL in different subpopulations with different illness trajectories. The PCS has been used as a primary outcome measure in rehabilitation trials [6, 29], in nutrition intervention trials [53] and is in general the most commonly reported 6-month HRQoL outcome measure [3, 6, 29-34]. The PF subscore has also been used as a primary outcome measure in critical illness [54]. Fundamentally, selection of an outcome measure assumes that the intervention is suitably designed with the primary outcome in mind. When evaluating rehabilitation trials if the primary outcome of a trial is health-related quality of life, then using the summative score (PCS, incorporating

all subdomains to reflect overall health-related quality of life) would be appropriate. In contrast, if the primary outcome is physical function, then it may be more appropriate to select the Physical Function subdomain as the measure used to evaluate the trial. It should be noted that HRQoL outcome measures have often been shown to not be sensitive enough to be affected by the biological efficacy of current post ICU interventions [63].

To date, little exploration of the most sensitive component of the SF-36 to use in trials of rehabilitation interventions has been conducted [55]. Physical and mental health factors account for 80-85% of the reliable variance in the 8 scales of the SF-36 [56]. A scoring assumption central to the summative scores (i.e. PCS and MCS) is that score aggregation could occur without score standardization or item weighing [57]. Our data challenge this assumption: in the presence of significant heterogeneity of physical HRQoL and disability post critical illness, individual domains are more appropriate outcome measures than summative scores for physical rehabilitation trials, given the responsiveness and predictive outcomes seen across patient sub-phenotypes. Of note the PF score has long been known to be the most valid scale for physical activity [58] and our data demonstrate that aggregating PF with the other components of the PCS decreases the clinimetric strength. The PF domain includes questions related to activities needed for daily living rather than also including return to work and questions about pain as found in the PCS. The PF domain includes several advanced mobility measures, independent activities of daily living, some activities of daily living as well as several items of the XSFMA, which may explain the concurrent validity findings, as this may be better viewed as construct validity. It may be that in the post critical illness population there is a more specific objective perception of physical function (the PF score, comprising of 10 questions), resulting in higher responsiveness than broader

subjective limitations in daily life (the RP score, comprising of 4 questions, or General Health comprising of 5 questions) or perception of pain (the BP score, comprising of 2 questions). However, the PF score also has significant ceiling effects (in those that recover) and floor effects (in those with persistent disability), suggesting the need for concurrent measurement of other more specific outcome measures such as the XSFMA-F which showed excellent validity with the SF-36 PF to address this. Notably, using the PF domain score at 6 months can predict poorer physical HRQoL outcomes and may help to guide further community or out-patient based individualised rehabilitation treatment.

Strengths and limitations

A major strength of these analyses are the data themselves- few long-term cohort studies exist with serial contemporaneous HRQoL and physical function data to allow detailed clinimetric testing of outcome measures. The cohort size was large relative to other long term cohort studies with serial contemporaneous HRQoL and physical function data. It is widely accepted, and accords with common sense, that the imputation of missing data on HRQoL for a deceased participant is inappropriate [59]. This is in keeping with approaches applied to randomised controlled trials [60] and is an approach used by others (with specific expertise in imputation) within the field of rehabilitation [59, 61]. This would also be consistent with analyses applied to this cohort which we have recently published [24]. Those patients who died were older, had a longer length of stay and more co-morbidities, and a 2-year follow-up period may not be appropriate for this sub-cohort.

A fundamental issue with clinimetric property assessment of summed scores like the PCS is the content overlap [57], as the used subscores are in part textual identical with the

summed score and there also was a high contentual intersection with the XSFMA-F/B and ADL scores. This is difficult to overcome, as the PCS is near ubiquitous in its use for measurement of physical HRQoL. The use of trajectory clustering techniques decreased the risk of bias relative to a researcher driven approach. The retrospective nature of this analysis mandates that the conclusions are tested prospectively. Trajectory cluster validity is limited by 22 (13.8%) of patients being not classifiable and understanding why these patients have unclear trajectories requires prospective analysis, using a mixed-methods approach. The XSFMA F/B scores have only been validated in German, limiting its use, though it was derived from the English SFMA [62]. Other tools such as the Functional Status Score for the Intensive Care Unit (FSS_ICU) or the Physical Function in Intensive Care Test scored (PFIT-s) may be of use, having been validated in several countries and languages [35]. While the focus of this manuscript has been on self-reported outcome measures, the subjective nature of these does constitute a limitation and comparative assessment with objective measures in sub-cohorts may be warranted.

