A systematic review of measurements of physical function in critically ill adults

Claire J Tipping, Paul J Young, Lorena Romero, Manoj K Saxena, Joel Dulhunty and Carol L Hodgson

Existing clinical trial data suggest that interventions aimed at encouraging early activity and mobilisation in the intensive care unit may reduce duration of mechanical ventilation and ICU length of stay, while improving physical function in survivors.1 However, the applicability of these data to the Australian and New Zealand setting is uncertain. A program of research evaluating early activity and mobility in the ICU is therefore underway in Australia and New Zealand, commencing with an observational study of current mobilisation practices in ventilated ICU patients (ClinicalTrials.gov NCT01674608). Our intention is to follow the observational study with a multicentre Phase II trial.

The Australian and New Zealand Intensive Care Society Clinical Trials Group recently published a consensus statement on the use of end points for Phase II trials involving critically ill patients.2 This statement highlights the importance of end points that measure limitations in physical function (so-called functional end points) and suggests that such end points could be used in trials investigating mobility interventions in the ICU in particular.

The aim of this systematic review was to identify, describe and evaluate measurements of physical function that have been used to assess early mobilisation in critically ill adults.3 A secondary aim was to evaluate the available evidence for the measurement properties and risk of bias associated with the functional end points.3,5 These data may assist in the design of Phase II trials of early activity and mobility interventions and may also have relevance to evaluating the quality of survival after critical illness.

Methods
This review had two parts: Search Strategy 1 identified studies of ICU rehabilitation or mobilisation that measured physical function; and Search Strategy 2 identified studies that examined the measurement properties and risk of bias associated with the identified end points when used with ICU patients.3,5 These data may assist in the design of Phase II trials of early activity and mobility interventions and may also have relevance to evaluating the quality of survival after critical illness.

Search Strategy 1
A literature search of Ovid MEDLINE, Embase, CINAHL, Cochrane Library and PEDro (Physiotherapy Evidence Database) was conducted in June 2012 using the search strategy outlined in Table 1. Additional studies were identified through reference and citation tracking, personal communication with a content expert, and by contacting authors of eligible trials. Review papers and meta-analyses were reviewed for publications that may otherwise have been missed. Two independent reviewers extracted data relating to patient and hospital outcomes and assessed the methodological quality of included studies.
Inclusion criteria
Eligible studies were prospective randomised controlled trials (RCTs) or prospective controlled clinical trials that: investigated mobilisation or activity-related rehabilitation in the ICU setting; reported a measurement of physical function in the ICU, such as mobility, activity, changing of body position, strength, balance, or quality of life with functional status; and were published in the past 10 years.

Exclusion criteria
Studies were excluded if they were conducted in a paediatric population (age < 16 years), long-term weaning units (or intermediate ICUs requiring long-term ventilation) or a post-ICU rehabilitation setting, or if the article was not published in English. Studies of musculoskeletal interventions other than early mobilisation or functional rehabilitation (eg, electrical stimulation or inspiratory muscle training) and those that did not include at least one measure of physical function were excluded.

Data extraction
For all studies meeting the inclusion criteria and none of the exclusion criteria, the type of study, sample size, median (or mean) patient age, measure of physical function, timing of measurement, and effect size were extracted independently by two of us (C T, C H). Measures of physical function were classified into domains based on the World Health Organization classification.

Quality assessment of studies
For RCTs meeting the inclusion criteria, the quality criteria were extracted.
reported in each RCT, methodological quality was assessed using the COSMIN (Consensus-based standards for the selection of health status measurement instruments) guidelines.3-5

Search Strategy 2
A further literature search of Ovid MEDLINE was conducted using the terms “critical illness”, “critical care”, intensive care unit”, “intensive care”, “mechanical ventilation”,

