



Published in final edited form as:

Crit Care Med. 2016 July ; 44(7): 1267–1277. doi:10.1097/CCM.0000000000001651.

Outcome Measurement in ICU Survivorship Research from 1970-2013: A Scoping Review of 425 Publications

Alison E. Turnbull, DVM, MPH, PhD^{1,2,3}, Anahita Rabiee, MD^{1,2}, Wesley E. Davis, BA^{1,2}, Mohamed Farhan Nasser, MBBS¹, Venkat Reddy Venna, MBBS¹, Rohini Lolitha, MBBS¹, Ramona O. Hopkins, PhD^{4,5,6}, O. Joseph Bienvenu, MD, PhD^{1,7}, Karen A. Robinson, MSc, PhD^{3,8,9}, and Dale M. Needham, FCPA, MD, PhD^{1,2,10}

¹Outcomes After Critical Illness and Surgery Group, Johns Hopkins University, Baltimore, MD

²Division of Pulmonary and Critical Care Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD

³Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

⁴Department of Medicine, Pulmonary and Critical Care Division, Intermountain Medical Center, Murray, UT

⁵Center for Humanizing Critical Care, Intermountain Healthcare, Murray, UT

⁶Psychology Department and Neuroscience Center, Brigham Young University, Provo, UT

⁷Department of Psychiatry and Behavior Sciences, School of Medicine, Johns Hopkins University, Baltimore, MD

⁸Department of Health Policy & Management, Bloomberg School of Public Health, Johns Hopkins University

⁹Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD

¹⁰Department of Physical Medicine and Rehabilitation, School of Medicine, Johns Hopkins University, Baltimore

Abstract

Objectives—To evaluate the study designs and measurement instruments used to assess physical, cognitive, mental health, and quality-of-life outcomes of survivors of critical illness over more than 40 years as a first step toward developing a core outcome set of measures for future trials to improve outcomes in ICU survivors.

Design—Scoping review

Corresponding Author: Alison E. Turnbull, DVM, MPH, PhD, Pulmonary and Critical Care Medicine, Johns Hopkins University, 1830 E. Monument St, 5th Floor, Baltimore, MD 21205, Phone: (410) 955-2190; Fax: (410) 955-0036; turnbull@jhmi.edu.

All authors contributed to the conception and/or design of this study. AET, AR, WED, MFN, VRV, RL, KAR and DMN contributed to the acquisition of data. All authors contributed to the analysis and interpretation of data. AET drafted the manuscript, and all authors critically revised it for important intellectual content and approved the final version to be submitted.

The remaining authors have disclosed that they do not have any potential conflicts of interest.

Setting—Published articles that included 1 post-discharge measure of a physical, cognitive, mental health, or quality of life outcome in 20 survivors of critical illness published between 1970 and 2013. Instruments were classified using the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) framework.

Interventions—None.

Measurements and Main Results—We reviewed 15,464 abstracts, and identified 425 eligible articles, including 31 (7%) randomized trials, 116 (27%) cross-sectional studies, and 278 (65%) cohort studies. Cohort studies had a median (interquartile range) sample size of 96 (52-209) survivors, with 38% not fully reporting loss to follow-up. A total of 250 different measurement instruments were used in these 425 articles. Among eligible articles, 25 (6%) measured physical activity limitations, 40 (9%) measured cognitive activity limitations, 114 (27%) measured mental health impairment, 196 (46%) measured participation restriction, and 276 (65%) measured quality of life.

Conclusions—Peer-reviewed publications reporting patient outcomes after hospital discharge for ICU survivors have grown from 3 in the 1970s to >300 since 2000. Although there is evidence of consolidation in the instruments used for measuring participation restriction and quality of life, the ability to compare results across studies remains impaired by the 250 different instruments used. Most articles described cohort studies of modest size with a single follow-up assessment using patient-reported measures of participation restriction and quality of life. Development of a core outcomes set of valid, reliable and feasible measures is essential to improving the outcomes of critical illness survivors.

Keywords

Patient Outcome Assessment; Follow-up studies; Intensive care; Disability Evaluation; Quality of Life; Cognition Disorders

Introduction

The demand for critical care services and the proportion of intensive care unit (ICU) patients surviving until hospital discharge have risen steadily over recent decades.(1–4) As a result of these trends, the number of ICU survivors is growing. However, critical illness survivorship often comes at a cost, with many survivors experiencing new and long-lasting physical,(5, 6) cognitive,(7) and mental health sequelae,(8–10) as well as impaired quality of life.(11, 12)

In response to this growing population of ICU survivors, many professional and scientific organizations, including the American Thoracic Society, NHLBI and the Multi-society Task Force for Critical Care Research, have recommended prioritizing research on the outcomes of survivors of critical illness after hospital discharge.(13–19) Although there is strong support for such research, there is no consensus on the most important outcomes and measurement instruments for assessment, creating challenges in comparing and synthesizing results across the large and growing number of studies in this field.(20)

Consensus about the most important outcomes and measurement instruments for studies of ICU survivors will not occur organically.(21) Widespread adoption of a recommended and

standardized minimum collection of outcomes, known as a core outcome set, will require deliberate effort, input from relevant stakeholders, and evaluation of the psychometric and other properties of existing measurement instruments.(22) A proposed methodology for this process is presented in Figure E1. (Supplemental digital content 1) A key initial step in this process is to identify and summarize existing outcome domains and measurement instruments via a scoping review.(23–26)

Hence, our objective was to document the study designs and measurement instruments used to assess outcomes of survivors of critical illness between 1970 and 2013. We undertook a scoping review of publications on ICU survivors' physical, cognitive, mental health, and quality of life outcomes, using the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) framework (27, 28).

Methods

Study design

The scoping review was conducted according to Arksey and O'Malley's 5-stage framework (29) and reported using recent recommendations.(24–26) Scoping reviews are an increasingly popular approach to summarizing the breadth and nature of research activity in a field.