Implications for outcome selection and trial design

As HRQoL outcome measures have often shown lack of sensitivity in post ICU interventions [63], our data offers two potential methodological solutions: Firstly, the described sub population characteristics, especially those relating to education could be used as population refinement tools for trials, either as inclusion/ exclusion criteria or for differential outcome measures set a priori. This may or may not be feasible where large samples are required, though a differential effect between sub populations has been used in phase II trials (NCT02358512). Secondly an adaptive trial design could use a) the presence of a floor effect as a predictor of a poor trajectory (i.e. a non-responder) in a multi-arm, multi-

stage fashion that explores treatments, doses with an option to exclude non-responders [64]; b) the characteristics (e.g. education or socioeconomic status) for population enrichment that narrow down recruitment to those who are likely to benefit most [65] or c) the PF score in conjunction with other markers e.g. CRP (as a marker of persistent inflammation) in a biomarker adaptive design [66] to stratify patients. Lack of data to inform adaptive trial design remains one of the barriers to their use and this study offers suggestions to overcome this [67].

Both subscore and summary score responsiveness varied over time in both cohorts, with a plateau seen after 6 months. These data imply that physical HRQoL endpoints may be more suited to earlier timepoints (e.g. 3 and 6 months), and other, more responsive endpoints are needed at 1-2 years such as measures of disability.

Conclusion

Within sepsis survivors, two distinct recovery trajectories of physical recovery could be demonstrated. Older patient with more co-morbidities and lower educational achievements are more likely to have a trajectory associated with persistent physical impairment. In regard to trajectory prediction, the Physical Function score of the SF-36 was more responsive than the Physical Component Score of the SF-36 and could be considered for primary outcomes. Future trials should consider adaptive trial designs that can deal with non-responders or sub-cohort specific outcome measures more effectively.

List of abbreviations

ADL: activities of daily living

BP: Body Pain

GH: General Health

HRQoL: Health-Related Quality of Life

ICU: Intensive Care Unit

PCS: Physical Component Score

PF: Physical Function

PROMS: Patient reported outcome measures

RP: Role Physical

SF-36: Short Form-36 Questionnaire

XSFMA-F/B: Extra Short Form Musculoskeletal Function Assessment regarding physical function (F) and disability (B)

Declarations

Ethics approval

The study protocol of the SMOOTH-Study was approved by the institutional review board of the University of Jena, 26 January 2011 (No.3001/111).

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

Zudin A. Puthucheary, Jochen S. Gensichen, Aylin S. Cakiroglu, Richard Cashmore, Lara Edbrooke, Christoph Heintze, Konrad Neumann, Tobias Wollersheim, Linda Denehy and Konrad F.R. Schmidt declare that they have no conflict of interest.

Funding

The SMOOTH-Study was supported by the Center for Sepsis Control and Care (CSCC) at Jena University Hospital, Germany, funded by the German Federal Ministry of Education and Research (grant 01 E0 1002) and the German Sepsis Society.

AC was was supported by the Francis Crick Institute (which receives its core funding from Cancer Research UK (FC010110), the UK Medical Research Council (FC010110), and the Wellcome Trust (FC010110)).

Authors' contributions

Study concept and design: ZP, KS

Data acquisition: KS, JG, ChH

Analysis of data: ZP, AC, KN

Interpretation of data and drafting of the manuscript: ZP, AC, RC, LE, ChH, KN, TW, LD, KS

Critical revision and approval of the manuscript: ZP, AC, RC, LE, ChH, KN, TW, LD, KS

Acknowledgements

We would like to thank Michelle Kaufmann (Jena University Hospital) for editorial support and the SMOOTH Study group for patient recruitment, data collection and analysis within the SMOOTH-Study: Anne Bindara-Klippel, Fiene Eissler, Carolin Fleischmann, MD, Ursula Jakobi, MD, Susan Kerth, Heike Kuhnsch, Friederike Mueller, MD, Konrad Reinhart, MD, Mercedes Schelle, André Scherag, MSc PhD, Nico Schneider MSc, Paul Thiel, MD, Susanne Worrack MSc (Jena University Hospital); Andrea Geist, Torsten Schreiber, MD (Zentralklinik Bad Berka); Christian Berhold, MD, Marcel Corea, MD, Adrian Freytag, MD, Herwig Gerlach, MD, MBA, Rainer Kuehnemund, MD, Josefa Lehmke, MD, Peter Lehmkuhl, MD, Lorenz Reil, MD, Guenter Tiedemann, MD, Susanne Toussaint, MD (Vivantes Clinics Berlin); Anton Goldmann, MD, Michael Oppert, MD, Didier Keh, MD, Sybille Rademacher, MD, Claudia Spies, MD, Lars Toepfer, MD (Charite University Medicine Berlin); Frank Klefisch, MD (Paulinen Hospital Berlin); Armin Sablotzki, MD (St Georgs Hospital Leipzig); Frank

Oehmichen, MD, Marcus Pohl, MD (Bavaria Clinic Kreischa); Andreas Meier-Hellmann, MD (Helios Clinic Erfurt); Farsin Hamzei, MD (Moritz-Clinic Bad Klosterlausnitz); Christoph Engel, MD, Christine Pausch, PhD (University of Leipzig); Gustav Schelling, MD (University Clinic LMU Munich).

