J Rehabil Med 2006; 38: 237–242

Taylor & Francis
Taylor & Francis Group

MEASURING PATIENT-REPORTED OUTCOMES AFTER DISCHARGE FROM INPATIENT REHABILITATION SETTINGS[1]

Wendy J. Coster, PhD[1], Stephen M. Haley, PhD[2] and Alan M. Jette, PhD[2]

From the [1]Department of Occupational Therapy & Rehabilitation Counseling, Sargent College of Health and Rehabilitation Sciences and [2]Health and Disability Research Institute, Boston University, Boston, USA

Objective: **To examine the sensitivity of the Short Form Activity Measure for Post-Acute Care (AM-PAC) in comparison to the Functional Independence Measure (FIM™) across a 12-month period after discharge from rehabilitation hospital.**

Design: **Prospective longitudinal study. Patients were recruited while receiving inpatient services from facilities in the north-east USA and interviewed 1, 6 and 12 months thereafter.**

Patients: **Convenience sample of 516 patients at baseline (65% retention at the final follow-up) receiving rehabilitation services for neurological, lower extremity orthopedic, or complex medical conditions. Mean age 68.3 years; 47% male.**

Main outcome measures: **AM-PAC Physical and Movement, Personal Care and Instrumental, and Applied Cognitive Activity scales; FIM™ Motor and Cognitive scales.**

Results: **All 3 AM-PAC scales were sensitive to both positive and negative change across the follow-up period. Standardized response means for the AM-PAC were consistently larger than for the FIM™ across patient and severity groups. A greater percentage of patients showed positive change that exceeded the minimal detectable change on the AM-PAC than on the FIM™ at both 6- and 12-month follow-ups.**

Conclusion: **The short-form AM-PAC scales are more sensitive measures of change in functional activity performance over time in the general population of persons who receive inpatient rehabilitation services compared to the FIM™. Thus, the AM-PAC offers a short, comprehensive, and sensitive measure of positive and/or negative change in patients' ability to perform important activities of daily life.**

Key words: outcomes, rehabilitation, measurement.

J Rehabil Med 2006; 38: 237–242

Correspondence address: Wendy Coster, Department of Occupational Therapy and Rehabilitation Counseling, Sargent College of Health and Rehabilitation Sciences, Boston University, 635 Commonwealth Avenue, Boston, MA 02215, USA. E-mail: wjcoster@bu.edu

Submitted September 22, 2005; accepted January 24, 2006

The ability to track rehabilitation outcomes accurately over time is essential in order to disentangle and better understand the factors that lead to more and less successful return to performance of daily activities. This important aspect of rehabilitation research has been seriously hampered, however, by limitations in the measures currently in widespread use. Problems identified in the literature include discrepancies in estimates of function when different instruments are used to assess daily activity performance (1–3) and the significant ceiling effects found for measures that focus only on basic activities of daily living when they are used to assess patients for follow-up after discharge from inpatient settings (4, 5). For example, although the Functional Independence Measure (FIM™) (6) is widely used in inpatient rehabilitation programs and research, it does not examine performance of the broader range of daily activities required for community function. Thus, it may have limited capacity to measure the extent to which patients have successfully resumed home and community responsibilities. Conversely, generic measures of health status such as the SF-36 (7) may show floor effects in persons with severe functional limitations, have many items that are not relevant in early stages of recovery, or that are difficult to answer by people using alternate means of mobility (8–10). More comprehensive alternatives to these measures exist (11, 12), but are limited in practical application because of their greater length.

The impact of these limitations is far from trivial. Some studies have suggested that the most commonly used measures may underestimate residual functional limitations in some populations, and thus do not provide an accurate picture of needs for continuing services beyond discharge from the inpatient treatment setting (5, 13). A related concern is that projections regarding likely extent of functional recovery in some populations may be inaccurate if the measures used to track outcomes are insensitive to lesser, but still clinically significant, degrees of progress (3, 14, 15). In addition, when measures that focus only on the performance of relatively basic activities are used to evaluate rehabilitation interventions, we run the risk of overlooking the impact of these interventions on other areas of daily living that may be related to the patient's long-term health and well-being (13). Thus, there is a serious need for improved measures to be used for follow-up after discharge from inpatient care settings.

