Functional Status Score for the ICU: An International Clinimetric Analysis of Validity, Responsiveness, and Minimal Important Difference

Minxuan Huang, ScM[1,2]; Kitty S. Chan, PhD[3]; Jennifer M. Zanni, PT, DScPT[4]; Selina M. Parry, PhD[5]; Saint-Clair G. B. Neto, PT[6,7,8]; Jose A. A. Neto, MSc, PT[7,8]; Vinicius Z. M. da Silva, MSc, PT[9,10]; Michelle E. Kho, PT, PhD[4,11]; Dale M. Needham, FCPA, MD, PhD[1,2,4]

[1]Outcome After Critical Illness and Surgery (OACIS) Group, Johns Hopkins University School of Medicine, Baltimore, MD.

[2]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD.

[3]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD.

[4]Department of Physical Medicine and Rehabilitation, Johns Hopkins University, Baltimore, MD.

[5]Department of Physiotherapy, School of Health Sciences, University of Melbourne, Melbourne, VIC, Australia.

[6]Department of Physical Therapy, Universidade Catolica de Brasilia, Brasilia, Brazil.

[7]Department of Physical Therapy, Hospital Santa Luzia, Brasilia, Brazil.

[8]Department of Research and Education, D'OR Institute, Brasilia, Brazil.

[9]Health Sciences Program, Escola Superior de Ciencias da Saude, Brasilia, Brazil.

[10]Department of Physical Therapy, Hospital de Base do Distrito Federal, Brasilia, Brazil.

[11]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada.

Dr. Huang had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors have read and approved the final article. Drs. Needham, Chan, and Huang developed the study concept and design. Dr. Huang conducted statistical analysis, and all authors interpreted the data. Drs. Huang, Chan, Parry, and Needham drafted the article, and all authors have provided critical revisions for important intellectual content. This study was supervised by Dr. Needham.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccmjournal).

Supported, in part, by funding from the National Institutes of Health (R24HL111895).

Dr. Huang's institution received funding from the National Institutes of Health (NIH) (R24HL111895). Dr. Chan received support for this article research from the NIH. Her institution received funding from the NIH and Biogen. Dr. Kho received funding from the Canadian Institutes of Health Research and received support for this article research from the Canadian Institutes of Health Research Fellowship and Bisby Prize. She currently holds a Canada Research Chair in Critical Care Rehabilitation and Knowledge Translation. Dr. Needham received support for this article research from the NIH. His institution received funding from the NIH. The remaining authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: dale.needham@jhmi.edu

DOI: 10.1097/CCM.0000000000001949

Objectives: To evaluate the internal consistency, validity, responsiveness, and minimal important difference of the Functional Status Score for the ICU, a physical function measure designed for the ICU.

Design: Clinimetric analysis.

Settings: Five international datasets from the United States, Australia, and Brazil.

Patients: Eight hundred nineteen ICU patients.

Intervention: None.

Measurements and Main Results: Clinimetric analyses were initially conducted separately for each data source and time point to examine generalizability of findings, with pooled analyses performed thereafter to increase power of analyses. The Functional Status Score for the ICU demonstrated good to excellent internal consistency. There was good convergent and discriminant validity, with significant and positive correlations ($r = 0.30–0.95$) between Functional Status Score for the ICU and other physical function measures, and generally weaker correlations with nonphysical measures ($|r| = 0.01–0.70$). Known group validity was demonstrated by significantly higher Functional Status Score for the ICU scores among patients without ICU-acquired weakness (Medical Research Council sum score, ≥ 48 vs < 48) and with hospital discharge to home (vs healthcare facility). Functional Status Score for the ICU at ICU discharge predicted post-ICU hospital length of stay and discharge location. Responsiveness was supported via increased Functional Status Score for the ICU scores with improvements in muscle strength. Distribution-based methods indicated a minimal important difference of 2.0–5.0.

Conclusions: The Functional Status Score for the ICU has good internal consistency and is a valid and responsive measure of physical function for ICU patients. The estimated minimal important difference can be used in sample size calculations and in interpreting studies comparing the physical function of groups of ICU patients. (*Crit Care Med* 2016; XX:00–00)

Key Words: Australia; Brazil; cross-sectional studies; intensive care; reproducibility of results; United States

Critically ill patients frequently experience long-lasting impairments in physical functioning after discharge from the ICU (1–5). There is a growing body of research aimed at evaluating ICU-based interventions that may reduce these impairments and growing interest in measures of physical function for critically ill adults (6–8).

The Functional Status Score for the ICU (FSS-ICU) is a physical function measure specifically designed for the ICU that has not had comprehensive evaluation of its clinimetric performance (9, 10). The FSS-ICU includes five functional tasks (rolling, transfer from spine to sit, sitting at the edge of bed, transfer from sit to stand, and walking). Each task is evaluated using an eight-point ordinal scale ranging from 0 (not able to perform) to 7 (complete independence; **Web Table 1**, Supplemental Digital Content 1, http://links.lww.com/CCM/B952; for example, scale, instrument, and scoring details are available at http:www.ImproveLTO.com). The total FSS-ICU score ranges from 0 to 35, with higher scores indicating better physical functioning.

Our objective was to evaluate the internal consistency, construct and predictive validity, responsiveness, and minimum important difference (MID) of the FSS-ICU in ICU patients across different in-patient assessment time points and across international ICU settings.

METHODS

This analysis was conducted in accordance with the consensus-based standards for the selection of health measurement instruments guideline for evaluating the measurement properties of instruments (11).