Author details

1. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK
2. Adult Critical Care Unit, Royal London Hospital, London, UK
3. Institute of General Practice and Family Medicine, Jena University Hospital, Germany
4. Institute of Family Medicine, University Hospital of the Ludwig-Maximilians-University, Germany
5. The Francis Crick Institute, London, UK
6. Physiotherapy Department, the University of Melbourne, Australia
7. Institute of General Practice and Family Medicine, Charité University Medicine Berlin, Germany
8. Institute of Biometry and Clinical Epidemiology, Charité University Medicine Berlin, Germany
9. Department of Anesthesiology and Operative Intensive Care Medicine, Charité University Medicine Berlin, corporate member of Freie Universität Berlin, Humboldt Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany
10. Allied Health Department, Peter MacCallum Cancer Centre, Australia
11. Center of Sepsis Care and Control, Jena University Hospital, Germany

Figures

Figure 1: Trajectory of physical recovery over 24 months

indicated by the Physical Component Score (PCS) of the SF-36, mean (95%CI) of

(A) all patients and (B) two sub cohorts: green line: complete recovery, red line: persistent impairment.

*represents P<0.05 for unpaired two-tailed Student's T-tests. Dotted line represents population norms.

Figure 2: Distribution of characteristics of both cohorts

For each figure, red columns represent the persistent impairment cohort, green columns represent the complete recovery cohort, broken down by A: Age; B: Education Status; C: number of co-morbidities; D: Family Status.

Figure 3: SF-36 components floor and ceiling effects

Red columns represent the persistent impairment cohort, and green the completed recovery cohort, both at 24-month. PF=Physical Function; RP=Role Physical; BP=Bodily Pain; GH=General Health.

*represents a value of >15% denoting relevant effect

Tables

Table 1: Baseline characteristics of different cohorts

	Persistent impairment	**NA**	**Complete Recovery**	**NA**	**Unclustered**	**NA**	
n	76		61		22		
Age (y)	65 (54.3-72)		56 (43-70)		63 (52-69.3)		**P=0.002***
Male Sex (n)[#]	47 (61.8%)		44 (72.1%)		16 (72.7%)		P=0.205
ICULOS	23.0 (12.8-39.5)	2	19 (10.0-31.0)	6	40.5 (15.3-48.3)	2	P=0.207
MV(d)	9 (2-20)	1	6 (2-22)	2	10 (4-29)	3	P=0.746
CCI	3 (1-5.8)		3 (1-5)	1	2.5 (1.8-6)		P=0.246
RRT (d)	0 (0-0.75)		0 (0-2.5)	3	0 (0-2.5)		P=0.650
Tracheostomy (n)[#]	20 (26.3%)	21	18 (29.5%)	13	11 (50%)	3	P=0.678
Intervention group (n)[#]	38 (50%)		38 (62.2%)		11(50%)		P=0.150
Education[‡$]	5 (1-9)		5 (2-9)		5 (2-9)		**P=0.039***
BMI	27.8 (24.4-32.5)		25.8 (22.6-29.1)	1	26.7 (23-30)	2	**P=0.006***
Family Status[‡$]	2 (1-6)	1	2(1-6)		2(1-4)	1	**P=0.021***
No. of ICD diagnoses at discharge	9 (6-15)		9 (5-11)		8 (6-15.8)		P=0.077

Data are shown as medians (interquartile ranges), except for percentages and mode (range). P-values represent Mann Whitney U tests between persistent impairment and complete recovery, except for #=Chi-Squared test. ICULOS= Intensive Care length of stay (days), MV(d)=Period of mechanical ventilation (days), CCI=Charlston Co-morbidity Index, RRT(d)=Renal Replacement Therapy (days), NA=not available.

[$]Indicated mode (range) with significance taken to be P<0.05, *represents p<0.05, [‡]Categories shown in Additional Table 1.

Table 2: SF-36 components floor and ceiling effects at 24 months after ICU discharge.

	Follow-Up Whole Cohort N=159		Completed recovery N=61		Persistent impairment N=76	
	Floor (0)	Ceiling (100)	Floor (0)	Ceiling (100)	Floor (0)	Ceiling (100)
PF	16 (10)	9 (6)	0 (0)	9 (15)*	16 (21)*	0 (0)
RP	71 (45)*	35(22)*	9 (15)*	27 (45)*	54 (71)*	3 (4.0)
BP	11 (7)	52(33)*	1 (2)	35 (57)*	9 (12)	7 (9.2)
GH	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
XSFMA-F	29(18)*	0(0)	29 (46)	0(0)	0(0)	0(0)

Data are shown as numbers of patients with percentages. Data of unclustered group (n=22) not shown (raw data shown in Additional Figure 2). PF= Physical Function; RP=Role Physical; BP=Bodily Pain; GH= General Health. XSFMA-F= Extra Short Form Musculoskeletal Function Assessment regarding physical function (F)

* represents a value of >15% denoting relevant effects [41].