Our research group has developed a new set of short but sensitive functional measures that can be used to monitor patient-reported rehabilitation outcomes across the full spectrum of service settings. We have described development of the Activity Measure for Post-Acute Care (AM-PAC)

[1] An earlier version of these findings was presented as part of a webcast by the Rehabilitation Research and Training Center for Measuring Rehabilitation Outcomes, Boston University, USA, February 26, 2004.

 ISSN 1650-1977
DOI: 10.1080/16501970600609774

in detail in previous reports (2, 16, 17). The next phase of our research, which is the focus of the present paper, examined the ability of the AM-PAC short forms to detect change in a population of rehabilitation patients across the 12 months following discharge from an inpatient setting and compared its performance to that of the FIM™ (6). Specifically, we examined whether the AM-PAC, as compared with the FIM™, achieved its intended purpose of reducing ceiling effects and being more sensitive to change in function over the 12 months after discharge. We chose this extended length of time because the literature has suggested that recovery in some patient groups may continue over this period (18, 19) and an important goal of this project was to provide measures that were adequate to detect such change if indeed it was occurring. We examined AM-PAC sensitivity across patients who were initially classified at different levels of severity as well as across patients with different types of conditions in order to determine whether the scales performed equally well for these different groups.

The specific research questions we addressed were:

- Do the AM-PAC scales show less ceiling effect than the FIM™?
- How sensitive to change are the AM-PAC scales across the 12 months after inpatient rehabilitation?
- Are the AM-PAC scales equally sensitive to change across groups with different types of conditions and different levels of severity?
- What is the relative sensitivity of the AM-PAC scales compared to that of the FIM™ across the 12-month follow-up?

METHODS

Subjects

Participants were adults age 18 years and older admitted to a large tertiary care centre or 1 of 2 rehabilitation hospitals in the Boston area (USA) who had a primary diagnosis of neurological disorder (e.g. stroke, Parkinson's disease, traumatic brain injury), lower extremity orthopedic trauma (e.g. hip fracture, amputation) or medically complex conditions (e.g. chronic obstructive pulmonary disease, heart disease, diabetes, liver disease). Specific inclusion criteria were that the person was currently receiving and/or about to be referred to skilled rehabilitation services (physical therapy, occupational therapy, or speech and language pathology); was able to speak and understand English; and had a prognosis for survival of one year, as determined by the primary physician or a facility recruiter via medical record review. In addition, patients were excluded if the facility recruiter judged that they were unable to give informed consent based on information in the medical record and/or discussions with treating clinicians. Specifically, the presence of any of the following criteria indicated ineligibility: (*i*) any orientation deficit, (*ii*) difficulty remembering the day's events, (*iii*) receptive or expressive communication deficits that precluded the patient from communicating responses reliably (verbally or non-verbally).

Study procedures were approved by the University Institutional Review Board as well as the research review committees of the participating institutions.

Procedure

AM-PAC and FIM™ data were collected via patient interview at 3 time points: 1, 6 and 12 months after discharge from acute care or rehabilitation hospital. An on-site recruiter at the facility explained the study to potential participants, answered any questions and obtained signed consent forms. The data collector abstracted information from the medical record including basic demographic and medical diagnostic information. All data were entered into files on laptop computers without personal identifiers.

Patient interviews were conducted by trained interviewers at the subjects' current living location or at a mutually convenient location. Research staff contacted the subject 1–2 weeks before each interview was scheduled to occur to set up a convenient time and location for the interview. A window of 6 weeks from the due date to be interviewed was applied. Subjects not interviewed within this time interval were dropped from that time point. Each interview lasted about 45 minutes to one hour.

Instruments

AM-PAC. The initial content domains and item definition for the AM-PAC item pool were guided by the World Health Organization's International Classification of Functioning, Disability and Health (ICF) (20) definition and categories of *activity*. Subsequent factor analyses and Rasch analyses of data from a sample of over 400 persons receiving rehabilitation services led to the definition of 3 separate activity scales: Physical and Movement, Personal Care and Instrumental, and Applied Cognitive. Coverage range, unidimensionality, reliability, and validity of these scales were confirmed in subsequent analyses (1, 2).