<table>
<thead>
<tr>
<th>Study</th>
<th>Study type</th>
<th>n</th>
<th>Patient age* (mean)</th>
<th>Physical function outcome measure</th>
<th>Primary outcome</th>
<th>Timing of assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Titsworth 2012</td>
<td>Prospective cohort study</td>
<td>166</td>
<td>59</td>
<td>I-MOVE</td>
<td>Yes</td>
<td>Hourly</td>
</tr>
<tr>
<td>Winkelman 2012</td>
<td>Prospective study</td>
<td>75</td>
<td>66.6 (mean)</td>
<td>Katz ADL scale</td>
<td>No</td>
<td>ICU DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRC scale (muscle strength)</td>
<td>No</td>
<td>ICU DC</td>
</tr>
<tr>
<td>Nordon-Craft 2011</td>
<td>Case series</td>
<td>19</td>
<td>48</td>
<td>FIM scale (three items)</td>
<td>ns</td>
<td>Baseline and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Timed up-and-go</td>
<td>ns</td>
<td>Baseline and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Five-times sit-to-stand test</td>
<td>ns</td>
<td>Baseline and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2-minute walk test</td>
<td>ns</td>
<td>Baseline and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRC scale (muscle strength)</td>
<td>ns</td>
<td>Baseline and hospital DC</td>
</tr>
<tr>
<td>Needham 2010</td>
<td>Prospective before–after quality improvement project</td>
<td>57</td>
<td>52</td>
<td>Functional mobility during PT/OT session</td>
<td>No</td>
<td>Daily</td>
</tr>
<tr>
<td>Zanni 2010</td>
<td>Prospective observational cohort study</td>
<td>32</td>
<td>49</td>
<td>FSS-ICU</td>
<td>No</td>
<td>ICU and hospital DC †</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Duration of unsupported sitting on edge of bed</td>
<td>No</td>
<td>ICU and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ambulation (max distance)</td>
<td>No</td>
<td>ICU and hospital DC †</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRC scale (muscle strength)</td>
<td>No</td>
<td>ICU initial assessment and ICU and hospital DC</td>
</tr>
<tr>
<td>Burtin 2009</td>
<td>Randomised controlled trial</td>
<td>90</td>
<td>57 (mean)</td>
<td>6-minute walk test</td>
<td>Yes</td>
<td>Hospital DC †</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SF-36</td>
<td>No</td>
<td>Hospital DC †</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Berg Balance Scale</td>
<td>No</td>
<td>ICU and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FAC</td>
<td>No</td>
<td>ICU and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Handheld dynamometer (quadriceps force)</td>
<td>No</td>
<td>ICU and hospital DC</td>
</tr>
<tr>
<td>Schweickert 2009</td>
<td>Randomised controlled trial</td>
<td>104</td>
<td>57.7 (mean)</td>
<td>Return to functional independence (six activities from Katz ADL scale, and mobility using FIM scale)</td>
<td>Yes</td>
<td>Every 48 hours ICU and hospital discharge †</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Barthel Index</td>
<td>No</td>
<td>Hospital DC †</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ambulation (max distance)</td>
<td>No</td>
<td>Hospital DC †</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Handheld dynamometer (hand grip strength)</td>
<td>No</td>
<td>ICU and hospital DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRC scale (manual muscle testing)</td>
<td>No</td>
<td>Hospital DC</td>
</tr>
<tr>
<td>Morris 2008</td>
<td>Prospective cohort study</td>
<td>330</td>
<td>54.7</td>
<td>Days until first out of bed</td>
<td>No</td>
<td>Daily</td>
</tr>
<tr>
<td>Thomsen 2008</td>
<td>Pre–post cohort study</td>
<td>104</td>
<td>57.9 (mean)</td>
<td>Functional mobility during PT/OT session</td>
<td>No</td>
<td>Daily</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ambulation (yes/no, distance)</td>
<td>Yes</td>
<td>Daily</td>
</tr>
<tr>
<td>van der Schaaf 2008</td>
<td>Prospective observational cohort study</td>
<td>109</td>
<td>60</td>
<td>Barthel Index</td>
<td>Yes</td>
<td>3–7 days after ICU DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Handheld dynamometer (hand grip strength)</td>
<td>No</td>
<td>3–7 days after ICU DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FAC</td>
<td>No</td>
<td>3–7 days after ICU DC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MRC scale (muscle strength)</td>
<td>No</td>
<td>ICU DC</td>
</tr>
<tr>
<td>Bailey 2007</td>
<td>Feasibility study</td>
<td>103</td>
<td>63</td>
<td>Functional mobility during PT/OT session</td>
<td>No</td>
<td>Daily</td>
</tr>
</tbody>
</table>