Table 3: Concurrent Validity of physical function assessment tools

0.00-0.25	**Little**
0.25-0.50	**Fair**
0.50-0.75	**Moderate**
0.75-1.0	**Excellent**

	All					**Complete recovery**					**Persistent impairment**				
	PCS	PF	XSFMA-F	XSFMA-B	ADL	PCS	PF	XSFMA-F	XSFMA-B	ADL	PCS	PF	XSFMA-F	XSFMA-B	ADL
PCS		0.87	-0.80	-0.75	-0.61		0.82	-0.71	-0.60	-0.42		0.60	-0.62	-0.55	-0.44
PF	0.87		-0.89	-0.82	-0.73	0.82		-0.81	-0.65	-0.61	0.60		-0.81	-0.71	-0.62
XSFMA- F	-0.80	-0.89		0.91	0.84	-0.71	-0.81		0.81	0.58	-0.62	-0.81		0.84	0.78
XSFMA- B	-0.75	-0.82	0.91		0.79	-0.60	-0.65	0.81		0.41	-0.55	-0.71	0.84		0.71
ADL	-0.61	-0.73	0.84	0.79		-0.42	-0.61	0.58	0.41		-0.44	-0.62	0.78	0.71	

Data shown as coefficients of determination at 24 months after ICU discharge.

PCS=Physical Component Score of the SF-36, PF=Physical Function subscore, XSFMA-F/B=Extra Short Form Musculoskeletal Function Assessment regarding physical function (F) and disability (B) and ADL=Activities of Daily Living.

Table 4: Responsiveness of physical function scores at 6, 12 and 24 months post ICU discharge.

	All			Complete Recovery			Persistent impairment		
Month	6	12	24	6	12	24	6	12	24
PCS	0.36	0.70	0.47	1.00	1.44	1.14	0.01	0.25	0.15
PF	1.02	0.88	0.50	1.75	2.05	1.63	0.71	0.42	0.37
RP	0.68	0.34	0.31	0.73	1.07	1.16	0.70	0.07	0.03
BP	0.15	0.34	0.03	0.19	0.46	0.38	0.11	0.30	0.31
XSFMA-F		0.39	0.28		0.42	0.33		0.40	0.27
XSFMA-B		0.43	0.34		0.39	0.51		0.46	0.27
ADL		0.28	0.18		0.19	0.05		0.35	0.24

Responsiveness was measured using Cohens ‘d, with changes interpreted as minimal (0.0 to 0.2, dark grey) small (0.2 to 0.49, grey), moderate (0.5 to 0.79, yellow) and large (>0.80, green). Six month XSFMA-F/B data were used as baseline for responsiveness**.**

Additional files

Additional file 1

- Additional file 1.docx
- Additional Table 1: Categories of Educational Level and Family Status
- VT=Vocational Training
- GSCE=General Certificate of Secondary Education
- Additional Table 1.1: Addressed domains of used questionnaires

Additional file 2

- Additional file 2.docx
- Additional Table 2: Summary measures for Trajectory Projection
- eMethods of use of trajectory projection

Additional file 3

- Additional file 3.docx
- Additional Table 3: Baseline characteristics of the whole cohort and the 24 months follow-up cohort
- Values shown as medians and interquartile range [IQR] except for $representing mode (range). P-values represent two-tailed Mann-Whitney U-tests, except for #=Chi-Squared test. ICULOS= Intensive Care length of stay. MV (d)=period of mechanical ventilation (days), CCI=Charlston Co-morbidity Index, RRT(d)=Renal Replacement Therapy (days), PCS=Physical Component Score of the SF-36, MCS =Mental Component Score recall 3 months prior to critical illness. XSFMA F/B= Extra

Short Musculoskeletal Function Assessment regarding Physical Function and Disability, 3m recall=recall 3 months prior to critical illness. NA=Not available, *Categories shown in Additional Table 1

- [1]47 patients without MV, 11 patients without available data, [2]209 patients without RRT, 5 patients without available data

- Additional Table 3.1: Baseline characteristics of the whole cohort split by loss to follow-up and death.
- Values shown as medians and interquartile range [IQR] except for [$]representing mode (range). ICULOS= Intensive Care length of stay. MV (d)=period of mechanical ventilation (days), CCI=Charlston Co-morbidity Index, RRT(d)=Renal Replacement Therapy (days), PCS=Physical Component Score of the SF-36, MCS =Mental Component Score recall 3 months prior to critical illness. XSFMA F/B= Extra Short Musculoskeletal Function Assessment regarding Physical Function and Disability, 3m recall=recall 3 months prior to critical illness. NA=Not available, *Categories shown in Table S1