Each of the 3 AM-PAC scales consists of 10 items that ask about either the difficulty (5-point rating) or use of assistance (6-point rating) to perform specified daily activities. Subjects were given a response card with the relevant response options in large print to use during this part of the interview. In this study, subjects were administered the community form of the AM-PAC, which includes both basic activities of daily living (e.g. completing grooming activities) and activities more typically performed at home, such as walking several blocks, putting dishes away, or looking up a telephone number.

The community form is linked to an inpatient version of the AM-PAC, which contains only activities likely to be performed in that setting. The linked format supports continuous tracking of a patient across the full spectrum of settings using a single scale. The items for each version were selected from a common item pool based on their ability to provide useful measurement information across a broad spectrum of function. Scales were derived from Rasch analyses conducted on the item pool therefore the scores are interval-level data (see details published elsewhere (17)). They range from 0 to 100 with higher scores reflecting greater function (less difficulty, less use of assistance). Test-retest reliability estimates for the longer AM-PAC versions from which these short forms were derived ranged from 0.91 to 0.97 (21); separate reliability analyses of the short scales have not been conducted.

FIM™. The patient interview version of the FIM used for the follow-up interviews has been tested in similar populations using both phone and in-person interview methods. These studies found acceptable reliability of the resulting scores (22). Raw total scores for the Motor and Cognitive subscales were transformed to logit-based scores on a 0–100 scale using the tables published by Heinemann et al. (23). Because raw score totals have frequently been used in clinical research, we also repeated some analyses using the raw score totals to compare sensitivity of these 2 scoring methods.

AM-PAC scales were administered in a pre-assigned randomized order to minimize loss of data on particular measures due to patient fatigue.

Data analysis

We used the standardized response mean (SRM) to assess sensitivity to change because it provides an estimate of the magnitude of change that is not influenced by sample size. The SRM is calculated as the mean change in scores from time 1 to time 2, divided by the standard deviation of these changes (24, 25). It is interpreted like an effect size, thus larger SRMs indicate greater difference in mean amount of change. Sensitivity of the AM-PAC and FIM™ were examined across 2 intervals: 1–6 months, and 6–12 months. AM-PAC subscale estimates were compared to those for the 2 FIM subscales (Motor and Cognitive). We also calculated the percentage of subjects at floor and ceiling at each

time point for each instrument to examine the extent to which change calculations might be affected by restrictions of range.

In addition, we calculated the minimal detectable change (MDC) (26). This statistic, also known as the reliable change index (27), examines the extent to which patient change exceeds the amount of variability accounted for by measurement error (SEM). Accordingly, it is calculated using a reliability coefficient, which in this case was the test-retest reliability coefficient r. We conducted 2 related analyses with the MDC: (*i*) an overall statistic (similar to a z-score), which identifies whether the mean change for the group as a whole was significantly larger ($p < 0.05$) than the SEM; (*ii*) the percent of subjects whose individual score changed by more than the number of scale points that might be accounted for by measurement error.

Reports from the interviewers that many participants had experienced setbacks in their recovery prompted us to investigate the proportion of the sample showing negative change on each scale at each time point. These numbers were substantial and increased across the 12 month follow-up period, meaning that the mean change score and SRM for the total group would be seriously affected by negative and positive scores canceling each other out. Therefore, in order to examine sensitivity to change in both directions, calculations for each scale were performed separately for the group who showed no or positive change (change score ≥ 0) and those showing negative change (change score <0).

All analyses were conducted using the SAS® program version 9.0 (28).

RESULTS

The initial sample consisted of 516 patients. There were slightly more women than men (53% vs 47%) and a greater percentage classified in the complex medical category (44%) compared to the lower extremity orthopedic (32%) and neurological (24%) categories. The mean age of participants was 68.3 years, however the range extended from 19 to 100 years with about 20% of subjects younger than age 50 years. Using Modified Rankin Scale (29) categories about 23% were classified at baseline as having severe disability, 52% as moderate, and 25% as slight.

