Comprehensive measurement of long-term outcomes and costs of rehabilitation in patients with stroke
Meijeren-Pont, W. van
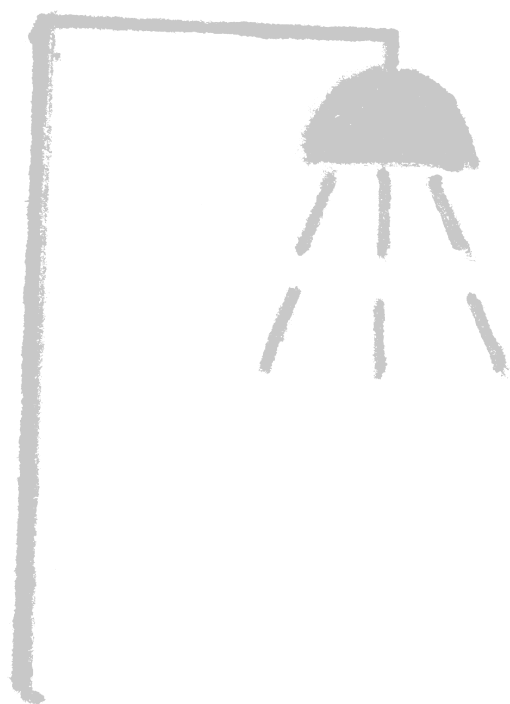
Comparison of the responsiveness of the Utrecht Scale for Evaluation of Rehabilitation (USER) and the Barthel Index in stroke patients

van Meijeren-Pont W | Volker G | Vliet Vlieland TPM | Goossens PH

Abstract

Objective
To compare the responsiveness of the Utrecht Scale for Evaluation of Rehabilitation (USER) to the responsiveness of the Barthel Index in stroke patients in an inpatient rehabilitation facility.

Design
Observational study.

Setting
Inpatient rehabilitation facility.

Subjects
Consecutive stroke patients admitted for clinical rehabilitation.

Interventions
Not applicable.

Main measures
The USER and the Barthel Index were administered by a nurse at admission and discharge. The Effect Size and Standardized Response Mean (SRM) were calculated as measures of responsiveness.

Results
From 198 (78%) of the 254 patients who were included in the study period, both admission and discharge data were available. At admission the mean score of the USER subscale Functional independence was 43.1 (SD = 18.9) and at discharge the mean score was 59.3 (SD = 13.8). The mean score of the Barthel Index at admission was 13.3 (SD = 5.4) and at discharge 18.4 (SD = 3.3). The Effect Size of the USER subscales Mobility, Self-care, Cognitive functioning, Pain, Fatigue and Mood were 0.85, 0.77, 0.48, 0.19, 0.40 and 0.28, respectively and of the Barthel Index 0.94. The results for the SRM were in the same range.

Conclusion
In inpatient rehabilitation after stroke, the USER was less responsive than the Barthel Index.

Introduction

The absolute number of people globally affected by stroke has substantially risen over the past decades, and acute treatment is more effective[1], resulting in an overall increase of the global burden of stroke[2]. The consequences for individual patients are often substantial[3] and thus many of them are in need of rehabilitation. In the Netherlands the majority of patients is discharged home after hospitalization for stroke, and receive rehabilitation in primary care, whereas less than 10% are referred to a rehabilitation centre for medical specialist rehabilitation[4], either in an inpatient or outpatient setting.

Regarding the measurement of outcomes of rehabilitation, appropriate outcome measures are required, both reflecting problems with specific body functions as well as addressing general functioning. Administration of such measures is not only useful for individual patients, with the aim to evaluate the rehabilitation process against set goals[5], but also from the perspective of benchmarking and monitoring the quality of care[6] on the group level.

There are many general outcome measures for stroke patients available, of which the Barthel Index[7] and the Utrecht Scale for Evaluation of Rehabilitation (USER)[8] are two examples. Particularly those two measures were included in the basic set of performance indicators that were accepted as measures of effect of clinical rehabilitation in the Netherlands since 2013[9]. The USER was found to have a good correlation with the Barthel Index and Functional Independence Measure in patients with different diagnoses, including stroke, spinal cord injury, amputation, chronic pain, multiple sclerosis, multiple trauma and various neuromuscular diseases[10]. In addition, it was found that the correlations between the USER Functional independence score and the Barthel Index score at admission and discharge in stroke patients were good[11].

For clinical practice it is important to compare similar outcome measures to see whether one offers more advantages over another, either in terms of responsiveness[12,13] or simplicity and ease-of-use, or in some other domain.

So far, little is known as to what extent all USER subscales are sensitive to changes over time in inpatient stroke rehabilitation nor have their responsiveness been directly compared with that of the Barthel Index. Therefore, the aim of the present study was to determine the responsiveness of each subscale of the USER as compared to the Barthel Index in stroke patients in inpatient rehabilitation.