Research Question

What study designs and measurement instruments were used in research on the physical, cognitive, mental health, and quality-of-life outcomes of ICU survivors published between 1970 and 2013? The study was approved by the institutional review board of Johns Hopkins University (NA_00087504).

Identification of eligible studies

As of November 7, 2013 we searched five electronic databases: PubMed, EMBASE, PsycINFO®, Cumulative Index of Nursing and Allied Health Literature (CINAHL), and the Cochrane Controlled Trials Registry (CENTRAL) using search strategies including a combination of keywords and controlled vocabulary for the concepts of “intensive care” combined with “outcome assessment”, “health status”, “functional status”, “quality of life” and “follow-up”. (Table E2, Supplemental digital content 2) In addition, we conducted a hand search of personal files, reference lists of relevant narrative and systematic review articles, and consulted with experts for eligible articles.

Study selection

We sought peer-reviewed, published studies of 20 adult ICU survivors assessed after hospital discharge. Articles were excluded if their primary intent was to evaluate or describe the psychometric properties of a measurement instrument, if outcomes were only assessed via qualitative methods (e.g., semi-structured interview or primarily open-ended questions), or if survival was the only outcome reported. Articles also were excluded if a majority (>50%) of the study population was 1) <16 years of age, 2) had neurological injury, 3) had undergone cardiac surgery, or 4) had not been admitted to an ICU.

Data abstraction

Abstract and full-text screening was managed using DistillerSR[®] (2014 Evidence Partners, Ottawa, Canada). Two independent reviewers screened the titles and abstracts of each retrieved citation and the full text of potentially eligible articles to make a final determination of eligibility.

Data abstraction from eligible articles was completed between May 2014 and March 2015 by reviewers (AR, WED, MFN, VRV, RL) who were trained via didactic training, and followed a 23-page written operations manual. Before collecting data, reviewers were required to demonstrate a high level of proficiency by abstracting data from articles previously abstracted by AET as gold standard quality assurance evaluations. Additionally, abstraction was conducted in duplicate with discrepancies resolved through consensus, or adjudication by the authors (AET, AR, or DMN) when needed.

For each eligible article, the following data were abstracted using a standardized data collection form, with data entry into a REDCap database (30): 1) publication year, 2) study design, 3) number of eligible ICU survivors alive at hospital discharge and at the last outcome assessment, 4) number of ICU survivors assessed for any outcome at the last outcome assessment, and 5) the measurement instruments used to assess each outcome. Measurement instruments were defined as “custom-made” if the instrument's psychometric properties had not been previously evaluated and the instrument had not been used in at least 1 previously-published study.

Physical, cognitive, mental health, and participation outcome measures were categorized according to the ICF domains as follows: 1) structure and function impairment, 2) activity limitation, and 3) participation restriction. Quality of life outcomes were added as a fourth domain, consistent with previous recommendations (Figure 1). (28) For each outcome measure, the following data were collected: 1) measurement instruments used, 2) existence and mode of any baseline assessment prior to hospitalization (i.e., none, patient interview, proxy interview, chart review), 3) number of follow-up visits in which the outcome was measured, and 4) the duration of follow-up.

Summary and reporting

Study characteristics for categorical variables were summarized as percentages and continuous variables as medians with interquartile ranges. Changes in the proportion of articles using custom-made measures before versus after 2000 (i.e., 1970 – 1999 versus 2000 – 2013) were assessed using a z-test for equality of proportions. The heterogeneity of measurement instruments within a domain was summarized using the ratio of eligible articles to unique outcome instruments, which we termed the “article-to-instrument” ratio (A:I). Generally, higher A:I ratios are produced by domains in which many articles utilize a relatively small number of standardized measurement instruments. However, interpretation of the A:I ratio requires consideration of the total number of articles within a domain, and the use of test batteries, requiring a large number of instruments, as standard practice for evaluating cognitive activity limitations. All descriptive statistics and plots were generated using R (version 3.0.1; R Development Core Team).

Results

Data synthesis

A total of 15,464 non-duplicate citations were reviewed, of which 1,207 were selected for full-text screening, yielding 425 eligible articles (Figure E3, Supplemental digital content 3). Although the number of articles reporting post-discharge outcomes of ICU survivors increased from 3 in the 1970s to more than 300 since 2000, the proportion of all critical care articles reporting post-discharge outcomes has remained small (Figure 2). The 425 articles eligible for this scoping review represented <2% of the 26,169 critical care articles identified in PubMed between 1970 and 2013 (Table E4, Supplemental digital content 4).

Methodological characteristics of eligible studies

A total of 31 (7%) of the 425 eligible articles were randomized controlled trials (RCTs), with the majority of articles (65%) being cohort studies, 60% of which had a single follow-up assessment and a median sample size of 107 (IQR 53 – 255) (Table 1). The median and mode time of final outcome assessment was 6 months for cohort studies with a single follow-up assessment and 12 months for randomized trials and cohort studies with multiple follow-up assessments. Across all study designs, 38% of articles did not fully report retention rate. Among reporting articles, the median participant retention rate was 80% (IQR 65% - 93%).

Outcome Assessments by ICF Domain

Among the 425 eligible articles, 150 (35%) measured structure and function impairment, 62 (15%) measured activity limitation, 196 (46%) measured participation restriction, and 276 (65%) measured quality of life. A total of 190 papers (45%) measured outcomes in >1 domain and 58 papers (14%) measured >2 domains. There were 250 unique outcome instruments, including all custom-made measures (n=46), used in the 425 eligible articles, for an A:I ratio of 1.7.

Within measures of physical structure and function impairment, pulmonary impairment measures (evaluated using pulmonary function testing) were the most common, found in 32 articles (84% of 38) (Table 2). For measures of neuromuscular impairment, there were 7 articles, with 6 (86%) of them published after 1999 and the most common measures being electromyography/nerve conduction study (57%) and manual muscle testing using the Medical Research Council grading system (43%). The A:I ratio was 2.5 (20:8) for measures of pulmonary impairment and 1.2 (6:5) for neuromuscular impairment.

