SHORT REPORT

Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke

I-P Hsueh, J-H Lin, J-S Jeng, C-L Hsieh

....

Objectives: To compare the reliability, validity, and responsiveness of the motor subscale of the functional independence measure (FIM), the original 10 item Barthel index (BI), and the 5 item short form BI (BI-5) in inpatients with stroke receiving rehabilitation.

Methods: 118 inpatients with stroke at a rehabilitation unit participated in the study. The patients were tested with the FIM motor subscale and original BI at admission to the rehabilitation ward and before discharge from the hospital. The distribution, internal consistency, concurrent validity, and responsiveness of each measure were examined.

Results: The BI and FIM motor subscale showed acceptable distribution, high internal consistency (α coefficient ≥ 0.84), high concurrent validity (Spearman's correlation coefficient, $r_s \geq 0.92$, intraclass correlation coefficient (ICC) ≥ 0.83), and high responsiveness (standardised response mean ≥ 1.2, $p < 0.001$). The BI-5 exhibited a notable floor effect at admission but this was not found at discharge. The BI-5 showed acceptable internal consistency at admission and discharge (α coefficient ≥ 0.71). The concurrent validity of the BI-5 was poor to fair at admission ($r_s = 0.74$, ICC ≤ 0.55) but was good at discharge ($r_s \geq 0.92$, ICC ≥ 0.74). It is noted that the responsiveness of the BI-5 was as high as that of the BI and the FIM motor subscale.

Conclusions: The results showed that the BI and FIM motor subscale had very acceptable and similar psychometric characteristics. The BI-5 appeared to have limited discriminative ability at admission, particularly for patients with severe disability; otherwise the BI-5 had very adequate psychometric properties. These results may provide information useful in the selection of activities of daily living measures for both clinicians and researchers.

T here are many published activities of daily living (ADL) indices for patients with stroke. Choosing an objective and scientific ADL measure is difficult but important for both clinicians and researchers. The Barthel index (BI)[1] and the functional independence measure (FIM)[2] are the most widely used measures of disability within Europe.[3] The FIM was developed to be a more comprehensive and responsive measure of disability than the BI.[4] However, a recent study found that the BI and the FIM had similar psychometric characteristics (reliability, validity, and responsiveness) in patients undergoing rehabilitation, suggesting that the FIM has no advantage over the BI.[5] Another recent study compared the psychometric properties of a five item short form BI (BI-5) with the original 10 item BI and found that both versions were psychometrically equivalent in patients undergoing neurorehabilitation.[6] However, psychometric characteristics

are sample dependent.[7] Further comparison of the psychometric properties of these ADL measures is required to determine the generalisability of the above results. The purpose of this prospective study was to compare the reliability, validity, and responsiveness of the motor subscale of the FIM, the original 10 item BI, and the BI-5 concurrently in a cohort of inpatients with stroke who were undergoing rehabilitation.

METHODS

Subjects

Subjects were recruited from the stroke registry of a study for the construction of an ADL scale conducted at National Taiwan University Hospital between 1 December 1999 and 31 May 2000. Patients were included in this study if they met two criteria: firstly, diagnosis (*International classification of diseases*, ninth revision, clinical modification codes) of cerebral haemorrhage (431), cerebral infarction (434), or other (432, 433, 436, 437); and, secondly, informed consent for participation, obtained from the patient or a family member responsible for the patient's care. Subjects were not included in further testing or analysis in the study if they suffered another stroke or from other major diseases during the study period.

Procedures

The BI and the motor subscale of the FIM were administered to patients at admission to the rehabilitation ward and again before hospital discharge. Both measures were administered separately on the same day by two occupational therapists in accordance with a counterbalanced sequence. Each patient's ADL performance was rated primarily by interviewing the patients, their primary caregiver, or their nurse. Observation of performance was applied if necessary. The therapists were blinded to both the purpose of the study and the results of each other's assessments during the study period.

Instruments

The FIM[2] is an 18-item scale and is scored from 18 (total assistance in all areas) to 126 (complete independence in all areas). The FIM consists of 13 motor (or physical) and 5 social-cognitive items, assessing self care, sphincter control, transfer, locomotion, communication, social interaction, and cognition. The results from the first 13 items (FIM motor) are summed to develop a motor score with a range of 13 to 91, and these items and figures were used in this study. The reliability and validity of the FIM have been studied extensively.[4][8]

The BI[1] has 10 items of ADL: feeding, grooming, bathing, dressing, bowel and bladder care, toilet use, ambulation, transfers, and stair climbing. The total score ranges from 0 to

....