At follow-up, 417 (81%) of participants were seen at one month, 370 were interviewed at 6 months (72%), and 336 (65%) were seen again at 12 months. Of those lost to follow-up at the final time point, 50 (9.7%) had died and 130 (25%) dropped out, either because they could not be located ($n = 63$) or they chose not to continue ($n = 67$). Those lost to follow-up did not differ from participants with respect to demographic background, impairment group, or baseline AM-PAC scores. At each time point more than 90% of participants were living at home. Additional demographic information about the sample is provided in Table I.

For analyses involving change across time periods, the number of subjects varied from the totals given above because some subjects were missing data for one of the follow-up assessments.

Table I. *Sample characteristics (n = 516)*

Variable	
Age (years)	
Mean (SD)	68.3 (14.97)
Range	19–100
Gender (*n*,%)	
Male	243 (47.1)
Female	273 (52.9)
Race/ethnicity (*n*,%)	
White	462 (89.5)
Black	27 (5.2)
Other	27 (5.3)
Education (*n*,%)	
High school or less	228 (44.2)
Beyond high school	262 (50.8)
Missing	26 (5)
Living location prior to hospitalization (*n*,%)	
Home	477 (92.4)
Assisted living	11 (2.1)
Health care facility	9 (1.8)
Missing	19 (3.7)
Living situation (*n*,%)	
Alone	169 (32.7)
With spouse	147 (28.4)
With family	134 (26.0)
With non-family	16 (3.1)
Pre-existing medical conditions (*n*,%)	
Cardiopulmonary	339 (65.7)
Diabetes	151 (29.3)
Chronic pain	57 (11)

Ceiling effects

The Physical and Movement Activity scale showed the least amount of ceiling effect (less than 1% even at 12-month follow-up). Personal Care and Instrumental had slightly higher rates (16% by 12 months) while the Applied Cognitive scale showed the largest effects (about 27% at one month and 44% at 12-month follow-up. The percentage at ceiling on the FIM Motor ranged from about 10% at one month to 15% at 12 months. A very large percentage of patients (70%) were at ceiling at one month on the FIM Cognitive scale. This figure declined to 53% by 12 months.

No patients scored at floor on the AM-PAC or FIM at any time points.

Sensitivity across diagnostic groups

All AM-PAC scales were sensitive to both positive (Fig. 1A) and negative (Fig. 1B) change across the follow-up periods. Variations in SRM among the 3 patient diagnostic groups (neurological, orthopedic, complex medical) were typically small (range –0.02–0.10) with only a few differences exceeding this amount. Mean positive change ranged between 8.35 and 13.23 points from 1 to 6 months, and between 4.64 and 7.56 points from 6 to 12 months. SRMs were around 1.0 (i.e. approximately 1 SD difference) at the 6-month follow-up and somewhat lower at 12 months. Despite the decline in magnitude of positive change, these results indicate that considerable progress in functional performance was still occurring more than 6 months post-hospitalization in all patient groups.

FIM™ Motor positive SRM was similar to that of the corresponding AM-PAC scales (Physical & Movement and Personal Care & Instrumental) at 6 months, but smaller at 12-month follow-up. In contrast, the FIM™ Cognitive SRMs were

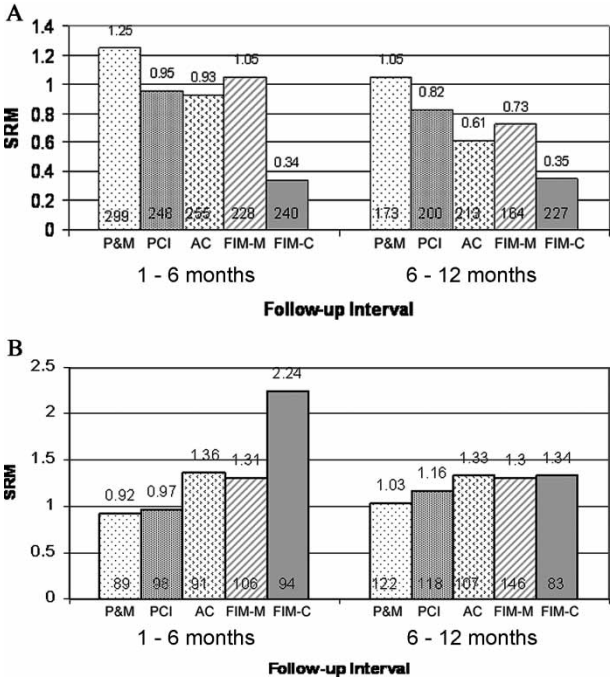

Fig. 1. Standardized response means (SRM) for Activity Measure for post-acute care (AM-PAC) and FIM™. (A) positive and (B) negative changes. P&M: Physical and Movement; PCI: Personal Care and Instrument, AC: Applied Cognitive.