For measures of mental health impairment, there were 114 articles using 39 unique measurement instruments (A:I = 2.9). Of these articles, 103 (90%) were published after 1999, 73 (63%) using >1 instrument, and 65 (57%) measuring >1 aspect of mental health impairment (e.g. depression and PTSD) (Table 3). The Hospital Anxiety and Depression Scale (31, 32) was the most common measure of depressive (60%) and anxiety (74%) symptoms, while the Impact of Event Scale (IES, 37%), (33, 34) Post-Traumatic Stress Syndrome 10-Questions instrument (24%), (35, 36) and the IES-Revised (19%), (33, 34) were

the most frequent measures of post-traumatic stress. Only 11 articles were published prior to 2000.

For measures of physical activity limitations, there were 25 articles using 8 unique outcome instruments (A:I = 3.1). There were 20 articles (80%) published after 1999, with the 6-Minute Walk Test (37, 38) used in 80% (Table E5, Supplemental digital content 5).

For measures of cognitive limitations, there were 40 articles, with 37 (93%) published after 1999, 37 (93%) using 1 measure of general intelligence, and 75% using >1 measure (many used extensive batteries of tests; Tables 4 and E6, Supplemental digital content 6). The Mini-Mental State Examination (MMSE) and the Trail Making Test Parts A & B were the most commonly used measures.

For measures of participation restriction, there were 196 articles with 131 (67%) published after 1999. There were 71 unique measurement instruments (A:I = 2.8), with 10 different instruments used in 3 articles (Table 5). "Return to work", assessed in 87 articles (44%), was the most common assessment of participation restriction, but measurement of this outcome was not standardized across studies. Excluding articles that reported on "return to work," there were 84 articles (43%) with a baseline assessment of participation restriction. Comparing pre- versus post-2000, the proportion of articles assessing participation restriction using custom-made instruments decreased from 35% to 4% ($p < 0.001$), with increases in use of Katz's Activities of Daily Living (39, 40) and Lawton's Instrumental Activities of Daily Living (41) from 15% to 24%, and 2% to 11%, respectively.

For quality of life, there were 276 articles with 222 (80%) published after 1999. There were 58 unique instruments (A:I ratio = 4.8), with 9 different instruments used in 3 articles (Table 6). There were 80 articles (29%) with a baseline assessment of quality of life. Comparing pre- versus post-2000, the proportion of articles measuring quality of life using custom-made instruments decreased from 20% to 3% ($p < 0.001$). After 1999, 63% of articles measuring quality of life used the Short Form-36 (13, 15, 42) and 19% used EQ-5D-3L (13, 15, 43, 44).

Discussion

This scoping review of measures of physical, cognitive, mental health and quality of life outcomes in ICU survivors evaluated 425 peer-reviewed articles, with the majority using patient-reported measures from 1 outcome domain assessed at 1 time-point after hospital discharge. This body of literature demonstrates a large number of different measurement instruments used to assess physical, cognitive, mental health, and quality-of-life outcomes. Although the number of articles reporting post-discharge outcomes of ICU survivors has increased from 3 in the 1970s to more than 300 since 2000, our ability to compare results or reach conclusions remains impeded by the use of 250 unique measurement instruments. Also, the number of RCTs assessing post-discharge outcomes remains small ($N=31$), often with modest sample sizes (50% with $N < 100$). Meta-analyses might help synthesize these small trials, but are challenging given the heterogeneity in measurement instruments used. Hence, purposeful steps are required to identify core outcome sets of agreed-upon, validated

measures.(22) This scoping review will inform the process of establishing a core outcome sets for studies of ICU survivors by providing an overview of the existing research.

Previous reviews of post-discharge outcome assessment in ICU survivors have summarized research on post-discharge outcomes between 1970 and 2003. (45, 46) Our findings emphasize the changes that have taken place in the more than 10 years since these undertakings. For example, a 2001 review recommended the use of the Sickness Impact Profile as a measure of health-related quality of life. This measure was used in 22% of articles assessing quality of life between 1970 and 1999 but only 5% of articles between 2000 and 2013 (Table 6).

We did not attempt to understand the forces driving the proliferation of outcome instruments or trends in the use of specific instruments. Proliferation may continue because there has not been a formal needs-evaluation in the field of ICU outcomes research. Investigators may develop new instruments because they are unaware of existing tools or because they judge existing tools to lack face or content validity. As of June 2015, six critical care projects are registered for core outcomes set development to address this challenge.(47)

There were 196 and 276 papers assessing participation restriction and quality-of-life, respectively, but only 25 and 40 assessing physical and cognitive activity limitations, respectively. Measures of participation restriction and quality-of-life are patient-reported outcomes assessed by survey, making them more feasible to study (including administration by mail or phone) rather than performance-based tests of physical limitations typically evaluated in-person within a standardized setting. However, performance-based and patient-reported outcomes measures evaluate different aspects of patient outcomes that are not always highly correlated.(48) Conducting both performance-based tests and patient-reported outcomes in the same population is necessary to elucidate mechanistic relationships between physiological outcomes and patient-important outcomes.

The number of different outcome instruments used in ICU survivorship research substantially varies across domains. For example, 43 and 39 standardized instruments were used to measure participation restriction and quality of life, respectively, while only 5 and 8 instruments were used for neuromuscular impairment and physical activity limitations. There is some evidence of consolidation around key instruments for participation restriction and quality of life. The A:I ratio rose from 1.5 prior to 2000 to 3.4 after 2000 for participation restriction, with greater use of Katz's Activities of Daily Living and Lawton's Instrumental Activities of Daily Living instruments, and from 1.9 to 5.8 for quality of life as investigators adopted the SF-36 and EQ-5D following their introduction in the 1990s and subsequent recommendation for use in critical care survivorship research.(13)

Importantly, 38% of papers did not fully report loss to follow-up. Studies estimating the prevalence or incidence of post-ICU morbidities risk survivorship bias if they fail to recognize that survivors lost to follow-up may have different levels of impairment than evaluated survivors. Hence, it is critical that investigators always report the number of individuals completing follow-up and the reasons for non-participation in each stage of a study as described by the 2006 Strengthening the Reporting of Observational Studies in

Epidemiology (STROBE) statement and checklist.(49) Editors and peer reviewers can help strengthen the methodological rigor of the field by holding investigators to this standard.