ADL = activities of daily living. DC = discharge. FAC = Functional Ambulation Classification. FIM = Functional Independence Measure. FSS-ICU = Functional Status Score for the ICU. I-MOVE = Independent Mobility Validation Examination. max = maximum. MRC = Medical Research Council. ns = not specified. OT = occupational therapy. PT = physiotherapy. SF-36 = Short Form 36. * Median age, unless otherwise specified. † Statistical significance (P < 0.05).
“artificial respiration” and the measures of physical function identified by Search Strategy 1, with a search filter reported previously for finding measurement properties of measurement instruments with a sensitivity of 93.1% and precision of 9.4%. Each title and abstract was examined for relevance by two of us (C T, C H), and the full text was reviewed if the inclusion criteria were met. Reference lists of sourced articles were manually reviewed to identify additional relevant articles. Authors were contacted directly if it was unclear whether outcome measures had been assessed for reliability and validity in the ICU.

Inclusion criteria
Articles were included if they: described a measure of physical function reported in Search Strategy 1; were conducted in an ICU population; and reported inter-rater and intra-rater reliability, face validity, structural validity, content validity, criterion validity and cross-cultural validity.

Results
Measures of physical function identified
Search Strategy 1 identified 3407 articles, of which 77 were reviewed in full. No further articles were identified on review of the reference lists in sourced articles (Figure 1). Eleven studies of ICU rehabilitation met the inclusion criteria and were included in this review (Table 2). Most were observational cohort studies assessing the feasibility of early mobility strategies in the critical care setting. There were two RCTs that assessed the effect of early mobilisation on recovery after critical illness. Quality assessment of the RCTs is presented in Table 3. Measurement properties of the primary outcome variable in each RCT are shown in Table 4.

Overall, 19 measures of physical function were identified in the 11 included studies (Table 5). In accordance with the WHO classification system, they were classified into domains that evaluate (i) mobility: including balance, lying, sitting, standing, shifting the body’s centre of gravity; (ii) muscle function: strength; (iii) walking and moving: including walking independently, walking with assistance, walking short and long distances; (iv) self-care: activities of daily living (ADL) such as washing, dressing, toileting, grooming and eating; and one additional area that the WHO describes as future work, which includes self-reported quality of life (QOL).

The most commonly used measurements of physical function were part of the WHO classification of mobility — that is, the assessment of highest functional mobility in the ICU, measured either at each therapy session or daily, and the maximum distance ambulated, measured at ICU and/or hospital discharge. Measures of self-care using the Barthe Index were included in two of the 11 studies. The two RCTs used different measures of function as the primary outcome: Burtin and colleagues used the 6-minute walk test (6MWT) at hospital discharge, whereas Schweickert and colleagues used a composite measure of functional independence (using the Katz ADL scale, Functional Independence Measure [FIM] scale, and independent walking) at hospital discharge.
Table 5. Summary of measures of physical function in the included studies