considerably smaller than AM-PAC SRMs at both the 6- and 12-month follow-ups (see Fig. 1A).

An increasing proportion of the sample showed negative change (i.e. reduction in reported functional activity performance) across the follow-up period. The proportions of subjects identified by AM-PAC and FIM™ negative scores were similar. At the 12-month follow-up approximately 40% of the participants had physical function scores (AM-PAC Physical & Movement or Personal Care & Instrumental scales, or FIM Motor scale) that had declined across the 6 months since the previous interview. The magnitude of this change was substantial and in many cases exceeded that seen in the positive change group (see Fig. 1B). Although some differences in magnitude of negative SRM across clinical groups were observed, the meaningfulness of these differences is unclear because some subgroups had relatively few (< 10) subjects.

Sensitivity across severity levels

The AM-PAC detected significant change at both follow-up points for all severity groups in both positive and negative directions. There were some differences in standardized response mean across groups, but these were generally in the range of 0.10 to 0.30, indicating that sensitivity did not differ substantially by initial severity of disability.

Minimal Detectable Change

Mean positive change on the AM-PAC Physical & Movement scale significantly ($p < 0.05$) exceeded the MDC at both 6- and 12-month follow-ups, while change on the Personal Care and Instrument scale was significant only at the 6-month follow-up. Mean change on the Applied Cognitive scale did not exceed the MDC at either point. Positive FIM™ Motor change exceeded the MDC at 6 months, but not at 12 months, while FIM Cognitive did not exceed the MDC at either point. Negative FIM change exceeded the MDC for both scales at both time points. On the AM-PAC only Personal Care and Instrument negative change was significant at 6 months, however all 3 scales exceeded this value for negative change at 12 months.

The AM-PAC Physical & Movement and Personal Care & Instrumental scales detected a slightly higher or similar percentage of participants with positive change that exceeded the MDC compared to the FIM Motor scale at both follow-up points. In contrast, the percentage identified with positive change in the cognitive area by the AM-PAC scale was twice that identified by the FIM Cognitive scale. At both follow-ups, the FIM™ identified more participants with negative scores that exceeded the MDC. These patterns were consistent across patient diagnostic groups and severity groups (Fig. 2).

Replication of these analyses using FIM raw score totals indicated that, in general, the SRM estimates based on raw scores were smaller than those computed from logit-based scores, especially for negative change estimates. For example, the negative change SRM for FIM Cognitive at 6 months was 1.39 using the raw scores vs 2.24 using the logit-based scores.

DISCUSSION

These results support the utility of the short-form AM-PAC scales to capture both positive and negative change in functional activity performance over 12 months in persons receiving rehabilitation services for a variety of clinical conditions. Overall, the AM-PAC scales appeared more sensitive to change than the comparable FIM™ scales (using logit-based scores), especially for positive change. The magnitude of the AM-PAC SRM at the 12-month follow-up assessment indicates that more functional recovery had occurred than was detected by the FIM™. These estimates are also larger than those reported by other studies using measures, such as the Barthel Index (30). These results are consistent with other studies that have raised concerns that reliance on basic activities on daily living measures in research may provide an insufficient picture of longer-term outcomes for persons who receive inpatient rehabilitation services (4, 13).