This scoping review has strengths and limitations. To our knowledge, this is the largest review of post-discharge outcomes in ICU survivors, with 15,464 abstracts screened and 425 full-text articles included. Although the search process was thorough, eligible studies may have been inadvertently excluded. To minimize this risk, all abstracts and manuscripts were reviewed in duplicate, and content experts were consulted to review personal files for eligible publications and to evaluate if all key articles were included. Moreover, although important to the development of patient-important outcome measures, assessing ICU survivors' outcomes using qualitative methods was outside the scope of this review.

Finally, while frequency of use is relevant for establishing comparability with prior research, in designing a core outcome set, rigorous evaluation of psychometric properties and feasibility of the instrument in the target population is also critical.(50) For example, the MMSE was the most commonly used measures of general intelligence (Table E4), but a recent evaluation showed it to have poor sensitivity in detecting cognitive impairment compared with detailed neuropsychological tests.(51) However, the most commonly used measure of physical activity limitation, the 6-minute walk test, was recently demonstrated to have strong psychometric properties.(38) Early results of foundational work evaluating the psychometric properties of measures of neuromuscular impairment and physical activity limitations has recently been published.(52) Unless there is considerable investment in research to evaluate the psychometric properties and performance of existing instruments, trials using these instruments as primary outcome measures should be interpreted cautiously.

In conclusion, this scoping review of outcomes after hospital discharge in survivors of critical illness identified 425 articles utilizing 250 different measurement instruments. The majority of articles described observational studies of <200 ICU survivors assessed at a single time point. While the number of studies assessing patient outcomes after discharge has grown, and measurement of participation restriction and quality of life have consolidated around dominant instruments, further work to obtain consensus around a core outcome set of valid, reliable and feasible measures is essential to advance research aimed at improving the outcomes of survivors of critical illness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge Clinical Informationist, Carrie Price, MLS for her assistance with literature searches, as well as the research staff who assisted with data collection and management, including Victor Dinglas, Krishi Nunna, Aparna Nallaganula, Deepti Baheti, Abhi Freye, Bernice Frimpong, Ayush Singh, Parker Ruhl, Mariela Pinedo, Marion Schmidt, and Abhinav Singh Verma.

Funding Support: This research was supported by the National Heart, Lung, and Blood Institute [R24HL111895].

Copyright form disclosures: Dr. Turnbull received funding (This research was supported by the National Heart, Lung, and Blood Institute [R24HL111895]) and received support for article research from the National Institutes of Health (NIH). Dr. Rabiee received support for article research from the NIH. Her institution received funding from

the NHLBI peer reviewed grant R24HL111895. Dr. Davis received support for article research from the NIH. His institution received funding from NHLBI peer reviewed grant R24HL111895 and from the NHLBI peer reviewed grant R24HL111895. Dr. Hopkins lectured for the Michigan Hospital Association (Talk on ICU outcomes at the Keystone: ICU Workshop) and received support for article research from the NIH. Her institution received grant support from the NIH (Peer reviewed grant) and from the Intermountain Research and Medical Foundation (Peer Reviewed grant). Dr. Bienvenu received support for article research from the NIH. His institution received funding from NHLBI. Dr. Robinson received support for article research from the NIH. Her institution received funding from NIH. Dr. Needham received support for article research from the NIH. His institution received funding from NIH, AHQR, and the Gordon & Betty Moore Foundation.

References

1. Carson SS. The epidemiology of critical illness in the elderly. *Crit Care Clin.* 2003; 19:605–617. [PubMed: 14601710]
2. Needham DM, Bronskill SE, Calinawan JR, et al. Projected incidence of mechanical ventilation in Ontario to 2026: Preparing for the aging baby boomers. *Crit Care Med.* 2005; 33:574–579. [PubMed: 15753749]
3. Zilberberg MD, de Wit M, Shorr AF. Accuracy of previous estimates for adult prolonged acute mechanical ventilation volume in 2020: update using 2000–2008 data. *Crit Care Med.* 2012; 40:18–20. [PubMed: 21926594]
4. Zimmerman JE, Kramer AA, Knaus WA. Changes in hospital mortality for United States intensive care unit admissions from 1988 to 2012. *Crit Care Lond Engl.* 2013; 17:R81.
5. Herridge MS, Tansey CM, Matté A, et al. Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med.* 2011; 364:1293–1304. [PubMed: 21470008]
6. Fan E, Dowdy DW, Colantuoni E, et al. Physical Complications in Acute Lung Injury Survivors: A 2-Year Longitudinal Prospective Study. *Crit Care Med.* 2013
7. Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. *N Engl J Med.* 2013; 369:1306–1316. [PubMed: 24088092]
8. Davydow DS, Desai SV, Needham DM, et al. Psychiatric morbidity in survivors of the acute respiratory distress syndrome: a systematic review. *Psychosom Med.* 2008; 70:512–519. [PubMed: 18434495]
9. Davydow DS, Gifford JM, Desai SV, et al. Depression in general intensive care unit survivors: a systematic review. *Intensive Care Med.* 2009; 35:796–809. [PubMed: 19165464]
10. Parker AM, Sricharoenchai T, Raparla S, et al. Posttraumatic Stress Disorder in Critical Illness Survivors: A Metaanalysis. *Crit Care Med.* 2015
11. Dowdy DW, Eid MP, Sedrakyan A, et al. Quality of life in adult survivors of critical illness: a systematic review of the literature. *Intensive Care Med.* 2005; 31:611–620. [PubMed: 15803303]
12. Dowdy DW, Eid MP, Dennison CR, et al. Quality of life after acute respiratory distress syndrome: a meta-analysis. *Intensive Care Med.* 2006; 32:1115–1124. [PubMed: 16783553]
13. Angus DC, Carlet J. 2002 Brussels Roundtable Participants: Surviving intensive care: a report from the 2002 Brussels Roundtable. *Intensive Care Med.* 2003; 29:368–377. [PubMed: 12536269]
14. Angus DC, Mira JP, Vincent JL. Improving clinical trials in the critically ill. *Crit Care Med.* 2010; 38:527–532. [PubMed: 19851096]
15. Spragg RG, Bernard GR, Checkley W, et al. Beyond mortality: future clinical research in acute lung injury. *Am J Respir Crit Care Med.* 2010; 181:1121–1127. [PubMed: 20224063]
16. Lieu TA, Au D, Krishnan JA, et al. Comparative effectiveness research in lung diseases and sleep disorders: recommendations from the National Heart, Lung, and Blood Institute workshop. *Am J Respir Crit Care Med.* 2011; 184:848–856. [PubMed: 21965016]
17. Needham DM, Davidson J, Cohen H, et al. Improving long-term outcomes after discharge from intensive care unit: report from a stakeholders' conference. *Crit Care Med.* 2012; 40:502–509. [PubMed: 21946660]
18. Deutschman CS, Ahrens T, Cairns CB, et al. Multisociety task force for critical care research: key issues and recommendations. *Am J Respir Crit Care Med.* 2012; 185:96–102. [PubMed: 22210788]