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>WHO domain</th>
<th>No. of studies</th>
<th>Original population</th>
<th>Description</th>
<th>Key reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barthel Index</td>
<td>Self-care</td>
<td>2</td>
<td>Neuromuscular and musculoskeletal disorders</td>
<td>10 items rated on 2-, 3- or 4-point scale. Total score ranges from 0 (dependent) to 100 (independent). Performance can be based on observation or interview.</td>
<td>Mahoney 1965^{18}</td>
</tr>
<tr>
<td>Return to functional independence</td>
<td>Self-care</td>
<td>1</td>
<td>ICU</td>
<td>Developed for use in an ICU mobility trial. Defined by ability to independently perform six Katz ADL scale items (bathing, dressing, eating, toileting, transferring, and remaining continent). Uses FIM 1–6 rating scale. Patients with ADL score ≥ 5 deemed as being independent.</td>
<td>Schweickert 2009^{1}</td>
</tr>
<tr>
<td>Katz ADL scale</td>
<td>Self-care</td>
<td>1</td>
<td>Older adults in aged care facilities</td>
<td>Six-item measure assessing ADL (bathing, dressing, eating, toileting, transferring, and remaining continent). Each item is marked either 0 (total dependence) or 1 (total independence), with a total score ranging from 0 to 6.</td>
<td>Wallace 2008^{19}</td>
</tr>
<tr>
<td>FIM scale</td>
<td>Self-care</td>
<td>1</td>
<td>Rehabilitation</td>
<td>18 items rated on a scale of 1–7. Total score ranges from 7 (total dependence) to 126 (total independence). Mixture of motor, ADL and cognitive items.</td>
<td>Turner-Stokes 1999^{20}</td>
</tr>
<tr>
<td>Berg Balance Scale</td>
<td>Walking</td>
<td>1</td>
<td>Stroke</td>
<td>14 static and dynamic balance tests, each rated from 0 (unable to perform task) to 4 (able to perform task).</td>
<td>Berg 1989^{21}</td>
</tr>
<tr>
<td>Timed up-and-go</td>
<td>Walking</td>
<td>1</td>
<td>Frail elderly</td>
<td>Measures time taken to stand up from a chair and walk around a cone 3 m away and back to sit in the chair. A standardised chair (with armrest and height of 44–47 cm) should be used. Challenges balance, transfers and mobility. 1 of 19 patients was able to complete on hospital discharge.</td>
<td>Podsiadlo 1991^{22}</td>
</tr>
<tr>
<td>SF-36</td>
<td>QOL</td>
<td>1</td>
<td>Primary care</td>
<td>36-item self-completed questionnaire. Covers all aspects of health including functional status, emotional and social wellbeing, and overall evaluation of health.</td>
<td>Chrispin 1997^{23}</td>
</tr>
<tr>
<td>FSS-ICU</td>
<td>Mobility and walking</td>
<td>1</td>
<td>ICU</td>
<td>Developed for use in one ICU mobility trial. Uses FIM scoring system, 7-point scale. Five items to suit ICU mobility including ability to roll, move from lying to sitting, sit on edge of bed, move from sitting to standing, and ambulate. Total score ranges from 5 (total dependence) to 35 (total independence).</td>
<td>Thrush 2012^{24}</td>
</tr>
<tr>
<td>6-minute walk test</td>
<td>Walking</td>
<td>1</td>
<td>Cardiorespiratory patients</td>
<td>Participant walks as far as possible in 6 minutes, with no external assistance. Standardised feedback is given. Standard corridor length (20–50 m) is used. ATS recommends two tests completed &gt; 1 hour apart. ICU patients were not able to complete two tests due to fatigue.</td>
<td>ATS statement 2002^{25}</td>
</tr>
<tr>
<td>2-minute walk test</td>
<td>Walking</td>
<td>1</td>
<td>Older adults</td>
<td>Participant walks as far as possible, with or without a gait aid, in 2 minutes, using a 50 m corridor. Assessor is able to provide verbal encouragement but no physical assistance. 1 of 19 patients was able to complete on hospital discharge.</td>
<td>Brooks 2004^{26}</td>
</tr>
<tr>
<td>FAC</td>
<td>Walking</td>
<td>2</td>
<td>Stroke</td>
<td>Assesses ability to mobilise on indoor and outdoor terrains. Performance rated on a scale from 1 (non-functional ambulatory) to 6 (independent ambulatory).</td>
<td>Pohl et al 2002^{27}</td>
</tr>
</tbody>
</table>

Quality assessment

Search Strategy 2 identified 77 articles, of which six were assessed in full (Table 6).23,32-36 Three of the functional outcome measures in the ICU population — the Medical Research Council (MRC) scale score,32-34,37 the handheld dynamometer,33,35,38 and the Short Form 36 (SF-36)23 for health-related QOL — have been reported to have good to excellent reliability. The Independent Mobility Validation Examination (I-MOVE) has been assessed for inter-rater reliability and face validity and shown to have excellent reliability in the acute hospital setting,28 but no studies were identified in the ICU setting.