Study Design

We performed a clinimetric evaluation of the FSS-ICU using data from five international datasets: two from the United States (9, 12), one from Australia (13, 14), and two from Brazil. All datasets were approved by the appropriate ethics review boards, and where required, informed consent was obtained.

The USA-Kho dataset ($n = 34$) was a randomized pilot trial of neuromuscular electrical stimulation (NMES) that enrolled patients requiring mechanical ventilation for less than or equal to 4 days in three medical and surgical ICUs in an academic medical center in Baltimore, MD, between 2008 and 2013 (15, 16). The randomized intervention of NMES versus a sham control group did not have a significant effect on the FSS-ICU score, so intervention and control groups were pooled for this analysis.

The USA-Needham dataset ($n = 59$) was a quality improvement (QI) project that enrolled patients requiring mechanical ventilation for more than or equal to 4 days in a single medical ICU at an academic medical center in Baltimore, MD, during 2007 (9, 12). This project used a structured QI framework to improve functional mobility via physical and occupational therapy. The QI versus pre-QI periods did not have a significant difference in the FSS-ICU score, so both periods were pooled for this analysis.

The Australia dataset ($n = 66$) included consecutive enrolled patients requiring mechanical ventilation for more than 48 hours in two mixed medical-surgical ICUs and received routine care in Melbourne, VIC, Australia, between 2012 and 2014 (13).

The Brazil-da Silva dataset ($n = 99$) included consecutive patients admitted in a single mixed (trauma, neurosurgical, and cardiovascular) ICU and received routine physical therapy (no intervention) in at a public hospital in Brasilia, Brazil, in 2014, using a Portuguese version of FSS-ICU developed with independent forward and backward language translation. The FSS-ICU data were collected as part of the routine care of physical therapy evaluation.

The Brazil-Neto dataset ($n = 561$) included consecutive patients more than or equal to 60 years old admitted in four ICUs (three medical-surgical and one surgical) and received routine physical therapy (no intervention) at a private hospital in Brasilia, Brazil, between 2013 and 2014, using a Portuguese version of FSS-ICU translated by the Brazilian investigators. The FSS-ICU data were collected as part of the routine care of physical therapy evaluation.

Study Measures

The FSS-ICU was evaluated prior to hospitalization (via proxy, evaluating the 2-mo period prior to hospitalization) and at ICU awakening, ICU discharge, and hospital discharge for both U.S. studies; at ICU awakening, ICU discharge, and hospital discharge for the Australian study; and at ICU admission and ICU discharge for both Brazilian studies.

Well-established measures of physical function, available within the datasets, were used to assess convergent and known-group validity of the FSS-ICU. These measures were the Lawton instrumental Activity of Daily Living (ADL) score (17) (range, 0–8, with higher scores indicating better status), the Katz ADL score (18) (range, 0–6, with higher scores indicating better status), manual muscle testing (MMT, using the Medical Research Council sum score, range, 0–60, with higher scores indicating greater strength, and < 48 indicating ICU-acquired weakness [ICUAW]) (19, 20), hand grip strength (kg, as percent predicted using normative data [21, 22]), ICU mobility scale (IMS; range, 0–10, with higher score indicating better mobility) (23), ICU and hospital length of stay (LOS), and hospital discharge location (home vs healthcare facility).

To assess discriminant validity, measures that were available and expected to have little to no relationship with FSS-ICU were used. These included body mass index (BMI), continence status (from ADL scale), hemodialysis status, home oxygen use at hospital discharge, steroid and insulin use on the hospital ward and at hospital discharge.

We used two outcome measures to assess predictive validity of FSS-ICU, similar to prior research (13, 24–26): post-ICU hospital LOS (i.e., number of days between ICU and hospital discharge) and hospital discharge location (home vs healthcare facility).

To assess FSS-ICU's responsiveness, changes in FSS-ICU scores across two time points (ICU awakening/admission to ICU discharge, ICU discharge to hospital discharge, and ICU awakening to hospital discharge) were evaluated and were compared with changes across the same two time points for the MMT and ADLs.

Statistical Analysis

Initially, analyses were conducted separately for each dataset and assessment time point to evaluate generalizability of these individual findings by time point, patient sample, and study setting, and then, pooled analyses across studies were performed, whenever feasible and appropriate (i.e., when there were similar results among individual datasets), to increase statistical power. All analyses were performed using Stata 13.1 (Stata, College Station, TX).

Floor and Ceiling Effects

Floor and ceiling effects were evaluated by examining the percentage of assessments with the minimum and maximum FSS-ICU scores, respectively.

Internal Consistency

Pearson correlations were used to identify pairwise correlations between the five FSS-ICU items, and Cronbach α was used to examine the internal consistency of the FSS-ICU total score (27).