Our comparisons indicate that the Physical and Movement and Personal Care and Instrumental Activity scales of the AM-PAC have little ceiling effect. In contrast, a significant percentage of patients were at ceiling on the Applied Cognitive scale at each time point. This result was not unexpected given that the primary reasons for receiving inpatient rehabilitation services are typically problems with physical function rather than cognitive function. In addition, as we have discussed elsewhere (2), it has proven more difficult to construct items that measure an upper range of typical cognitive function (i.e. more

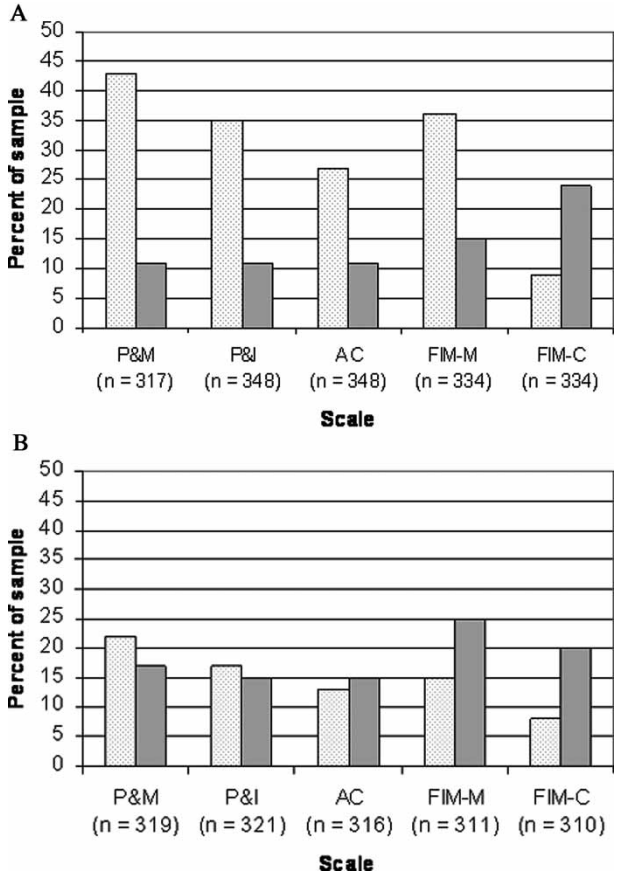

Fig. 2. Percent of sample with change exceeding minimal detectable change (A) at 1–6 months and (B) at 6–12 months. □ +change; ■ –change. For abbreviations see Fig. 1.

"difficult" items) that are not primarily reflective of level of education. The AM-PAC Applied Cognitive scale did have less of a ceiling effect than the FIM™ at both follow-ups.

The very large negative change SRM for the FIM-Cognitive scale at the 6-month follow-up is likely linked to the scale's ceiling effect. As noted in the literature (30), logit-based scores at the extreme on scales with few items may change dramatically with a small change (e.g. 1 point) in raw score. Because 70% of the participants were at ceiling on this scale at 1 month, small declines in some of these participants across the follow-up period had significant impact on the summary statistics such as the SRM.

It is important to note that our results apply only to logit-based FIM™ summary scores. Our analyses indicate that use of FIM™ raw summary scores for outcomes measurement (as is still common) will likely under-represent the extent of patient change, both positive and negative.

The results of this study also raise a number of important issues that extend beyond establishing the properties of these particular measures. First, self-report measures like the AM-PAC may often be the only feasible means to obtain functional activity information from many service recipients once they have left the hospital. However, there is a need to further investigate the extent to which these patient self-report data provide a sufficient and valid basis for examining outcomes over time. A number of studies using other measures (31, 32) have documented differences between patient, clinician, and caregiver reports on the person's functioning. In the present study, patients identified by clinical staff as having cognitive limitations sufficient to raise questions about their ability to provide informed consent or to participate in an interview were excluded. Further work is needed to help identify characteristics of less accurate respondents and to examine the most common areas of agreement and disagreement among different respondents for the AM-PAC.

Another important issue raised by the results of this study is the extent to which patients with negative change scores have been accounted for in other long-term outcome studies, particularly those examining sensitivity using statistics such as the standardized response mean. While some reports indicate clearly that they excluded subjects whose function declined (5) and others anticipated and therefore measured such declines (e.g. in persons with amyotrophic lateral sclerosis (33), the majority of studies do not separate these 2 groups in their analyses. Given the substantial proportion of subjects in our sample who showed decline in function, unless positive and negative change groups are analyzed separately or the absolute value of the change scores is used in calculations, estimates of mean change for the group may be significantly attenuated by negative and positive scores canceling each other out. This is clearly an issue that needs further investigation as it has relevance to all longitudinal follow-up analyses with this population.