The Barthel Index has been extensively assessed and found to have very good measurement properties, including responsiveness, inter-rater reliability and face validity, in stroke and geriatric populations.39-42 No studies were found that assess the measurement properties of the Barthel Index in the ICU. Measures of physical function including the FIM scale,43 Functional Ambulation Classification (FAC),44 Berg Balance Scale21,45 and the five-times sit-to-stand test29 have been shown to have excellent reliability and validity in non-ICU patient populations.

The 6MWT has very good responsiveness, inter-rater reliability and content validity in patients with heart failure,46 community-dwelling adults, older adults,47 and cardiac and pulmonary rehabilitation settings.48,49 However, its use may not be appropriate in the critical care population during the ICU stay, because of the high level of patient acuity. The 6MWT may provide valuable information at hospital discharge, as it showed a response to change with...
a clinically significant difference between two groups in a single-centre RCT of intensive care rehabilitation. However, this study was not blinded, did not use intention-to-treat analysis, and had some loss to follow-up (Table 3).

From articles evaluating reliability and validity, two additional measures of physical function — the Physical Function ICU Test (PFIT) and the de Morton Mobility Index — were identified as having good reliability in the ICU or acute hospital setting. We did not include them in this review as they had not been published as an outcome in an ICU clinical trial, but they may be of future interest in Phase II or III studies of critically ill patients.

**Discussion**

**Key findings**

The main findings of this systematic review were that the most commonly used measurements of physical function were mobility, such as the highest functional mobility (eg, rolling, sitting, standing, walking) measured during physiotherapy or occupational therapy in the ICU, and walking, such as the maximum distance ambulated, measured at ICU or hospital discharge. As no studies have investigated the measurement properties of these measures, including inter-rater or intra-rater reliability, construct or population validity, in critically ill patients, the results relating to efficacy that are based on these end points may need to be interpreted with caution.

Reliability, validity and responsiveness have been established for the SF-36 as a measure of health-related QOL in critically ill patients, and for the MRC scale score and handheld dynamometer scores as measures of strength. The Functional Status Score for the ICU (FSS-ICU) was the only end point identified in the included studies that was specifically designed for use in the ICU. It has demonstrated clinical responsiveness in the ICU setting, but further evaluation is required to assess its construct validity, reliability and predictive ability.

**Clinical implications and significance**

The lack of data on measurement properties, including responsiveness, face validity, content validity, cross-cultural validity and inter-rater and intra-rater reliability, of the most commonly used end points in contemporary studies requires further comment. Measures that generate discrete numerical scores, such as strength of specific muscle groups (MRC scale score and handheld dynamometer), may be pragmatically attractive, but may not enable an intuitive translation into an understanding of functional ability, such as ability to sit on the edge of a bed and maintain posture, or to walk or dress. We therefore suggest that measures of physical function should be categorised into four types of