Additional file 4

- Additional file 4.docx
- Additional Table 4: Bivariable and multivariate logistic regression analysis of cohort membership characteristics
- Dependent variable: Allocation to persistent impairment cohort vs. complete recovery cohort. ICD=International Classification of Disease; ICULOS= Intensive Care Unit Length of Stay. * represents $p<0.05$

- Additional Table 4.1: Multinomial regression for the persistent impairment group, using the full recovery as the reference group. ICD=International Classification of Disease; ICULOS= Intensive Care Unit Length of Stay; * represents $p<0.05$

- Additional Table 4.2: Multinomial regression for the unclustered group, using the full recovery as the reference group. ICD=International Classification of Disease; ICULOS= Intensive Care Unit Length of Stay; * represents $p<0.05$

Additional file 5

- Additional file 5.docx
- Additional Tables 5A and B: Ceiling and floor effects
- Data are shown as n(%) over time for SF-36 components in patients with a persistent impairment trajectory (n=76) and in patients with a completed recovery trajectory (n=61) (Table 5A: only patients with completed recovery). PF= Physical Function; RP= Role Physical, BP=Bodily Pain, GH= General Health, XSFMA-F= Extra Short Form Musculoskeletal Function Assessment regarding physical function (F)

*represents a value of >15% denoting relevant effect. % may not=100 due to rounding effects.

Additional file 6

- Additional file 6.png
- Additional Figure 1: Area under receiver operating characteristic curve (AUROC)
- Logistic regression of predictors of cluster allocation

Additional file 7

- Additional file 7.png
- Additional Figure 2: Trajectories of unclustered patients (n=22)
- Data points are means of the SF-36 Physical Component Score (PCS) over 24 months after discharge from ICU.

References

1. Iwashyna TJ, Ely EW, Smith DM, Langa KM: **Long-term cognitive impairment and functional disability among survivors of severe sepsis**. *JAMA* 2010, **304**(16):1787-1794.
2. Herridge MS, Tansey CM, Matté A, Tomlinson G, Diaz-Granados N, Cooper A, Guest CB, Mazer CD, Mehta S, Stewart TE *et al*: **Functional Disability 5 Years after Acute Respiratory Distress Syndrome**. *New Engl J Med* 2011, **364**(14):1293-1304.
3. Morris PE, Berry MJ, Files DC, Thompson JC, Hauser J, Flores L, Dhar S, Chmelo E, Lovato J, Case LD *et al*: **Standardized Rehabilitation and Hospital Length of Stay Among Patients With Acute Respiratory Failure: A Randomized Clinical Trial**. *JAMA* 2016, **315**(24):2694-2702.
4. Walsh TS, Salisbury LG, Merriweather JL, et al.: **Increased Hospital-Based Physical Rehabilitation and Information Provision After Intensive Care Unit Discharge. The RECOVER Randomized Clinical Trial.** *JAMA Intern Med* 2015, **175**(6):901-910.
5. Denehy L, Skinner EH, Edbrooke L, Haines K, Warrillow S, Hawthorne G, Gough K, Hoorn SV, Morris ME, Berney S: **Exercise rehabilitation for patients with critical illness: a randomized controlled trial with 12 months of follow-up**. *Critical care (London, England)* 2013, **17**(4):R156.
6. Wright SE, Thomas K, Watson G, Baker C, Bryant A, Chadwick TJ, Shen J, Wood R, Wilkinson J, Mansfield L *et al*: **Intensive versus standard physical rehabilitation therapy in the critically ill (EPICC): a multicentre, parallel-group, randomised controlled trial**. *Thorax* 2018, **73**(3):213-221.

7. McCarthy B, Casey D, Devane D, Murphy K, Murphy E, Lacasse Y: **Pulmonary rehabilitation for chronic obstructive pulmonary disease**. *Cochrane database of systematic reviews* 2015(2):CD003793.

8. Bear DE, Puthucheary ZA: **Designing nutrition-based interventional trials for the future: addressing the known knowns**. *Critical care (London, England)* 2019, **23**(1):53.

9. Puthucheary ZA, Denehy L: **Exercise Interventions in Critical Illness Survivors: Understanding Inclusion and Stratification Criteria**. *American journal of respiratory and critical care medicine* 2015, **191**(12):1464-1467.

10. Dowdy DW, Eid MP, Sedrakyan A, Mendez-Tellez PA, Pronovost PJ, Herridge MS, Needham DM: **Quality of life in adult survivors of critical illness: a systematic review of the literature**. *Intensive care medicine* 2005, **31**(5):611-620.

11. Griffith DM, Salisbury LG, Lee RJ, Lone N, Merriweather JL, Walsh TS, Investigators R: **Determinants of Health-Related Quality of Life After ICU: Importance of Patient Demographics, Previous Comorbidity, and Severity of Illness**. *Critical care medicine* 2018, **46**(4):594-601.