Table 1 Distribution, internal consistency, validity, and responsiveness of the motor subscale of the functional independence measure (FIM), original 10 item Barthel index (BI), and 5 item BI (BI-5)

	FIM motor subscale		BI		BI-5	
	Admission	Discharge	Admission	Discharge	Admission	Discharge
Distribution						
Floor/ceiling effect (%)	5.8/0	3.5/0	18.2/0	4.7/0	46.6/0	13.6/0
Median (inter-quartile range)	28 (18–35)	43 (32–54)	5 (1.5–8)	10 (6–13)	1 (0–1)	4 (1–5)
Reliability						
Cronbach α	0.88	0.91	0.84	0.85	0.71	0.73
Concurrent validity						
Spearman correlation coefficient	0.74*	0.92*	0.92†	0.94†	0.74‡	0.94‡
ICC	0.55*	0.86*	0.83†	0.87†	0.36‡	0.74‡
Responsiveness						
Standardised response mean	1.3		1.2		1.2	
Change scores relation	0.75*		0.88†		0.78‡	
Wilcoxon Z (p value)	7.5 (<0.001)		7.4 (<0.001)		7 (<0.001)	

*FIM motor subscale versus BI-5; †FIM motor subscale versus BI; ‡BI versus BI-5. ICC, intraclass correlation coefficient.

20. It has been shown to be a reliable, valid, and responsive measure of basic ADL in patients with stroke.[9] As there are several scoring guidelines for the BI, we adopted the version of Collin et al,[10] which was also used in the studies mentioned in the introduction to this paper.

The BI-5 was derived from the original BI to simplify the testing procedure and decrease the time taken to administer the measure.[6] The 5 items are transfers, bathing, toilet use, stair climbing, and mobility. The BI-5 showed promising reliability, validity, and responsiveness in patients (about half of whom had multiple sclerosis and 15% of whom had had a stroke) undergoing rehabilitation.[6]

Analysis

The score ranges and distributions of each of the three measures were examined. The floor and ceiling effects (the percentages of the sample scoring the minimum and maximum possible scores) reflect the extent to which scores cluster at the bottom and top, respectively, of the scale range.

The internal consistency of each ADL measure was expressed using Cronbach's α coefficients. An α coefficient > 0.70 is considered adequate for group comparison.[11] Concurrent validity is usually established by a high correlation between the scale and an ideal. The interrelations between the three ADL measures at admission and discharge were examined using the Spearman's correlation coefficient (r_s) and the intraclass correlation coefficient (ICC). An ICC > 0.75 indicates excellent agreement. Because each of the three measures used has a different score range, the scores from each measure were transformed to a 0–100 range using the following formula[11]:

$$100 \times (\text{observed score} - \text{minimum possible score}) / \text{score range}$$

Responsiveness was examined using the standardised response mean, one type of effect size. The standardised response mean was calculated by dividing the mean change scores by the standard deviation of the change score in the same subjects. According to Cohen's criteria,[12] an effect size > 0.8 is large, 0.5–0.8 is moderate, and 0.2–0.5 is small. Furthermore, the relation between the change scores of the three ADL measures was examined using r_s. In addition, Wilcoxon matched pairs signed rank tests were performed to determine the significance of the change scores.

RESULTS

A total of 125 patients were enrolled in the study. Seven of these patients suffered from another stroke during the hospital stay and were therefore not included in further testing and analysis. Of the remaining 118 subjects studied, 50 were women and 68 were men, with a mean (SD) age of 67.5 (10.9) years. More than half (62.3%) of these patients had cerebral infarction and about 30% had cerebral haemorrhage. The median (interquartile range) number of days of hospital rehabilitation stay of these patients was 26 (range 14–45). The FIM motor subscale and BI scores indicated that the patients were severely disabled (table 1).

Distributions of the BI and the FIM motor subscale were acceptable (table 1). However, the BI-5 showed significant floor effects at admission, but this was not found at discharge. The interquartile score range of the BI-5 at admission was quite limited (table 1). The Cronbach's α of the three measures were ≥ 0.71, indicating acceptable internal consistency. However, the internal consistency of the BI-5 was slightly lower than those of the other two measures (table 1).

The correlations (r_s ≥ 0.92) and agreement (ICC ≥ 0.83) between the FIM motor and 10 item BI were high at admission and discharge, indicating high concurrent validity. However, as table 1 shows, the BI-5 had poor to fair concurrent validity at admission (r_s = 0.74, ICC ≤ 0.55) and fair to high validity at discharge (r_s ≥ 0.92, ICC ≥ 0.74).

The standardised response means (≥ 1.2) showed that the three disability measures were highly responsive in detecting changes in performance of ADL during the hospital stay. The relations between the change scores of the three measures were close (r_s ≥ 0.75, p < 0.001). The changes of the three measures were all significant (p < 0.001; table 1).

DISCUSSION

It has been suggested that a measurement tool should be selected based on empirical evidence and not on clinical relevance.[5] An assessment tool should be scientifically sound in terms of three basic psychometric properties: reliability, validity, and responsiveness.[13] In this study, the psychometric properties of three disability measures (FIM motor subscale, BI, and BI-5) for patients with stroke were concurrently and systematically compared. The findings of this study may provide useful information to both clinicians and researchers who need to choose between competing measures.