Concurrent Construct Validity

We used Pearson correlations (for continuous measures) and biserial correlations (for binary measures) to examine convergent and discriminant validity. To evaluate convergent validity, we hypothesized that the measures evaluated would be at least moderately correlated ($|r| > 0.40$) with the FSS-ICU. To evaluate discriminant validity, we hypothesized that measures evaluated would have negligible to weak correlations

TABLE 1. Prehospitalization Patient Characteristics for Individual and Combined Studies

Patient Characteristics	USA-Kho ($n = 34$)	USA-Needham ($n = 59$)	Australia ($n = 66$)	Brazil-da Silva ($n = 99$)	Brazil-Neto ($n = 561$)[a]	Combined ($n \leq 819$)
Age (yr), mean (SD)	55 (16)	54 (15)	58 (17)	66 (10)	75 (9)	70 (13)
Men, *n* (%)	17 (50)	19 (32)	40 (61)	35 (35)	276 (49)	387 (47)
Body mass index (kg/m^2), mean (SD)	27 (7)	29 (11)	28 (7)			28 (8)
ADL score,[b] mean (SD)	6 (1)	5 (2)				5 (1)
Instrumental ADL score,[c] mean (SD)	6 (3)	4 (3)				5 (3)
Functional Status Score for the ICU,[d] mean (SD)	34 (4)	31 (9)				32 (8)
Acute Physiology and Chronic Health Evaluation II severity of illness,[e] mean (SD)	25 (7)	26 (7)	21 (7)	14 (7)	12 (7)	14 (8)
ICU admission diagnosis,[f] *n* (%)						
Respiratory (including pneumonia)	25 (76)	39 (66)	14 (21)	30 (30)		108 (42)
Gastrointestinal	3 (9)	5 (8)	12 (18)	8 (8)		28 (11)
Sepsis, nonpulmonary	0 (0)	3 (5)	13 (20)	18 (18)		34 (13)
Cardiovascular	2 (6)	4 (7)	18 (27)	10 (10)		34 (13)
Trauma	0 (0)	0 (0)	5 (8)	20 (20)		25 (10)
Neurologic	0 (0)	4 (7)	0 (0)	13 (13)		17 (7)
Other	3 (9)	4 (7)	4 (6)	0 (0)		11 (4)
Hospital length of stay, mean (SD)	35 (21)	31 (20)	28 (15)	19 (4)		33 (20)

ADL = activities of daily living.

[a]The Brazil-Neto study does not have ICU admission diagnosis data that fit into the above categories.

[b]Activities of daily living (ADL) score has a range of 0–6 with higher score, indicating better functional status.

[c]Instrumental ADL score has a range of 0–8 with higher score, indicating better functional status.

[d]Functional Status Score for the ICU score has a range of 0–35 with higher score, indicating better functional status.

[e]Acute Physiology And Chronic Health Evaluation II score has a range of 0–71, with higher score indicating greater severity of illness within first 24 hr of ICU admission.

[f]Percentages may not sum to 100 (%) because of rounding.

TABLE 2. Construct Validity: Cross-Sectional Relationship of Functional Status Score for the ICU With Other Outcome Measures Across Publications and Assessment Time Points

Time Point by Publication	*n*	Construct and Convergent Validity (Pearson Correlation)							
		Instrumental Activities of Daily Living	ADL	Manual Muscle Testing	Hand Grip % Predicted	Hand Grip Strength (kg)	ICU Mobility Scale[b]	ICU LOS	Hospital LOS
Prehospitalization									
USA-Kho	32	0.48[d]	0.73[d]						−0.04
USA-Needham	46–50	0.57[d]	0.80[d]						
Combined	78–82	0.55[d]	0.80[d]						
ICU awakening/admission[e]									
USA-Kho	20–29		0.48[d]	0.81[d]	0.44[d]	0.41	0.46[d]		
USA-Needham	43–52			0.72[d]					
Australia	19–66			0.62[d]	0.30	0.30			
Brazil-da Silva	99			0.38[d]					
Brazil-Neto	561			0.32[d]					
Combined	20–802		0.39[d]	0.44[d]	0.40[d]	0.37[d]	0.46[d]		
ICU discharge									
USA-Kho	12–27		0.70[d]	0.70[d]	0.43[d]	0.50[d]	0.62[d]	0.03	
USA-Needham	39–47			0.76[d]				−0.18	
Australia[f]	20–66			0.68[d]	0.62[d]	0.70[d]	0.69[d]	−0.24	
Brazil-da Silva	99						0.95[d]	−0.20[d]	
Brazil-Neto	561			0.64[d]				−0.27[d]	
Combined	27–800		0.70[d]	0.60[d]	0.50[d]	0.59[d]	0.86[d]	−0.25[d]	
Hospital discharge									
USA-Kho	12–28		0.86[d]	0.80[d]	0.46[d]	0.51[d]			−0.17
USA-Needham	15–44		0.81[d]	0.80[d]					−0.34[d]
Australia	8–19			0.39	0.51[d]	0.61[d]			−0.46[d]
Combined	31–91		0.80[d]	0.80[d]	0.43[d]	0.49[d]			−0.26[d]

ADL = activities of daily living, LOS = length of stay.

[a]Biserial correlations evaluate a correlation when one variable is dichotomous and were used to evaluate correlation with continence, hemodialysis, home oxygen, steroids, and insulin use.

[b]This Hodgson ordinal ICU mobility scale evaluates patients' highest mobility level, during physical therapy assessment in the ICU, and ranges from 0 (lying in bed) to 10 (walking independently without a gait aid).

[c]Represents any use of this medication on hospital ward (for the ICU discharge time point) and on hospital discharge (for hospital discharge time point).

[d]$p < 0.05$.

[e]ICU awakening was defined using the De Jonghe criteria in USA-Kho and Australia studies; in the USA-Needham study, it was defined based on Richmond Agitation Sedation Scale and ability to follow instructions to perform the assessment.

[f]A total of 20 patients had hand grip data, and 41 had ICU mobility scale data.