19. Carson SS, Goss CH, Patel SR, et al. An Official American Thoracic Society Research Statement: Comparative Effectiveness Research in Pulmonary, Critical Care, and Sleep Medicine. *Am J Respir Crit Care Med.* 2013
20. Needham DM. Understanding and improving clinical trial outcome measures in acute respiratory failure. *Am J Respir Crit Care Med.* 2014; 189:875–877. [PubMed: 24735026]
21. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials.* 2007; 8:39. [PubMed: 18039365]
22. Williamson PR, Altman DG, Blazeby JM, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials.* 2012; 13:132. [PubMed: 22867278]
23. Idzerda L, Rader T, Tugwell P, et al. Can we decide which outcomes should be measured in every clinical trial? A scoping review of the existing conceptual frameworks and processes to develop core outcome sets. *J Rheumatol.* 2014; 41:986–993. [PubMed: 24584917]
24. Daudt HM. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC Med Res Methodol.* 2013; (13):48. [PubMed: 23522333]
25. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci.* 2010; 5:69. [PubMed: 20854677]
26. Colquhoun HL, Levac D, O'Brien KK, et al. Scoping reviews: time for clarity in definition, methods, and reporting. *J Clin Epidemiol.* 2014; 67:1291–1294. [PubMed: 25034198]
27. World Health Organization. *International Classification of Functioning, Disability and Health.* Switzerland: World Health Organization; 2001.
28. Iwashyna TJ, Netzer G. The burdens of survivorship: an approach to thinking about long-term outcomes after critical illness. *Semin Respir Crit Care Med.* 2012; 33:327–338. [PubMed: 22875378]
29. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005; 8:19–32.
30. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009; 42:377–381. [PubMed: 18929686]
31. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand.* 1983; 67:361–370. [PubMed: 6880820]
32. Herrmann C. International experiences with the Hospital Anxiety and Depression Scale—a review of validation data and clinical results. *J Psychosom Res.* 1997; 42:17–41. [PubMed: 9055211]
33. *Assessing Psychological Trauma and PTSD a Practitioner's Handbook.*
34. Bienvenu OJ, Williams JB, Yang A, et al. Posttraumatic stress disorder in survivors of acute lung injury: evaluating the Impact of Event Scale-Revised. *Chest.* 2013; 144:24–31. [PubMed: 23699588]
35. Schelling G, Stoll C, Haller M, et al. Health-related quality of life and posttraumatic stress disorder in survivors of the acute respiratory distress syndrome. *Crit Care Med.* 1998; 26:651–659. [PubMed: 9559601]
36. Stoll C, Kapfhammer HP, Rothenhäusler HB, et al. Sensitivity and specificity of a screening test to document traumatic experiences and to diagnose post-traumatic stress disorder in ARDS patients after intensive care treatment. *Intensive Care Med.* 1999; 25:697–704. [PubMed: 10470573]
37. Holland AE, Spruit MA, Troosters T, et al. An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. *Eur Respir J.* 2014; 44:1428–1446. [PubMed: 25359355]
38. Chan KS, Pfoh ER, Denehy L, et al. Construct validity and minimal important difference of 6-minute walk distance in survivors of acute respiratory failure. *Chest.* 2015; 147:1316–1326. [PubMed: 25742048]
39. Katz S, Downs TD, Cash HR, et al. Progress in development of the index of ADL. *The Gerontologist.* 1970; 10:20–30. [PubMed: 5420677]
40. Katz S, Akpom CA. A measure of primary sociobiological functions. *Int J Health Serv Plan Adm Eval.* 1976; 6:493–508.

41. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *The Gerontologist*. 1969; 9:179–186. [PubMed: 5349366]
42. Chrispin PS, Scotton H, Rogers J, et al. Short Form 36 in the intensive care unit: assessment of acceptability, reliability and validity of the questionnaire. *Anaesthesia*. 1997; 52:15–23. [PubMed: 9014540]
43. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy Amst Neth*. 1990; 16:199–208.
44. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care*. 2005; 43:203–220. [PubMed: 15725977]
45. Black NA, Jenkinson C, Hayes JA, et al. Review of outcome measures used in adult critical care. *Crit Care Med*. 2001; 29:2119–2124. [PubMed: 11700407]
46. Hennessy D, Juzwishin K, Yergens D, et al. Outcomes of elderly survivors of intensive care: a review of the literature. *Chest*. 2005; 127:1764–1774. [PubMed: 15888857]
47. Blackwood, Ba; Marshall, Jb; Rose, Lc. Progress on core outcome sets for critical care research. *Curr Opin Crit Care*. 2015; 21 Review.
48. Needham DM, Wozniak AW, Hough CL, et al. Risk Factors for Physical Impairment after Acute Lung Injury in a National, Multi-Center Study. *Am J Respir Crit Care Med*. 2014
49. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008; 61:344–349. [PubMed: 18313558]
50. Boers M, Kirwan JR, Wells G, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol*. 2014; 67:745–753. [PubMed: 24582946]
51. Pfoh ER, Chan KS, Dinglas VD, et al. Cognitive screening among acute respiratory failure survivors: a cross-sectional evaluation of the Mini Mental State Examination. *Crit Care Lond Engl*. 2015; 19:220.
52. Parry SM, Granger CL, Berney S, et al. Assessment of impairment and activity limitations in the critically ill: a systematic review of measurement instruments and their clinimetric properties. *Intensive Care Med*. 2015
53. Higgins, J.; Green, S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*. 2011. Internet Available from: www.cochrane-handbook.org