Finally, work is needed to identify appropriate external criterion measures that can be used to examine measure responsiveness, or the ability to detect meaningful change. There is currently no gold standard that can be used to examine the extent to which new clinical measures like the AM-PAC are able to capture changes in function that have relevance in daily life. Although some studies use patient-rated "improvement" ratings as the criterion measure for meaningful change, there are clear limitations of this method when it is applied over an extended period of time (34, 35). In this study, asking the patient to rate improvement since the first measurement point would require the patient to recall and compare his or her functioning a year earlier to the present. In the absence of a valid external criterion it is difficult to take investigation of measurement properties beyond establishing whether they are sensitive to statistically significant change, as we have done in the present study.

In conclusion, our evidence suggests that the AM-PAC short forms are sensitive to change in activity performance after discharge from inpatient rehabilitation for persons with a variety of medical conditions and levels of severity. Because the AM-PAC consists of linked Facility and Community forms, patients can be assessed using only items that are relevant to his or her current setting. However, because summary scores for both forms are on the same scale, patient change can be tracked over time across a broader range of function. Thus, the AM-PAC meets the critical need for a consistent measure that can be

used across rehabilitation settings to provide estimates of change that are not contaminated by unknown effects of using different instruments at different times. Use of more sensitive measures is critical to obtain an accurate picture of the course of recovery after an episode of inpatient rehabilitation care, whether for monitoring the outcomes of specific programs or for testing the effectiveness of interventions.

ACKNOWLEDGEMENTS

Supported in part by Grant H133B990005 from the National Institute on Disability and Rehabilitation Research (NIDRR), Rehabilitation Research and Training Center on Measuring Rehabilitation Outcomes and Grant R01 HD43568 from the National Institute of Child Health and Human Development (NICHD) and the Agency for Healthcare Research and Quality (AHRQ). The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NICHD, AHRQ or NIH.