($|r| < 0.30$). We hypothesized significant negative correlations between FSS-ICU and ICU and hospital LOS. For known-group validity, we conducted two-sample *t* tests for group differences in FSS-ICU by ICUAW status (MMT ≥ 48 vs < 48) and hospital discharge location (home vs healthcare facility). We hypothesized that patients without (vs with) ICUAW or discharged to home (vs healthcare facility) would have significantly higher FSS-ICU scores.

Predictive Validity

As done in prior research (13, 24, 25), we used two-sample *t* tests and linear and logistic regression models to test the association of FSS-ICU at ICU discharge with post-ICU hospital LOS and hospital discharge location. In addition, the area under a receiver operating characteristic curve (i.e., *C* statistic) was calculated for FSS-ICU with discharge location. We hypothesized that patients with higher FSS-ICU scores at ICU

Discriminant Validity (Biserial Correlation[a])					
Body Mass Index (kg/m²)	Continence Item From ADL	Hemodialysis Status	Need for Home Oxygen	Steroid Use[c]	Insulin Use[c]
0.09					
< 0.01	0.01				
−0.03	0.03				
−0.01	0.50[d]				
0.16					
−0.17					
−0.01	0.50[d]				
0.05	0.70[d]	0.06		−0.20	0.33
0.31[d]		−0.26		0.30	0.05
−0.21					
0.05	0.70[d]	−0.20		0.19	0.11
[d]0.12	0.78[d]	0.30	−0.12	−0.15	0.33
0.11	0.29	0.14	0.22	−0.12	−0.24
−0.37					
−0.05	0.42[d]	0.38[d]	0.05	−0.16	−0.06

discharge would have a shorter post-ICU hospital LOS and be discharged to home (vs healthcare facility).

Responsiveness

Responsiveness was examined in three ways. First, we tracked FSS-ICU scores across the expected recovery trajectory. Differences in mean FSS-ICU scores between consecutive time points were tested using paired *t* tests. Second, we calculated the effect size for changes over time (mean difference in FSS-ICU scores between two time points divided by the SD at first time point) (28). Third, we evaluated change over time in the FSS-ICU relative to patients' change in MMT and ADL scores, with changes categorized as "significant improvement" if MMT and ADL scores at the later assessment were greater than or equal to 1SD higher than the earlier assessments. A comparison group was comprised of patients

whose scores increased less than 1SD or declined over the period (29).

Estimating MID

We used the following distribution-based methods to estimate MID (30, 31): SEM, minimal detectable change 90 (MDC-90), 0.2SD, and 0.5SD (32).

RESULTS

Patient Characteristics

Across the five studies, the mean (SD) age of patients ranged from 54 (15) to 75 (9) years, and the mean (SD) Acute Physiology And Chronic Health Evaluation (APACHE) II score ranged from 12 (7) to 26 (7) (**Table 1**). Both Brazilian studies had older patients and lower APACHE II scores. There was a wide range of ICU admission diagnoses across the studies, with respiratory failure being the most common primary diagnosis (42% in the combined dataset).

Floor and Ceiling Effect

Minimal floor effect was observed (0.5%, 0.3%, and 0% at ICU admission/awakening, ICU discharge, and hospital discharge, respectively). Some ceiling effect was observed later during recovery (0.7% at ICU admission/awakening and then 11% and 21% at ICU and hospital discharge, respectively).

Internal Consistency

Good to excellent internal consistency was observed (33). The correlation coefficients for pairwise correlation between FSS-ICU items were all positive and significant ($p < 0.05$) in all datasets and at all time points. Across time points, Cronbach α for each study ranged from 0.90 to 0.94 (USA-Kho), 0.94 to 0.95 (USA-Needham), 0.91 to 0.93 (Australia), 0.78 to 0.91 (Brazil-da Silva), and 0.78 to 0.93 (Brazil-Neto).

Concurrent Construct Validity

Consistently, across studies and time points (**Table 2**), we observed significant and positive correlations between FSS-ICU and other physical measures and negative association with ICU and hospital LOS. These findings support concurrent validity. Known-group validity was supported by significantly higher FSS-ICU scores among survivors without ICUAW (MMT ≥ 48 vs < 48) and among those discharged to home (vs healthcare facility) (**Web Table 2**, Supplemental Digital Content 2, http://links.lww.com/CCM/B953).

Consistent with our hypotheses, most associations were not statistically significant between FSS-ICU and BMI, hemodialysis, need for home oxygen, and steroid and insulin use. These findings support discriminant validity (Table 2).

Predictive Validity

We found evidence of predictive validity for duration of post-ICU hospital LOS in the USA-Needham study and in combined results across all studies, with significantly higher FSS-ICU scores at ICU discharge for survivors with below versus above the median post-ICU hospital LOS (**Table 3**). Linear regression analysis suggested that for a 1-unit increase in FSS-ICU score, post-ICU hospital LOS decreased by 0.27 days ($p < 0.01$) in the combined results (Table 3). Prediction of discharge location was consistently significant across studies: survivors discharged to home were associated with a higher FSS-ICU at ICU discharge (23 vs 16 in combined results; $p < 0.01$). Logistic regression indicated that for 1-unit increase in FSS-ICU score, the odds of discharge to home increased by 11% ($p < 0.01$) in combined results. The *C*-statistic for discharge location was 0.75 in the combined

TABLE 3. Predictive Validity of Functional Status Score for the ICU at ICU Discharge for Duration of Post-ICU Hospital Stay and Discharge Location