**Table 6. Reliability and validity of physical function outcome measures in the intensive care unit**

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study</th>
<th>Population</th>
<th>n</th>
<th>Method</th>
<th>Outcome</th>
<th>Statistic</th>
<th>Result</th>
<th>Main result</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF-36</td>
<td>Chrispin 1997</td>
<td>ICU</td>
<td>166</td>
<td>Patient completed with nursing staff</td>
<td>Differences in scores with age and sex</td>
<td>Reliability coefficient</td>
<td>&gt;0.75</td>
<td>Good reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Construct validity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Content validity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRC scale</td>
<td>Fan 2010</td>
<td>Recovering from critical illness</td>
<td>19</td>
<td>Various professionals</td>
<td>Inter-rater reliability</td>
<td>ICC</td>
<td>0.98</td>
<td>Excellent reliability</td>
</tr>
<tr>
<td></td>
<td>Hough 2011</td>
<td>ICU</td>
<td>34</td>
<td>Doctors</td>
<td>Inter-rater reliability</td>
<td>ICC</td>
<td>0.83</td>
<td>Good reliability</td>
</tr>
<tr>
<td>Hermans 2012</td>
<td>ICU</td>
<td>75</td>
<td>Physiotherapy</td>
<td>Inter-rater reliability</td>
<td>ICC</td>
<td>0.95</td>
<td>Excellent reliability</td>
<td></td>
</tr>
<tr>
<td>Handheld dynamometer</td>
<td>Hermans 2012</td>
<td>ICU</td>
<td>46</td>
<td>Physiotherapy</td>
<td>Inter-rater reliability</td>
<td>ICC</td>
<td>0.97</td>
<td>Excellent reliability</td>
</tr>
<tr>
<td>Baldwin 2012</td>
<td>ICU</td>
<td>17</td>
<td>Physiotherapy</td>
<td>Inter-rater reliability</td>
<td>ICC</td>
<td>0.782–0.946</td>
<td>Good reliability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hermans 2012</td>
<td>ICU</td>
<td>12</td>
<td>Physiotherapy</td>
<td>Test–retest reliability</td>
<td>P</td>
<td>0.4</td>
<td>No significant difference</td>
</tr>
<tr>
<td>Vanpee 2011</td>
<td>ICU</td>
<td>39</td>
<td>Physiotherapy</td>
<td>Inter-rater reliability</td>
<td>ICC</td>
<td>0.76–0.96</td>
<td>Good reliability</td>
<td></td>
</tr>
</tbody>
</table>

FIM = Functional Independence Measure. ICC = interclass correlation coefficient. MRC = Medical Research Council. SF-36 = Short Form 36.
end point that could broadly reflect whether they are being used for proof-of-concept (Phase II) studies or potential practice-changing (Phase III) studies. The four categories are:

1. A single numerical score that evaluates a specific physical function or muscle function, such as strength or walking distance
2. A hierarchical scale that measures some aspect of physical function (eg, for mobility: sit, stand, walk, or highest physical mobility)
3. A “composite measure” that measures the ability to perform multidimensional physical activities (such as the WHO classification of self-care$^6$)
   a. Short- to medium-term end points: FIM, FSS-ICU, PFIT
   b. Long-term end points: Barthel Index and Katz ADL scale
4. A measure of the patient’s perception of his or her physical function (SF-36).

In the context of the observational studies that have explored the feasibility of delivering early mobilisation interventions, the research agenda now needs to focus on identifying candidate end points for Phase II and Phase III RCTs, and establishing their measurement properties. This may require consensus opinions augmented by input from clinicians, patients and caregivers. End points from our categories 1 and 2 may be suited to evaluating proof-of-concept of efficacy in Phase II RCTs, to assess whether a candidate intervention has a measurable effect in the short term (ie, at or before hospital discharge). End points from categories 3 and 4 may be more suited to Phase III RCTs, as they could be used to evaluate the effect of candidate interventions on complex multidimensional activities that are more relevant to medium- or long-term outcomes (ie, at or after hospital discharge). Some end points (eg, the 6MWT) may be suited for use in both Phase II and III studies.

The potential for change to be quantified in patients with a wide range of physical function may be limited when using a single outcome measure. For example, many patients may be too unwell to perform the 6MWT during their ICU stay, but this test may provide valuable information at hospital discharge and beyond. Therefore, it may be appropriate to use a combination of outcome measures, to allow appropriate, valuable and diverse patient information to be collected throughout the ICU and hospital stay.