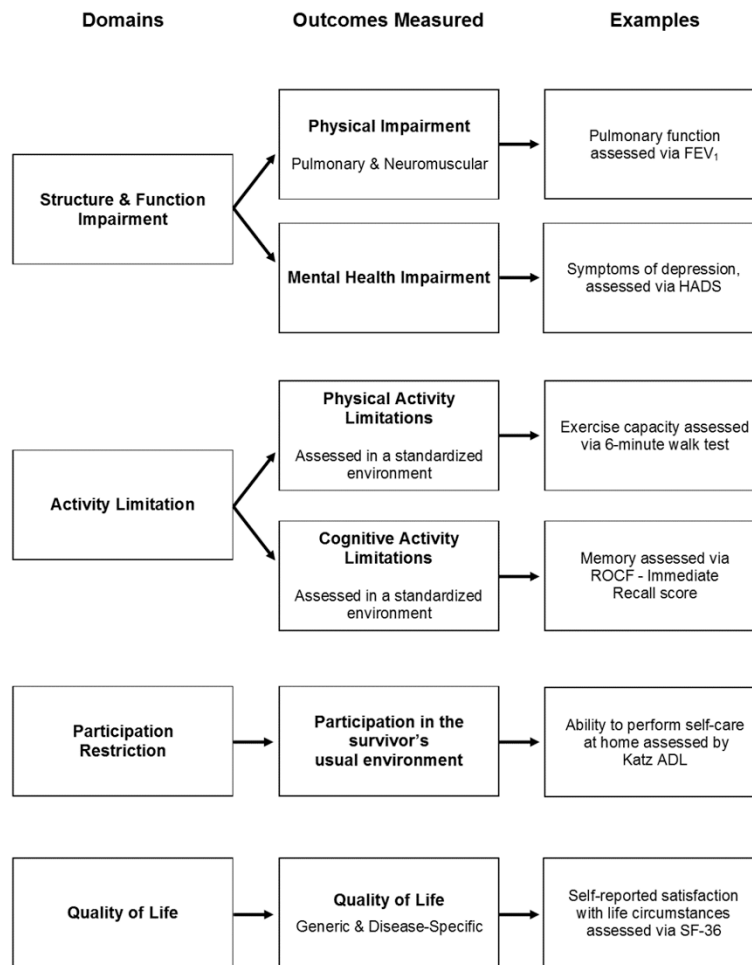


Figure 1. Categorization of outcome measures and domains
 Physical, cognitive, and mental health outcomes were categorized using the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) framework. Quality of life outcomes were added as a fourth domain, consistent with previous recommendations.(28)

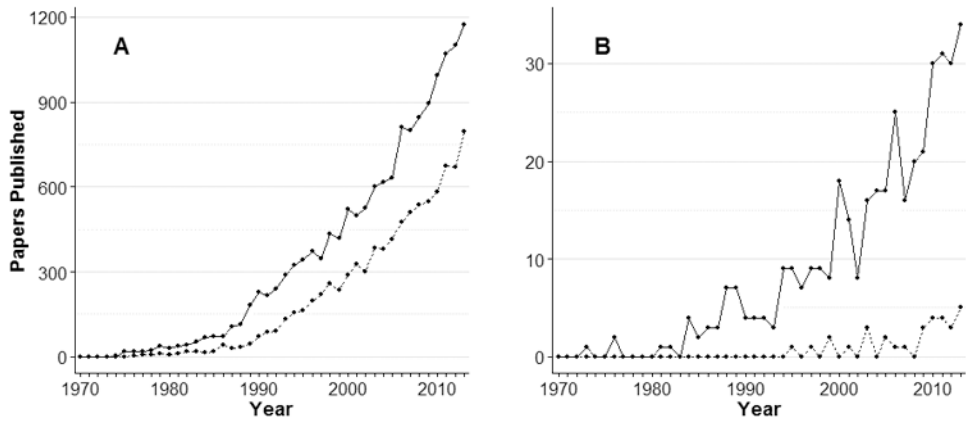


Figure 2. Critical care citations between 1970 and 2013

Panel A displays all critical care citations in PubMed, with the *solid line* excluding randomized trials, and the *dashed line* indicating randomized controlled trials as identified using an augmented version of the Cochrane Highly Sensitive Search Strategy. (53) Search terms are available in Table E1 and E2 of the Online Data Supplement. Panel B displays studies of ICU survivor outcomes included in this scoping review, with the *solid line* excluding randomized controlled trials and the *dashed line* indicating randomized controlled trials.

Table 1
Methodological characteristics of 425 studies measuring ICU survivor outcomes (1970 - 2013)

	Cross-sectional studies* N=116 (27%)	Cohort studies with 1 follow-up assessment* N=168 (39%)	Cohort studies with >1 follow-up assessment* N=110 (26%)	Randomized Controlled Trials N=31 (7%)
Months from discharge to last assessment: median (IQR) [†]	31 (18 - 52)	6 (6 - 12)	12 (12 - 12)	12 (5 - 12)
No. follow-up assessments: median (IQR)			2 (2 - 3)	1 (1 - 2)
No. participants assessed at last follow-up time: median (IQR) [‡]	65 (37 - 125)	107 (53 - 255)	80 (46 - 146)	87 (32 - 199)
Loss to follow-up reported [§]	80 (69%)	129 (77%)	40 (36%)	20 (65%)
Loss to follow-up: median (IQR) [§]	22% (11% - 38%)	22% (7% - 37%)	14% (2% - 25%)	13% (6% - 25%)

* In cross-sectional studies, participants were defined as eligible at the time of assessment. In cohort studies, eligible patients were defined as patients eligible for a follow-up assessment at any point in the study.