	Mean Functional Status Score for the ICU Score, Below vs Above Median Post-ICU Hospital LOS			Post-ICU Hospital LOS (Continuous)		Discharge Location			Discharge Location (Home vs Healthcare Facility)		Area Under the Receiver Operating Characteristics Curve (i.e., *C*-Statistic)
Time Point by Publication[a]	Below Median[b]	Above Median[b]	*p*[c]	Linear Regression Coefficient	*p*	Home	Healthcare Facility[d]	*p*[c]	Odds Ratio (95% CI)	*p*	
ICU discharge											
USA-Kho	23	19	0.19	−0.23	0.19	25	16	< 0.01	1.23 (1.05–1.45)	0.01	0.83
USA-Needham	22	13	< 0.01	−0.37	< 0.01	23	15	< 0.01	1.11 (1.02–1.20)	0.01	0.73
Australia	20	18	0.23	−0.24	0.08	22	16	< 0.01	1.09 (1.02–1.17)	< 0.01	0.72
Combined	21	17	< 0.01	−0.27	< 0.01	23	16	< 0.01	1.11 (1.06–1.17)	< 0.01	0.75

LOS = length of stay.

[a]Sample size by post-ICU hospital length of stay (LOS) (below median and above median): USA-Kho (12, 14), USA-Needham (21, 23), and Australia (32, 34). Sample size by discharge location (home and healthcare facility): USA-Kho (13, 15), USA-Needham (15, 29), and Australia (26, 37).

[b]Median LOS (d) in different publications: USA-Kho: 10; USA-Needham: 10; Australia: 14; and combined: 11.

[c]*p* values calculated using two-sample *t* test.

[d]Healthcare facilities include nursing home, other hospital's ICU or ward, or long-term ventilation facility.

analysis, indicating that FSS-ICU can adequately predict discharge location.

Responsiveness

Mean FSS-ICU scores at each time point are shown in **Web Figure 1** (Supplemental Digital Content 3, http://links.lww.com/CCM/B954; legend, Supplemental Digital Content 4, http://links.lww.com/CCM/B955). Consistent with the expected functional trajectory, the FSS-ICU score decreased from the baseline value prior to hospitalization to ICU admission/awakening and then increased at ICU and hospital discharge. Changes between each consecutive time point were statistically significant ($p < 0.01$). In combined analysis, the median (interquartile range) FSS-ICU score was 35 (33–35) prior to hospitalization, 5 (5–10) at ICU admission/awakening, 20 (10–30) at ICU discharge, and 29 (20–34) at hospital discharge.

Although not always statistically significant, increased FSS-ICU scores were generally observed with improvements in muscle strength (**Table 4**), supporting responsiveness. The effect size was 2.02 from ICU awakening/admission to ICU discharge, suggesting good responsiveness. Only the USA-Kho study, with data on 24–26 patients, could be used to evaluate the FSS-ICU's responsiveness to changes in ADL scores. This study showed a larger increase in FSS-ICU among survivors with more than 1-SD increase in ADL scores compared with those with negative or no change in ADL scores although this difference was significant only when comparing ICU discharge with hospital discharge (Table 4).

MID

In the combined results, MID estimates based on the SEM and 0.2SD were relatively consistent with 1.2–1.3 for ICU admission/

TABLE 4. Responsiveness to Change of Functional Status Score for the ICU Score Versus Manual Muscle Testing and Activity of Daily Living Scores

Time Point by Publication[a]	Change in FSS-ICU		Change in FSS-ICU		Effect Size for FSS-ICU
	MMT Score, Negative or No Change[b]	MMT Score, Significant Positive Change[b]	ADL Score, Negative or No Change[b]	ADL Score, Significant Positive Change[b]	
ICU awakening/admission to ICU discharge					
USA-Kho	5.8	9.8	7.1	7.5	0.84
USA-Needham	5.8[c]	12.0[c]			0.99
Australia	3.9[c]	11.9[c]			0.84
Brazil-Neto[d]	13.5[c]	16.5[c]			2.62
Combined	10.0[c]	17.5[c]	7.1	7.5	2.02
ICU discharge to hospital discharge					
USA-Kho	8.1	16.3	5.2[c]	13.8[c]	0.93
USA-Needham	5.3	8.5			0.69
Australia	13.0	7.0[e]			1.45
Combined	7.1	11.6	5.2[c]	13.8[c]	0.92
ICU awakening to hospital discharge					
USA-Kho	10.5[c]	17.8[c]	10.7	15.9	1.71
USA-Needham	8.4[c]	15.3[c]			1.78
Australia	21.4	19.7			2.51
Combined	11.1[c]	15.6[c]	10.7	15.9	1.85

ADL = activities of daily living, FSS-ICU = Functional Status Score for the ICU, MMT = manual muscle testing.

[a]Sample sizes for difference in Functional Status Score for the ICU (FSS-ICU) values by time point and outcome measure: 1) manual muscle testing (MMT) (no change and positive change): ICU awakening to ICU discharge: USA-Kho (8, 13), USA-Needham (7, 30), Australia (23, 43), and Brazil-Neto (465, 73); ICU discharge to hospital discharge: USA-Kho (3, 22), USA-Needham (6, 31), and Australia (1, 8); and ICU awakening to hospital discharge: USA-Kho (12, 15), USA-Needham (15, 22), and Australia (3, 6). 2) Activities of daily living (ADL) (no change and positive change): ICU awakening to ICU discharge: USA-Kho (2, 22); and ICU discharge to hospital discharge: USA-Kho (11, 15).