12. Jones JRA, Berney S, Connolly B, Waterland JL, Denehy L, Griffith DM, Puthucheary ZA: **Socioeconomic Position and Health Outcomes Following Critical Illness: A Systematic Review**. *Critical care medicine* 2019.

13. Puthucheary ZA, Rawal J, McPhail M, Connolly B, Ratnayake G, Chan P, Hopkinson NS, Padhke R, Dew T, Sidhu PS *et al*: **Acute skeletal muscle wasting in critical illness**. *JAMA* 2013, **310**(15):1591-1600.

14. Puthucheary ZA, Astin R, McPhail MJW, Saeed S, Pasha Y, Bear DE, Constantin D, Velloso C, Manning S, Calvert L *et al*: **Metabolic phenotype of skeletal muscle in early critical illness**. *Thorax* 2018.

15. Brown SM, Wilson EL, Presson AP, Dinglas VD, Greene T, Hopkins RO, Needham DM, with the National Institutes of Health NAN: **Understanding patient outcomes after acute respiratory distress syndrome: identifying subtypes of physical, cognitive and mental health outcomes**. *Thorax* 2017, **72**(12):1094-1103.

16. Batt J, dos Santos CC, Cameron JI, Herridge MS: **Intensive Care Unit–acquired Weakness: Clinical Phenotypes and Molecular Mechanisms**. *American journal of respiratory and critical care medicine* 2013, **187**(3):238-246.

17. McNelly AS, Rawal J, Shrikrishna D, Hopkinson NS, Moxham J, Harridge SD, Hart N, Montgomery HE, Puthucheary ZA: **An Exploratory Study of Long-Term Outcome Measures in Critical Illness Survivors: Construct Validity of Physical Activity, Frailty, and Health-Related Quality of Life Measures**. *Critical care medicine* 2016, **44**(6):e362-369.

18. **High Quality Care for All** [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/228836/7432.pdf]

19. Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M: **The importance of patient-reported outcomes in clinical trials and strategies for future optimization**. *Patient Relat Outcome Meas* 2018, **9**:353-367.

20. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan AW, King MT, the S-PROG, Hunn A, Bottomley A, Regnault A *et al*: **Guidelines for Inclusion of Patient-Reported**

Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA* 2018, **319**(5):483-494.

21. Herridge MS, Cheung AM, Tansey CM, Matte-Martyn A, Diaz-Granados N, Al-Saidi F, Cooper AB, Guest CB, Mazer CD, Mehta S *et al*: **One-year outcomes in survivors of the acute respiratory distress syndrome**. *The New England journal of medicine* 2003, **348**(8):683-693.

22. Gandotra S, Lovato J, Case D, Bakhru RN, Gibbs K, Berry M, Files DC, Morris PE: **Physical Function Trajectories in Survivors of Acute Respiratory Failure**. *Ann Am Thorac Soc* 2019, **16**(4):471-477.

23. Schmidt K, Worrack S, Von Korff M, Davydow D, Brunkhorst F, Ehlert U, Pausch C, Mehlhorn J, Schneider N, Scherag A *et al*: **Effect of a Primary Care Management Intervention on Mental Health-Related Quality of Life Among Survivors of Sepsis: A Randomized Clinical Trial**. *JAMA* 2016, **315**(24):2703-2711.

24. Schmidt KF, Schwarzkopf D, Baldwin LM, Brunkhorst FM, Freytag A, Heintze C, Reinhart K, Schneider N, von Korff M, Worrack S *et al*: **Long-Term Courses of Sepsis Survivors: Effects of a Primary Care Management Intervention**. *Am J Med* 2020, **133**(3):381-385 e385.

25. Schmidt K, Thiel P, Mueller F, Schmuecker K, Worrack S, Mehlhorn J, Engel C, Brenk-Franz K, Kausche S, Jakobi U *et al*: **Sepsis survivors monitoring and coordination in outpatient health care (SMOOTH): study protocol for a randomized controlled trial**. *Trials* 2014, **15**:283.

26. Bullinger M: **German translation and psychometric testing of the SF-36 Health Survey: preliminary results from the IQOLA Project. International Quality of Life Assessment**. *Social science & medicine* 1995, **41**(10):1359-1366.

27. Wollmerstedt N, Kirschner S, Bohm D, Faller H, Konig A: **[Design and evaluation of the Extra Short Musculoskeletal Function Assessment Questionnaire XSMFA-D]**. *Z Orthop Ihre Grenzgeb* 2003, **141**(6):718-724.

28. Engelberg R, Martin DP, Agel J, Obremsky W, Coronado G, Swiontkowski MF: **Musculoskeletal Function Assessment instrument: criterion and construct validity**. *J Orthop Res* 1996, **14**(2):182-192.