[†]Time from discharge to assessment was reported using available data (i.e. means, when mean not reported, we used medians or calculated a value based on range as follows: (maximum response time - minimum response time)/2), with 4 cross-sectional and 2 cohort studies with 1 follow-up assessment not reporting time to assessment.

[‡]Number of participants assessed was not reported for 7 (6%) cross-sectional studies, 8 (5%) cohort studies with 1 follow-up assessment, 14 (13%) cohort studies with 1 follow-up assessments, and 3 (10%) randomized clinical trials.

[§]Patients known to be deceased at the time of assessment were not considered lost to follow-up.

Table 2

Measurements of Physical Impairment*

Papers with any assessment of physical structure or function impairment	1970 - 2013 N = 38 32 (84%)	1970 - 1999 N = 13 12 (92%)	2000 - 2013 N = 25 20 (80%)
Any pulmonary impairment measurement			
Papers using >1 test to assess pulmonary impairment	30 (94%)	10 (83%)	20 (100%)
Article to instrument ratio †	4.0	2.0	2.5
Tests used to assess pulmonary impairment			
Forced Expiratory Volume in 1 Second (FEV1)	31 (97%)	11 (92%)	20 (100%)
Forced vital capacity (FVC)	29 (91%)	9 (75%)	20 (100%)
Carbon Monoxide Diffusion Capacity (DLCO)	24 (75%)	7 (58%)	17 (85%)
Total Lung Capacity (TLC)	21 (66%)	5 (42%)	16 (80%)
Other spirometry measures	19 (59%)	6 (50%)	13 (65%)
Other lung volumes	7 (22%)	2 (17%)	5 (25%)
Other pulmonary tests‡	2 (6%)	0 (0%)	2 (10%)
Any neuromuscular impairment measurement	7 (18%)	1 (8%)	6 (24%)
Papers using >1 instrument to assess neuromuscular impairment	3 (43%)	0 (0%)	3 (50%)
Article to instrument ratio	1.4	1.0	1.2
Instruments used to assess neuromuscular impairment			
Electromyography/Nerve Conduction Studies (EMG/NCS)	4 (57%)	1 (100%)	3 (50%)
Manual Muscle Testing (MMT) using Medical Research Council (MRC) score	3 (43%)	0 (0%)	3 (50%)
Grip Strength	2 (29%)	0 (0%)	2 (33%)
Other neuromuscular tests‡	2 (29%)	0 (0%)	2 (33%)

* No study included data on pre-hospitalization assessment of structure and function

† The article to instrument ratio is the quotient of the number of articles to the number of unique measurement instruments. A higher ratio indicates greater consolidation around a core set of measures.

‡ Represents an instrument used in only 1 eligible article

Table 3
Assessments of Mental Health Impairment from 1970 – 2013 (N=114*)

Assessment	No. (%)
Papers assessing >1 aspect of mental health impairment (e.g., anxiety & PTSD)	65 (57%)
Papers using >1 instrument to assess mental health impairment [‡]	73 (63%)
Article to instrument ratio [‡]	2.9
Assessment of baseline (pre-hospitalization) mental health impairment [§]	22 (19%)
Using same instrument as during follow-up assessment	11 (50%)
Baseline assessment obtained via:	
Patient interview	17 (77%)
Proxy interview	4 (18%)
Chart review	2 (9%)
Not Reported	3 (14%)
Any measure of depressive symptoms^{///} (N = 77, Article to instrument ratio = 8.6)	
Hospital Anxiety and Depression Scale (HADS): depression subscale	46 (60%)
Beck Depression Inventory (BDI or BDI-II)	14 (18%)
Center for Epidemiological Study - Depression (CES-D)	9 (12%)
Patient Health Questionnaire (PHQ-9)	3 (4%)
Other named instruments assessing depressive symptoms ^{****}	7 (9%)
Post-traumatic stress disorder symptoms^{///} (N = 70, Article to instrument ratio = 4.7)	
Impact of Event Scale (IES)	26 (37%)
Post-Traumatic Stress Syndrome 10-Questions (PTSS-10)	17 (24%)
Impact of Event Scale – Revised (IES-R)	13 (19%)
Clinician-Administered PTSD Scale for DSM-IV (CAPS)	8 (11%)
Symptom Check list-90-R (SCL-90-R®)	5 (7%)
PTSD Checklist-Civilian Version (PCL-C)	5 (7%)
Post-Traumatic Stress Syndrome 14-Questions (PTSS-14)	4 (6%)
Post-traumatic Stress Diagnostic Scale (PDS)	4 (6%)
Other named instruments assessing PTSD symptoms ^{****}	7 (10%)
Any measure of anxiety symptoms^{///} (N=57, Article to instrument ratio = 14.3)	
Hospital Anxiety and Depression Scale (HADS): anxiety subscale	42 (74%)
Beck Anxiety Inventory (BAI)	8 (14%)
State-Trait Anxiety Inventory (STAI)	7 (12%)
Other named instruments assessing anxiety symptoms ^{****}	1 (2%)
Any measure of other psychiatric impairments^{///} (N = 22, Article to instrument ratio = 2.0)	
ICU Memory Tool	14 (64%)

Assessment	No. (%)
Other named instruments ^{****}	10 (46%)

Abbreviations: PTSD, Post-traumatic Stress Disorder; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th Edition; ICU, Intensive Care Unit

* 11 papers were published prior to 2000. 3 used the Center for Epidemiological Study – Depression (CES-D), 3 used the Post-Traumatic Stress Syndrome 10-Questions (PTSS-10), 2 used the Impact of Event Scale – Revised (IES-R), and 1 each used the Beck Anxiety and Depression inventories, the State-Trait Anxiety Inventory (STAI), the Symptom Checklist-90-R (SCL-90-R®), and the geriatric depression scale.