[b]The "no change" category represents all patients who did not increase by more than 1 SD at last available FSS-ICU time points; except for Brazil-Neto study, no patient had a decrease in MMT or ADL scores of more than 1 SD. The SD used for different studies: MMT: USA-Kho (7), USA-Needham (10), Australia (6), Brazil-Neto (14), Combined (8); ADL: USA-Kho (3).

[c]$p < 0.05$ by two-sample t test for comparison of "negative or no change" versus "significant positive change" categories.

[d]In Brazil-Neto study, there were patients with significant decrease (> 1SD) of MMT score from ICU admission to ICU discharge, and the change in FSS-ICU in MMT score significant negative change group was 2.9 ($n = 22$). This is significantly different from the change in FSS-ICU in MMT score no change or positive change groups.

[e]Only one patient was found in this group.

awakening, 2.1–2.4 for ICU discharge, and 1.7–1.9 for hospital discharge (**Table 5**). Estimates based on MDC-90 and 0.50SD also were consistent but larger at 3.0–3.1, 5.3–5.4, and 4.3–4.5 for the same time points, respectively. Hence, the MID is estimated to be in the range of 2.0–5.0.

DISCUSSION

Using data from five studies across three continents, we evaluated internal consistency, validity, responsiveness, and MID of FSS-ICU, an outcome measure assessing physical function in critically ill patients (7, 10, 13). We found consistent and strong evidence of internal consistency and concurrent construct validity with expected findings for convergent, discriminant, and known-group validity tests. The similarity of these clinimetric analyses across individual studies demonstrates generalizability of results and supports pooling of data and analyses across studies, as done in prior research (34–36).

The findings of convergent validity between the FSS-ICU and MMT agree with a prior smaller analysis (13). Prior studies of the FSS-ICU also provided preliminary evidence of predictive validity and responsiveness (10, 13), which were expanded in our current analyses with larger sample size and more variables. Predictive validity was supported with FSS-ICU scores at ICU discharge significantly predicting post-ICU hospital LOS and hospital discharge location. An increase in FSS-ICU score was observed with improvement in muscle strength and ADLs, and FSS-ICU scores tracked the recovery trajectory of survivors from ICU awakening/admission to hospital discharge with a large effect size, supporting responsiveness. The MID for the FSS-ICU, based on multiple distribution-based methods, is estimated within a range of 2.0–5.0. These results were similar across various time points and the 5 datasets, supporting generalizability.

The results of this evaluation should be compared with similar evaluations of other published ICU-specific physical function measures, including the Physical Function in Intensive care Test scored (PFIT-s) (13, 24, 37), Chelsea Critical Care Physical Assessment tool (CPAx) (38–40), Perme mobility scale (41, 42), Acute Care Index of Function (ACIF) score (43), Surgical ICU Optimal Mobilization Score (SOMS) (25, 26, 44), and the IMS (23). With respect to floor and ceiling effects, for the FSS-ICU, we detected a minimal floor effect (≤ 0.5%), but some ceiling effects at hospital discharge (≤ 21%), which may limit the instrument's ability to detect improvement (45). However, these findings are favorably comparable with other ICU-specific physical measures (**Web Table 3**, Supplemental Digital Content 5, http://links.lww.com/CCM/B956). The CPAx has the lowest ceiling effects at ICU discharge (39, 40); however, it is important to note that CPAx differs from other ICU-specific measures (Web Table 3, Supplemental Digital Content 5, http://links.lww.com/CCM/B956) as it involves evaluation of both physical function (whole body activities and grip strength) and respiratory (ventilation, oxygenation, and secretion clearance) measures.

TABLE 5. Minimum Important Difference for Functional Status Score for the ICU: Distribution-Based Estimates

Study (Sample Size)	USA-Kho (*n* = 27–29)	USA-Needham (*n* = 44–52)	Australia (*n* = 19–66)[a]	Brazil-da Silva (*n* = 99)	Brazil-Neto (*n* = 561)	Combined (*n* = 91–807)
SEM						
ICU awakening/admission	1.8	1.7	1.8	1.0	1.1	1.3 (807)
ICU discharge	1.9	2.1	2.0	2.0	2.4	2.4 (800)
Hospital discharge	1.6	2.1	0.9			1.9 (91)
Minimal detectable change$_{90}$						
ICU awakening/admission	4.1	3.9	4.1	2.4	2.7	3.1 (807)
ICU discharge	4.3	4.9	4.7	4.8	5.7	5.4 (800)
Hospital discharge	3.7	4.9	2.2			4.5 (91)
0.5SD (moderate Cohen effect size)						
ICU awakening/admission	3.9	3.7	4.0	2.3	2.6	3.0 (807)
ICU discharge	4.2	4.7	4.6	4.6	5.4	5.3 (800)
Hospital discharge	3.6	4.7	2.1			4.3 (91)
0.2SD (small Cohen effect size)						
ICU awakening/admission	1.6	1.5	1.6	0.9	1.0	1.2 (807)
ICU discharge	1.7	1.9	1.8	1.8	2.2	2.1 (800)
Hospital discharge	1.4	1.9	0.8			1.7 (91)

[a]Only 19 patients at hospital discharge.