29. Kayambu G, Boots R, Paratz J: **Early physical rehabilitation in intensive care patients with sepsis syndromes: a pilot randomised controlled trial**. *Intensive care medicine* 2015, **41**(5):865-874.

30. Eggmann S, Verra ML, Luder G, Takala J, Jakob SM: **Effects of early, combined endurance and resistance training in mechanically ventilated, critically ill patients: A randomised controlled trial**. *PloS one* 2018, **13**(11):e0207428.

31. Fossat G, Baudin F, Courtes L, Bobet S, Dupont A, Bretagnol A, Benzekri-Lefevre D, Kamel T, Muller G, Bercault N *et al*: **Effect of In-Bed Leg Cycling and Electrical Stimulation of the Quadriceps on Global Muscle Strength in Critically Ill Adults: A Randomized Clinical Trial**. *JAMA* 2018, **320**(4):368-378.

32. Moss M, Nordon-Craft A, Malone D, Van Pelt D, Frankel SK, Warner ML, Kriekels W, McNulty M, Fairclough DL, Schenkman M: **A Randomized Trial of an Intensive Physical Therapy Program for Patients with Acute Respiratory Failure**. *American journal of respiratory and critical care medicine* 2016, **193**(10):1101-1110.

33. Needham DM, Dinglas VD, Bienvenu OJ, Colantuoni E, Wozniak AW, Rice TW, Hopkins RO, Network NNA: **One year outcomes in patients with acute lung injury randomised to initial trophic or full enteral feeding: prospective follow-up of EDEN randomised trial**. *BMJ (Clinical research ed* 2013, **346**:f1532.

34. Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Friedman LA, Bingham CO, 3rd, Turnbull AE: **Core Outcome Measures for Clinical Research in Acute Respiratory Failure Survivors. An International Modified Delphi Consensus Study**. *American journal of respiratory and critical care medicine* 2017, **196**(9):1122-1130.

35. [https://cran.r-project.org/web/packages/traj/ Accessed 06 Mar 2020]

36. Leffondre K, Abrahamowicz M, Regeasse A, Hawker GA, Badley EM, McCusker J, Belzile E: **Statistical measures were proposed for identifying longitudinal patterns of change in quantitative health indicators**. *J Clin Epidemiol* 2004, **57**(10):1049-1062.

37. Sylvestre MP, McCusker J, Cole M, Regeasse A, Belzile E, Abrahamowicz M: **Classification of patterns of delirium severity scores over time in an elderly population**. *Int Psychogeriatr* 2006, **18**(4):667-680.

38. Portney L.G WMP: **Foundations of Clinical Research**. In: *Foundations of Clinical Research.* Pearson; 2008: 125.

39. Nordon-Craft A, Schenkman M, Edbrooke L, Malone DJ, Moss M, Denehy L: **The physical function intensive care test: implementation in survivors of critical illness**. *Physical therapy* 2014, **94**(10):1499-1507.

40. Fries JF, Krishnan E, Rose M, Lingala B, Bruce B: **Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory**. *Arthritis Res Ther* 2011, **13**(5):R147.

41. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC: **Quality criteria were proposed for measurement properties of health status questionnaires**. *J Clin Epidemiol* 2007, **60**(1):34-42.

42. Davis E, Young D, Waters E, Gold L: **Children and Adolescents, Measuring the Quality of Life of**. In: *International Encyclopedia of Public Health (Second Edition).* Edited by Quah SR. Oxford: Academic Press; 2017: 539-545.

43. Portney LG, Watkins MP: **Foundations of clinical research, applications to practice**, 3rd Edition edn. Conneticut: Appleton and Lange; 2009.

44. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC: **The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study**. *Qual Life Res* 2010, **19**(4):539-549.

45. Fava GA, Tomba E, Sonino N: **Clinimetrics: the science of clinical measurements**. *International journal of clinical practice* 2012, **66**(1):11-15.

46. Kazis LE, Anderson JJ, Meenan RF: **Effect sizes for interpreting changes in health status**. *Medical care* 1989, **27**(3 Suppl):S178-189.

47. Husted J, Cook R, Farewell V, Gladman D: **Methods for assessing responsiveness: A critical review and recommendations**. *Journal of Clinical Epidemiology* 2000, **53**:459-468.

48. Cohen J: **Statistical Power Analysis for the Behavioural Sciences**, 2nd ed edn. Hillsdale: Lawrence Erlbaum; 1988.

49. Jones JRA, Berney S, Connolly B, Waterland JL, Denehy L, Griffith DM, Puthucheary ZA: **Socioeconomic Position and Health Outcomes Following Critical Illness: A Systematic Review**. *Critical care medicine* 2019, **47**(6):e512-e521.

50. Pollock RD, Carter S, Velloso CP, Duggal NA, Lord JM, Lazarus NR, Harridge SD: **An investigation into the relationship between age and physiological function in highly active older adults**. *The Journal of physiology* 2015, **593**(3):657-680; discussion 680.