There are other considerations with studies of early mobilisation that measure physical function. First, the difficulty with blinding participants and therapists highlights the importance of ensuring the blinding of assessors to increase the study quality. Second, the level of sedation or arousal in critically ill patients may be an important factor when determining the inter-rater and test–retest reliability of functional end points.$^{33,55}$

Novel measures of function in the ICU

One article described the PFIT as a functional measure designed for the ICU population.$^{50}$ It measures strength, endurance, cardiovascular capacity and functional level. The PFIT was shown to have good reliability and is responsive to change in the ICU population.$^{50}$ The de Morton Mobility Index was specially designed for use in the acute hospital and is valid in this setting, but it has also been extended for use in subacute settings.$^{51,52}$

Strengths and limitations

The strengths of this systematic review include the extensive search strategy, the rigorous approach to reviewing the data, and the multidisciplinary team of researchers. However, it does have some limitations. First, the search strategy may not have captured all of the studies of measurement properties of these outcome measures in an ICU population, although every attempt was made to search for these and to contact appropriate authors. Second, measures of physical function may be used in studies unrelated to early mobilisation and rehabilitation. Third, studies that measured these outcomes with undesirable results may not have been published (publication bias). Fourth, we included the SF-36 health-related QOL as a measure of physical function, as several studies have reported the “physical function” component separately to the entire survey. Fifth, we did not include studies that investigated functional status in a rehabilitation setting after discharge from the ICU, and this is a growing area of research in patients who have been critically ill. Finally, there is no information about the relationship between functional status at ICU or hospital discharge and functional status at 3, 6 or 12 months’ follow-up. This is the key question that needs to be addressed in future studies.

Future directions

Good functional survival is a key theme in intensive care medicine, and the growth of research in this area reflects this.$^7$ However, there is a scarcity of measures of physical function that are validated and reliable in this setting, and there is a clear imperative for these. Some new measures of physical function that have been developed for use within the ICU, including the PFIT and the FSS-ICU, are of interest.

Conclusion

The feasibility of early mobilisation has been established.$^{56}$ We identified 19 measures of physical function that have been used in clinical studies, and the choice of end point
will depend on whether the aim of a study is proof of concept (Phase II) or practice-changing (Phase III). Future studies of early mobilisation and rehabilitation conducted in the ICU need to ensure that measures of physical function are practical, responsive, valid and reliable in this unique setting. Where the goal is to return the patient to his or her previous level of both physical function and QOL, global measures of physical function that include self-care or activities of daily living (Barthel Index, I-MOVE or the Katz ADL scale), or QOL measures, may be more informative than simple measures of strength, mobility or best level of activity during the day.

Competing interests
None declared.

Author details
Claire J Tipping, Senior Physiotherapist1
Paul J Young, ICU Specialist2,3
Lorena Romero, Senior Librarian1
Manoj K Saxena, ICU Specialist4,5
Joel Dulhunty, Research Fellow6,7
Carol L Hodgson, Senior ICU Physiotherapist,1 and Senior Research Fellow8
1 The Alfred Hospital, Melbourne, VIC, Australia.
2 Intensive Care Unit, Wellington Regional Hospital, Wellington, New Zealand.
3 Medical Research Institute of New Zealand, Wellington, New Zealand.
4 St George Hospital, Sydney, NSW, Australia.
5 The George Institute For Global Health, Sydney, NSW, Australia.
6 Department of Intensive Care Medicine, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia.
7 Burns, Trauma and Critical Care Research Centre, University of Queensland, Brisbane, QLD, Australia.
8 Australian and New Zealand Intensive Care Research Centre, Monash University, Melbourne, VIC, Australia.
Correspondence: carol.hodgson@monash.edu

References


36 Baldwin CE, Paratz JD, Bersten AD. Muscle strength assessment in critically ill patients with handheld dynamometry: an investigation of reliability, minimal detectable change, and time to peak force generation. J Crit Care 2012; Apr 18 [Epub ahead of print].


41 Green J, Forster A, Young J. A test-retest reliability study of the Barthel Index, the Rivermead Mobility Index, the Nottingham Extended Activities of Daily Living Scale and the Frenchay Activities Index in stroke patients. Disabil Rehabil 2001; 23: 670-6.


52 de Morton NA, Davidson M, Keating JL. Validity, responsiveness and the minimal clinically important difference for the de Morton Mobility Index (DEMMI) in an older acute medical population. BMC Geriatr 2010; 10: 72.