For evaluation of validity, the PFIT-s, IMS, and CPAx also displayed concurrent construct validity with MMT (Web Table 3, Supplemental Digital Content 5, http://links.lww.com/CCM/B956). Similar to FSS-ICU, PFIT-s also showed construct validity with hand grip strength and IMS, and there is a strong positive correlation between FSS-ICU and PFIT-s (ρ = 0.85–0.87; $p < 0.005$) at ICU awakening and ICU discharge (13). Our analyses also demonstrated appropriate divergent validity of FSS-ICU.

For predictive validity, a higher FSS-ICU, along with higher PFIT-s, IMS, SOMS, and ACIF scores, predict shorter hospital LOS and/or discharge location to home. The PFIT-s, IMS, and CPAx also demonstrated moderate to large responsiveness to change via effect-size analyses. Although a prior study of the FSS-ICU demonstrated small responsiveness to change (effect size, 0.46) (13, 24, 37), our current analysis demonstrated a large effect size (2.02) for FSS-ICU from ICU awakening/admission to ICU discharge, suggesting good responsiveness.

There is growing interest in identifying a core set of outcome measures, which can be used across the continuum of recovery to measure response to interventions and monitor functional improvement. The FSS-ICU is a robust tool, which can be used to evaluate physical function in both the ICU setting and in the acute hospital setting for ICU survivors. The ability of FSS-ICU to be used in longer term follow-up beyond acute hospitalization may be impacted by a ceiling effect. It is also important to consider clinical utility: the FSS-ICU takes 10 to 30 minutes to complete (depending on patient's functional status), requires no additional equipment, and can be undertaken by the therapist at the bedside with standardized instructions readily available and thus can be easily integrated into routine critical care practice.

The strengths of our study include performing a range of clinimetric analyses using five international datasets with relatively large combined sample size (n = 819). Given that many of our findings were consistent across these datasets with different study designs, patient populations, and time points, generalizability of our findings is supported. However, there are potential limitations. First, we only assessed internal consistency of the FSS-ICU and did not evaluate interrater and test-retest reliability, which should be examined in future research. Second, because of the heterogeneity in study design and data collection among studies, some measurements were not available in all studies and at all assessment time points, limiting our sample size for some analyses particularly for analyses of validity and responsiveness, which may have contributed to nonsignificant findings. Third, the Brazil-Neto study evaluated FSS-ICU in Portuguese without undertaking independent forward and backward translation process; however, its results were similar to analyses from the other datasets. Further cross-cultural validation is needed. Fourth, we could not calculate the MIDs using an anchor-based method as recommended (30, 31) because of the lack of MIDs for MMT and other available physical measures that would be needed as anchors. However, the SEM has been recommended among distribution-based MID methods (31) and estimates based on the SEM converged with those from 0.2SD (31, 46). Future studies should compare anchor-based MIDs with distribution-based MIDs.

CONCLUSIONS

The FSS-ICU is an internally consistent, valid, and responsive measure of physical function in the ICU and acute hospital ward setting. The estimated range for the MID of 2.0–5.0 will facilitate sample size calculations and interpretation of future group comparison studies in ICU patients.

ACKNOWLEDGMENT

We thank all of the patients who participated in this study.

REFERENCES

1. Desai SV, Law TJ, Needham DM: Long-term complications of critical care. *Crit Care Med* 2011; 39:371–379
2. Herridge MS, Tansey CM, Matté A, et al; Canadian Critical Care Trials Group: Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med* 2011; 364:1293–1304
3. Fan E, Dowdy DW, Colantuoni E, et al: Physical complications in acute lung injury survivors: A two-year longitudinal prospective study. *Crit Care Med* 2014; 42:849–859
4. Hough CL: Improving physical function during and after critical care. *Curr Opin Crit Care* 2013; 19:488–495
5. Fletcher SN, Kennedy DD, Ghosh IR, et al: Persistent neuromuscular and neurophysiologic abnormalities in long-term survivors of prolonged critical illness. *Crit Care Med* 2003; 31:1012–1016
6. Tipping CJ, Young PJ, Romero L, et al: A systematic review of measurements of physical function in critically ill adults. *Crit Care Resusc* 2012; 14:302–311
7. 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; 41:744–762
8. Connolly B: Describing and measuring recovery and rehabilitation after critical illness. *Curr Opin Crit Care* 2015; 21:445–452
9. Zanni JM, Korupolu R, Fan E, et al: Rehabilitation therapy and outcomes in acute respiratory failure: An observational pilot project. *J Crit Care* 2010; 25:254–262
10. Thrush A, Rozek M, Dekerlegand JL: The clinical utility of the functional status score for the intensive care unit (FSS-ICU) at a long-term acute care hospital: A prospective cohort study. *Phys Ther* 2012; 92:1536–1545
11. Terwee CB, Mokkink LB, Knol DL, et al: Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res* 2012; 21:651–657
12. Needham DM, Korupolu R, Zanni JM, et al: Early physical medicine and rehabilitation for patients with acute respiratory failure: A quality improvement project. *Arch Phys Med Rehabil* 2010; 91: 536–542
13. Parry SM, Denehy L, Beach LJ, et al: Functional outcomes in ICU – what should we be using? – An observational study. *Crit Care* 2015; 19:127
14. Parry SM, Berney S, Warrillow S, et al: Functional electrical stimulation with cycling in the critically ill: A pilot case-matched control study. *J Crit Care* 2014; 29:695.e1–695.e7
15. Kho ME, Truong AD, Zanni JM, et al: Neuromuscular electrical stimulation in mechanically ventilated patients: A randomized, sham-controlled pilot trial with blinded outcome assessment. *J Crit Care* 2015; 30:32–39
16. Kho ME, Truong AD, Brower RG, et al: Neuromuscular electrical stimulation for intensive care unit-acquired weakness: Protocol and methodological implications for a randomized, sham-controlled, phase II trial. *Phys Ther* 2012; 92:1564–1579