51. Ferrante LE, Murphy TE, Leo-Summers LS, Gahbauer EA, Pisani MA, Gill TM: **The Combined Effects of Frailty and Cognitive Impairment on Post-ICU Disability among Older ICU Survivors**. *American journal of respiratory and critical care medicine* 2019, **200**(1):107-110.

52. Herridge MS, Chu LM, Matte A, Tomlinson G, Chan L, Thomas C, Friedrich JO, Mehta S, Lamontagne F, Levasseur M *et al*: **The RECOVER Program: Disability Risk Groups and 1-Year Outcome after 7 or More Days of Mechanical Ventilation**. *American journal of respiratory and critical care medicine* 2016, **194**(7):831-844.

53. Allingstrup MJ, Kondrup J, Wiis J, Claudius C, Pedersen UG, Hein-Rasmussen R, Bjerregaard MR, Steensen M, Jensen TH, Lange T *et al*: **Early goal-directed nutrition versus standard of care in adult intensive care patients: the single-centre, randomised, outcome assessor-blinded EAT-ICU trial**. *Intensive care medicine* 2017, **43**(11):1637-1647.

54. McDowell K, O'Neill B, Blackwood B, Clarke C, Gardner E, Johnston P, Kelly M, McCaffrey J, Mullan B, Murphy S *et al*: **Effectiveness of an exercise programme on physical function in patients discharged from hospital following critical illness: a randomised controlled trial (the REVIVE trial)**. *Thorax* 2017, **72**(7):594-595.

55. Connolly B, Denehy L: **Hindsight and moving the needle forwards on rehabilitation trial design**. *Thorax* 2018, **73**(3):203-205.

56. Ware JE, Jr., Kosinski M, Gandek B, Aaronson NK, Apolone G, Bech P, Brazier J, Bullinger M, Kaasa S, Leplege A *et al*: **The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. International Quality of Life Assessment**. *J Clin Epidemiol* 1998, **51**(11):1159-1165.

57. Ware JE, Kosinski M: **Interpreting SF-36 summary health measures: a response**. *Qual Life Res* 2001, **10**(5):405-413; discussion 415-420.

58. Ware JE, Jr., Gandek B, Kosinski M, Aaronson NK, Apolone G, Brazier J, Bullinger M, Kaasa S, Leplege A, Prieto L *et al*: **The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: results from the IQOLA Project. International Quality of Life Assessment**. *J Clin Epidemiol* 1998, **51**(11):1167-1170.

59. Biering K, Hjollund NH, Frydenberg M: **Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes**. *Clin Epidemiol* 2015, **7**:91-106.

60. Edbrooke L, Aranda S, Granger CL, McDonald CF, Krishnasamy M, Mileshkin L, Clark RA, Gordon I, Irving L, Denehy L: **Multidisciplinary home-based rehabilitation in inoperable lung cancer: a randomised controlled trial**. *Thorax* 2019, **74**(8):787-796.

61. Colantuoni E, Scharfstein DO, Wang C, Hashem MD, Leroux A, Needham DM, Girard TD: **Statistical methods to compare functional outcomes in randomized controlled trials with high mortality**. *BMJ (Clinical research ed* 2018, **360**:j5748.

62. Swiontkowski MF, Engelberg R, Martin DP, Agel J: **Short musculoskeletal function assessment questionnaire: validity, reliability, and responsiveness**. *J Bone Joint Surg Am* 1999, **81**(9):1245-1260.

63. Hoekstra M, Maslove DM, Veldhoen RA, Marshall JC, Muscedere J: **Insights Into a "Negative" ICU Trial Derived From Gene Expression Profiling**. *Critical care medicine* 2019, **47**(12):e941-e947.

64. Sydes MR, Parmar MK, Mason MD, Clarke NW, Amos C, Anderson J, de Bono J, Dearnaley DP, Dwyer J, Green C *et al*: **Flexible trial design in practice - stopping**

arms for lack-of-benefit and adding research arms mid-trial in STAMPEDE: a multi-arm multi-stage randomized controlled trial. *Trials* 2012, **13**:168.

65. Wang SJ, Hung HM: **Adaptive enrichment with subpopulation selection at interim: methodologies, applications and design considerations**. *Contemporary clinical trials* 2013, **36**(2):673-681.

66. Kaplan R, Maughan T, Crook A, Fisher D, Wilson R, Brown L, Parmar M: **Evaluating many treatments and biomarkers in oncology: a new design**. *J Clin Oncol* 2013, **31**(36):4562-4568.

67. Pallmann P, Bedding AW, Choodari-Oskooei B, Dimairo M, Flight L, Hampson LV, Holmes J, Mander AP, Odondi L, Sydes MR *et al*: **Adaptive designs in clinical trials: why use them, and how to run and report them**. *BMC medicine* 2018, **16**(1):29.