The most important finding of this study is that the BI and the FIM motor subscale have clearly acceptable and similar psychometric characteristics in inpatients with stroke. This is somewhat surprising, as the FIM has more items and a wider scoring range than the BI. Interestingly, van der Putten et al[14] also found that the FIM had no advantage over the BI in evaluating change of ADL performance for patients with multiple sclerosis or stroke. Furthermore, Hobart et al[5] found that the BI and the FIM showed similar psychometric characteristics in patients undergoing rehabilitation. These results

suggest that the FIM motor subscale has no advantage over the BI. The BI is quicker and simpler to rate than the FIM. The BI thus seems to be preferable to the FIM motor subscale in measuring basic ADL after stroke.

The psychometric properties of the BI-5 varied but seem acceptable for a simplified measure. At admission, the score distribution and validity of the BI-5 were not well supported. These results are contrary to Hobart and Thompson's findings that both versions were psychometrically equivalent.[6] A possible explanation is that we examined the psychometric properties of these instruments in inpatients with stroke who had severe disability. We found that almost half of our patients, compared with < 5% of the subjects in the study of Hobart and Thompson,[6] had scores of zero on the BI-5 at admission, thus showing that the BI-5 had limited score distribution and reduced validity. However, it is noted that the distribution and validity of the BI-5 at discharge were very adequate. Furthermore, the responsiveness of the BI-5 was high and similar to those of the BI and the FIM motor subscale. Therefore, considering its limitations, the BI-5 is easy to use and thus practical for use in both clinical and research settings.

A potential limitation of the present study is that we did not compare the interrater and intrarater reliability or predictive validity of the instruments. Further studies comparing comprehensive psychometric characteristics in patients with stroke in different stages may be needed to determine which instrument is preferable.

In summary, the results of this study indicate that both the BI and the FIM motor subscale have clearly acceptable and similar psychometric characteristics for patients with stroke during hospital rehabilitation. The psychometric properties of the BI-5 were adequate, with the exception of limited discrimination for patients with severe disability. These results may provide information useful in the selection of ADL measures for both clinicians and researchers.

ACKNOWLEDGEMENTS

....

Authors' affiliations

I-P Hsueh, C-L Hsieh, School of Occupational Therapy, College of Medicine, National Taiwan University, Taiwan, ROC
J-H Lin, School of Rehabilitation Medicine, Kaohsiung Medical University, Taiwan, ROC
J-S Jeng, Department of Neurology, National Taiwan University Hospital, Taiwan, ROC

Correspondence to: Dr C-L Hsieh, School of Occupational Therapy, College of Medicine, National Taiwan University, No 7, Chung-Chan S Rd, Taipei 100, Taiwan, ROC; mike26@ha.mc.ntu.edu.tw

REFERENCES

1 **Mahoney FI**, Barthel DW. Functional evaluation: the Barthel index. *Md State Med J* 1965;**14**:61–5.
2 **Hamilton BB**, Granger CV, Shervin FS, *et al.* A uniform national data system for medical rehabilitation. In: Further MJ, ed. *Rehabilitation outcomes: analysis and measurements.* Baltimore: Paul H Brooks, 1987.
3 **Haigh R**, Tennant A, Biering-Sorensen F, *et al.* The use of outcome measures in physical medicine and rehabilitation within Europe. *J Rehabil Med* 2001;**33**:273–8.
4 **Cohen ME**, Marino RJ. The tools of disability outcomes research functional status measures. *Arch Phys Med Rehabil* 2000;**81**(suppl 2):S21–9.
5 **Hobart JC**, Lamping DL, Freeman JA, *et al.* Evidence-based measurement: which disability scale for neurologic rehabilitation? *Neurology* 2001;**57**:639–44.
6 **Hobart JC**, Thompson AJ. The five item Barthel index. *J Neurol Neurosurg Psychiatry* 2001;**71**:225–30.
7 **Nunnally JC**, Bernstein IH. *Psychometric theory.* New York: McGraw-Hill, 1994.
8 **Daving Y**, Andren E, Nordholm L, *et al.* Reliability of an interview approach to the functional independence measure. *Clin Rehabil* 2001;**15**:301–10.
9 **Hsueh IP**, Lee MM, Hsieh CL. The psychometric characteristics of the Barthel activities of daily living index in stroke patients. *J Formos Med Assoc* 2001;**100**:526–32.
10 **Collin C**, Wade DT, Davies S, *et al.* The Barthel ADL index: a reliability study. *Int Disabil Stud* 1988;**10**:61–3.
11 **Ware JE**, Snow KK, Kosinski M, *et al. SF-36 health survey: manual and interpretation guide.* Boston: The Health Institute, New England Medical Centre, 1993.
12 **Cohen J**. *Statistical power analysis for the behavioural sciences.* Hillsdale: Lawrence Erlbaum Associates, 1988.
13 **Sharrack B**, Hughes RAC, Soudain S, *et al.* The psychometric properties of clinical rating scales used in multiple sclerosis. *Brain* 1999;**122**:141–59.
14 **van der Putten JJMF**, Hobart JC, Freeman JA, *et al.* Measuring change in disability after inpatient rehabilitation: comparison of the responsiveness of the Barthel index and the functional independence measure. *J Neurol Neurosurg Psychiatry* 1999;**66**:480–4.