REFERENCES

1. Coster WJ, Haley SM, Andres PL, Ludlow LH, Bond T, Ni PS. Refining the conceptual basis for rehabilitation outcome measurement: personal care and instrumental activities domain. Med Care 2004; 42 (suppl 1): 62–72.
2. Coster WJ, Haley SM, Ludlow LH, Andres PL, Ni PS. Development of an applied cognition scale to measure rehabilitation outcomes. Arch Phys Med Rehabil 2004; 85: 2030–2035.
3. Duncan P, Jorgensen H, Wade D. Outcome measures in acute stroke trials. A systematic review and some recommendations to improve practice. Stroke 2000; 31: 1429–1438.
4. Duncan PW, Samsa GP, Weinberger M, Goldstein LB, Bonito A, Witter DM, et al. Health status of individuals with mild stroke. Stroke 1997; 28: 740–745.
5. Lai SM, Studenski S, Duncan PW, Perera S. Persisting consequences of stroke measured by the Stroke Impact Scale. Stroke 2002; 33: 1840–1844.
6. Guide for the Uniform Data Set for Medical Rehabilitation (Adult FIM™) version 5.1. Buffalo: State University of NY; 1997.
7. Ware J, Kosinski M. The SF-36® Physical and Mental Health Summary Scales: A Manual for User's of Version 1.0, 2nd edn. Boston, MA: Health Assessment Lab; 2001.
8. Lai SM, Perera S, Duncan PW, Bode R. Physical and social functioning after stoke: comparison of the Stroke Impact Scale and the Short Form-36. Stroke 2003; 34: 488–493.
9. Kersten P, Mullee MA, Smith JA, McLellan L, George S. Generic health status measures are unsuitable for measuring health status in severely disabled people. Clin Rehabil 1999; 13: 219–228.
10. Andresen EM, Fouts BS, Romeis JC, Brownson CA. Performance of health-related quality-of-life instruments in a spinal cord injured population. Arch Phys Med Rehabil 1999; 80: 877–884.
11. Harvey RF, Jellinek HM. Functional performance assessment: a program approach. Arch Phys Med Rehabil 1981; 62: 456–460.
12. Harvey RF, Silverstein B, Venzon MA, Kilgore KM, Fisher WP, Steiner M, et al. Applying psychometric criteria to functional assessment in medical rehabilitation. III. Construct validity and predicting level of care. Arch Phys Med Rehabil; 1992; 73: 887–892.
13. Desrosiers J, Rochette A, Noreau L, Bravo G, Hebert R, Boutin C. Comparison of two functional independence scales with a participation measure in post-stroke rehabilitation. Arch Gerontol Geria 2003; 37: 157–172.
14. Jette AM, Haley SM, Ni P-S. Comparison of functional status tools used in post-acute care. Health Care Fin Rev 2003; 24: 13–24.
15. Patel AT, Duncan PW, Lai SM, Studenski S. The relation between impairments and functional outcomes poststroke. Arch Phys Med Rehabil 2000; 81: 1357–1363.
16. Haley SM, Coster WJ, Andres PL, Ludlow LH, Ni PS, Bond TL, et al. Activity outcome measurement for post-acute care. Med Care 2004; 42 (suppl 1): 49–61.
17. Haley SM, Andres PL, Coster WJ, Kosinski M, Ni PS, Jette AM. Short-form activity measure for post-acute care (AM-PAC). Arch Phys Med Rehabil 2004; 85: 649–660.
18. van Balen R, Essink-Bot ML, Steyerberg EW, Cools HJM, Habbema JDF. Quality of life after hip fracture: a comparison of four health status measures in 208 patients. Disabil Rehabil 2003; 25: 507–519.
19. Magaziner J, Hawkes W, Hebel JR, Zimmerman SI, Fox KM, Dolan M, et al. Recovery from hip fracture in eight areas of function. J Gerontol A Biol Sci Med Sci 2000; 55: M498–507.
20. World Health Organization. International Classification of Functioning, Disability and Handicap (ICF). Geneva: World Health Organization; 2001.
21. Andres PL, Haley SM, Ni PS. Is patient-reported function reliable for monitoring post-acute outcomes? Am J Phys Med Rehabil 2003; 82: 614–621.
22. Petrella R, Overend T, Chesworth B. FIM after hip fracture: is telephone administration valid and sensitive to change? Am J Phys Med Rehabil 2002; 81: 639–644.
23. Heinemann AW, Linacre JM, Wright BD, Hamilton BB, Granger C. Measurement characteristics of the Functional Independence Measure. Top Stroke Rehabil 1994; 1: 1–15.
24. Stratford PW, Binkley JW, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. Phys Ther 1996; 76: 1109–1123.
25. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000; 53: 459–468.
26. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V. Measuring the whole or the parts? Validity, reliability, and responsiveness of the disabilities of the arm, shoulder and hand outcome measure in different regions of the upper extremity. J Hand Ther 2001; 14: 128–146.
27. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consul Clin Psychol 1991; 59: 12–19.
28. SAS program version 9.0. SAS Institute, Inc.; 2004.
29. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. Scot Med J 1957; 2: 200–215.
30. McHorney CA, Haley SM, Ware JE. Evaluation of the MOS SF-36 physical functioning scale (PF-10): II. Comparison of relative precision using likert and Rasch scoring methods. J Clin Epidemiol 1997; 50: 451–461.
31. Malzer RI. Patient performance level during inpatient physical rehabilitation: therapist, nurse and patient perspective. Arch Phys Med Rehabil 1988; 69: 113–122.
32. Rogers JC, Holm MB, Beach S, Schulz R, Cipriani J, Fox A, et al. Concordance of four methods of disability assessment using performance in the home as the criterion method. Arthrit Rheum 2003; 49: 640–647.
33. Norquist JM, Fitzpatrick R, Jenkinson C. Rasch measurement in the assessment of amytrophic lateral sclerosis patients. J App Meas 2003; 4: 249–257.
34. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. Control Clin Trials 1989; 10: 407–415.
35. Liang MH. Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. Med Care 2000; 38 (suppl II): II-84–90.