17. Lawton MP, Brody EM: Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist* 1969; 9:179–186
18. Katz S, Ford AB, Moskowitz RW, et al: Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *JAMA* 1963; 185:914–919
19. Fan E, Ciesla ND, Truong AD, et al: Inter-rater reliability of manual muscle strength testing in ICU survivors and simulated patients. *Intensive Care Med* 2010; 36:1038–1043
20. De Jonghe B, Sharshar T, Lefaucheur JP, et al; Groupe de Réflexion et d'Etude des Neuromyopathies en Réanimation: Paresis acquired in the intensive care unit: A prospective multicenter study. *JAMA* 2002; 288:2859–2867
21. Massy-Westropp NM, Gill TK, Taylor AW, et al: Hand grip strength: Age and gender stratified normative data in a population-based study. *BMC Res Notes* 2011; 4:127
22. Bohannon RW: Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Arch Phys Med Rehabil* 1997; 78:26–32
23. Hodgson C, Needham D, Haines K, et al: Feasibility and inter-rater reliability of the ICU Mobility Scale. *Heart Lung* 2014; 43:19–24
24. Denehy L, de Morton NA, Skinner EH, et al: A physical function test for use in the intensive care unit: Validity, responsiveness, and predictive utility of the physical function ICU test (scored). *Phys Ther* 2013; 93:1636–1645
25. Piva S, Dora G, Minelli C, et al: The surgical optimal mobility score predicts mortality and length of stay in an Italian population of medical, surgical, and neurologic intensive care unit patients. *J Crit Care* 2015; 30:1251–1257
26. Schaller SJ, Stäuble CG, Suemasa M, et al: The German validation study of the surgical intensive care unit optimal mobility score. *J Crit Care* 2016; 32:201–206
27. Knapp TR: Coefficient alpha: Conceptualizations and anomalies. *Res Nurs Health* 1991; 14:457–460
28. Kazis LE, Anderson JJ, Meenan RF: Effect sizes for interpreting changes in health status. *Med Care* 1989; 27:S178–S189
29. Hough CL, Lieu BK, Caldwell ES: Manual muscle strength testing of critically ill patients: Feasibility and interobserver agreement. *Crit Care* 2011; 15:R43
30. Yost KJ, Eton DT: Combining distribution- and anchor-based approaches to determine minimally important differences: The FACIT experience. *Eval Health Prof* 2005; 28:172–191
31. Revicki D, Hays RD, Cella D, et al: Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61:102–109
32. Cohen J: Statistical Power Analysis for the Behavioral Sciences. Second Edition. Hillsdale, NJ, Lawrence Earlbaum Associates, 1988
33. Bland JM, Altman DG: Cronbach's alpha. *BMJ* 1997; 314:572
34. Chan KS, Friedman L, Dinglas VD et al: Evaluating physical outcomes in ARDS survivors: Validity, responsiveness & minimal important different of 4-meter gait speed test. *Crit Care Med* 2016; 44:859–868
35. 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
36. Pfoh E, Chan K, Dinglas V, et al: Cognitive screening among acute respiratory failure survivors: A cross-sectional evaluation of the mini mental state examination. *Critical Care* 2015;5:220
37. Nordon-Craft A, Schenkman M, Edbrooke L, et al: The physical function intensive care test: Implementation in survivors of critical illness. *Phys Ther* 2014; 94:1499–1507
38. Corner EJ, Wood H, Englebretsen C, et al: The Chelsea critical care physical assessment tool (CPAx): Validation of an innovative new tool to measure physical morbidity in the general adult critical care population; an observational proof-of-concept pilot study. *Physiotherapy* 2013; 99:33–41
39. Corner EJ, Hichens LV, Attrill KM, et al: The responsiveness of the Chelsea critical care physical assessment tool in measuring functional recovery in the burns critical care population: An observational study. *Burns* 2015; 41:241–247
40. Corner EJ, Soni N, Handy JM, et al: Construct validity of the Chelsea critical care physical assessment tool: An observational study of recovery from critical illness. *Crit Care* 2014; 18:R55
41. Perme C, Nawa RK, Winkelman C, et al: A tool to assess mobility status in critically ill patients: The Perme intensive care unit mobility score. *Methodist Debakey Cardiovasc J* 2014; 10:41–49
42. Nawa RK, Lettvin C, Winkelman C, et al: Initial interrater reliability for a novel measure of patient mobility in a cardiovascular intensive care unit. *J Crit Care* 2014; 29:475.e1–475.e5
43. Bissett B, Green M, Marzano V, et al: Reliability and utility of the acute care index of function in intensive care patients: An observational study. *Heart Lung* 2016; 45:10–14
44. Kasotakis G, Schmidt U, Perry D, et al: The surgical intensive care unit optimal mobility score predicts mortality and length of stay. *Crit Care Med* 2012; 40:1122–1128
45. de Vet H, Terwee CB, Mokkink LB, et al: Measurement in Medicine - A Practical Guide. Cambridge, Cambridge University Press, 2011
46. Wyrwich KW, Tierney WM, Babu AN, et al: A comparison of clinically important differences in health-related quality of life for patients with chronic lung disease, asthma, or heart disease. *Health Serv Res* 2005; 40:577–591