[†] 32 papers used two instruments and 41 used >2 instruments

[‡] The article to instrument ratio is the quotient of the number of articles to the number of unique measurement instruments. A higher ratio indicates greater consolidation around a core set of measures.

[§] All baseline measures were obtained retrospectively except for one study of solid organ transplant patients.

// Categories are not mutually exclusive

** Represents instruments used in <3 eligible articles

Table 4
Assessments of Cognitive Activity Limitation from 1970 - 2013* (N = 40)

Assessment	No. (%)
Papers using >1 instrument to assess cognitive activity limitations	30 (75%)
Cognitive Domains measured †	
General intelligence	37 (93%)
Executive function	22 (55%)
Memory	19 (48%)
Attention	18 (45%)
Language / verbal fluency or production / naming	17 (43%)
Visuospatial Construction or Ability	16 (40%)
Working memory	16 (40%)
Mental Processing Speed	12 (30%)
Alertness	2 (5%)

* Three (8%) papers were published prior to 2000

† Domains are not mutually exclusive

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Assessments of Participation Restriction

Table 5

Assessment	1970 - 2013	1970 - 1999	2000 - 2013
	(N = 196)	(N = 65)	(N = 131)
Papers using >1 instrument to assess participation restriction	77 (39%)	28 (43%)	49 (37%)
Article to instrument ratio*	2.8	1.5	3.4
Assessment of baseline (pre-hospitalization) participation restriction ^{††}	84 (43%)	35 (54%)	49 (37%)
Baseline assessment obtained via:			
Patient interview	79 (40%)	31 (48%)	48 (37%)
Proxy interview	42 (21%)	17 (26%)	25 (19%)
Chart review	10 (5%)	4 (6%)	6 (5%)
Not reported or unclear	18 (9%)	3 (5%)	15 (12%)
Instruments used to assess participation restriction			
Return to work	87 (44%)	31 (48%)	56 (53%)
Katz Activities of Daily Living (ADL)	41 (21%)	10 (15%)	31 (24%)
Glasgow Outcome Scale (GOS)	16 (8%)	4 (6%)	12 (9%)
Lawton Instrumental Activities of Daily Living (IADL)	15 (8%)	1 (2%)	14 (11%)
Karnofsky Performance Status Scale	11 (6%)	4 (6%)	7 (5%)
Barthel Index (BI)	11 (6%)	1 (2%)	10 (8%)
Functional Independence Measure (FIM) ^{‡‡}	6 (3%)	1 (2%)	5 (4%)
New York Heart Association (NYHA) Functional Classification	5 (3%)	2 (3%)	3 (2%)
Cerebral Performance Category (CPC) Scale	5 (3%)	0 (0%)	5 (4%)
Modified Rankin Scale (MRS)	3 (2%)	0 (0%)	3 (2%)
Other Named Instruments [‡]	32 (16%)	11 (17%)	21 (16%)
Custom-made Instrument	28 (14%)	23 (35%)	5 (4%)

*The article to instrument ratio is the quotient of the number of articles to the number of unique measurement instruments. A higher ratio indicates greater consolidation around a core set of measures.

[†]Excludes return to work, all but 2 papers used the same instrument used during follow-up

[‡]Represents instruments used in <3 eligible articles

Table 6
Assessments of Quality of Life

	1970 - 2013 (N = 276)	1970 - 1999 (N = 54)	2000 - 2013 (N = 222)
Papers using >1 instrument to assess quality of life	45 (16%)	12 (22%)	33 (15%)
Article to instrument ratio*	4.8	1.9	5.8
Assessment of baseline (pre-hospitalization) quality of life [†]	80 (29%)	20 (37%)	60 (27%)
Baseline quality of life assessment obtained via: [‡]			
Patient interview	68 (25%)	18 (33%)	50 (23%)
Proxy interview	44 (16%)	8 (15%)	36 (16%)
Chart review	3 (1%)	0 (0%)	3 (1%)
Not reported or unclear	4 (1%)	2 (4%)	1 (<1%)
Generic Quality of Life	258 (94%)	51 (94%)	207 (93%)
Short Form 36 (SF-36) [§]	151 (55%)	12 (22%)	139 (63%)
EQ-5D	43 (16%)	0 (0%)	43 (19%)
Sickness Index Profile (SIP)	23 (8%)	12 (22%)	11 (5%)
Nottingham Health Profile (NHP)	17 (6%)	6 (11%)	11 (5%)
Spitzer Quality of Life Index	11 (4%)	8 (15%)	3 (1%)
Fernandez's scale	8 (3%)	2 (4%)	6 (3%)
Rosser Index	5 (2%)	5 (9%)	0 (0%)
Visual Analogue Scale	6 (2%)	1 (2%)	5 (2%)
EORTC QLQ-C30	4 (1%)	0 (0%)	4 (2%)
Other Named Instruments****	22 (8%)	8 (15%)	14 (6%)
Custom-made Instrument	18 (7%)	11 (20%)	7 (3%)
Disease-Specific Quality of Life	18 (7%)	3 (6%)	15 (7%)
St. George's Respiratory Questionnaire	10 (4%)	1 (2%)	9 (4%)
Other Named Instruments****	7 (8%)	2 (15%)	5 (6%)

Abbreviation: European Organization for Research and Treatment of Cancer, Core Quality of Life Questionnaire, EORTC QLQ-C30

* The article to instrument ratio is the quotient of the number of articles to the number of unique measurement instruments. A higher ratio indicates greater consolidation around a core set of measures.

[†] 98% of papers assessed baseline quality of life using the same instrument used during follow-up

[‡] Categories are not mutually exclusive

[§] 3 papers used Rand-36 scoring

^{||} Two papers used the EQ-5D-5L version first developed in 2007, while the rest used EQ-5D-3L

** These instruments were used in < 3 studies in the review