Methods

This observational study was executed in Basalt Rehabilitation Leiden, the Netherlands. Patients were eligible for the present study if they were admitted for inpatient rehabilitation after between 1 January 2014 and 31 December 2015. It concerned the retrospective analysis of data which was routinely gathered according to a clinical care pathway for stroke patients who were admitted for clinical rehabilitation. For the retrospective analysis of such data, no medical ethical consent is needed according to Dutch law. The handling of data as well as the analysis and reporting was done according to Good Research Practice guidelines[14] and General Data Protection Regulation (GDPR)[15].

Rehabilitation treatment was conducted by a multidisciplinary team, consisting of a rehabilitation specialist, nurse, physical therapist, occupational therapist, speech therapist, social worker, psychologist and other professionals if needed. Treatment goals were set in cooperation with the patients, based on a comprehensive clinical assessment. Every 4 weeks multidisciplinary team conferences were held to evaluate the achievement of treatment goals. At discharge, the clinical assessment was repeated.

Criteria for admission to the rehabilitation centre, as judged by a rehabilitation physician, included recent stroke preventing the patient from living independently at home, being able to take part in at least two therapy sessions of 30 minutes each per day, having some learning ability and expecting to live independently, whether or not with spouse or caregiver, for a life expectancy of at least 1 year. Additional exclusion criteria are a diagnosis of dementia and (neuro)psychiatric conditions interfering with admittance in an open setting.

Clinical and stroke characteristics such as age, sex, stroke localization (left, right or other) and stroke type (ischemic or haemorrhagic) as well as length of stay were extracted from the medical records.

In all patients, the USER was administered by a nurse. It is a measure of functional independence that covers physical functioning (Mobility and Self-care), Cognitive functioning and additional domains of Pain, Fatigue and Mood. The USER has 30 items in total divided over six domains of which three are nurse-reported, and three patient-reported and nurse-recorded. Supplemental Appendix 1 shows the items in each subscale. The nurse-reported domains include Mobility (0-35), Self-care (0-35) and Cognitive functioning (0-50) and the nurse-recorded domains include Pain, Fatigue and Mood. Higher scores on the nurse-reported domains reflect better performance. The scores of the subscales Mobility and Self-care, are aggregated, and defined as Functional independence (0-70)[8]. An improvement of 3 points of the Functional independence scale is considered a small improvement, 7 points a moderate improvement and 14 points a large improvement[9]. For the patient-reported and nurse-recorded subscales

Pain and Fatigue, the scores range from 0 to100. The subdomains Dreariness, Grief, Fear and Anger all with a range from 0 to 100 were summed-up and defined the Mood scale (0-400). Higher scores on the Pain, Fatigue and Mood scales reflect more subjective complaints. The mean time to administer the USER was 10 minutes[8]. The inter-rater reliability of USER was found to be satisfactory to good in patients with diagnoses stroke, chronic pain, spinal cord injury, amputation and other[8].

In addition, the Barthel Index was recorded. The Barthel Index is a fully nurse-reported, 10-item measurement instrument that scores independence in activities of daily living and yields a score between 1 and 20. Higher scores indicate more independency in activities of daily living. The Barthel Index requires no direct testing and should take only minutes to administer[7]. The reported inter-rater reliability in stroke patients ranges from good to very good[16,17].

The USER and Barthel Index were collected by a nurse at admission and discharge of the patient. At the end of the rehabilitation period the USER was repeated, except if the patient was discharged from the rehabilitation facility within 6 weeks. All nurses received training each year to administer the USER. The USER and Barthel Index scores were stored in the patients' electronic medical records. Since the USER and Barthel Index were included in the basic set of performance indicators and a part of daily practice, about 30 nurses were involved in administering the USER and Barthel Index.

If the USER was administered at both admission and discharge, the patient was included in the present analysis. Furthermore, if for a subscale paired measurements were present, the patient was included in the analysis for the concerning subscale. If not, the patient was excluded from the analysis for the concerning subscale. In case of missing items no imputation was executed.

Data analyses were performed in IBM SPSS version 22 v02 (IBM Corp, Armonk, NY, 2013). Descriptive analyses were used for the patient characteristics at admission. Patient characteristics were presented as percentages, means and standard deviations for normally distributed values or medians with 25-75 percentiles (interquartile range (IQR)) and minimum and maximum for non-normally distributed variables.

Characteristics of eligible patients who were admitted in the study period, but not included in the present analysis, were compared with characteristics of patients included in the study by means of the Fisher's exact test, the independent sample's t-test and the Mann-Whitney U-test.

The difference between admission and discharge was examined by means of the paired samples t-test, as the data were normally distributed. The mean change, with the 95%

confidence interval (CI) between admission and discharge, was calculated for all USER subscales and the Barthel Index. The same was calculated without the patients with the lowest or highest possible scores. The lowest score was defined as 0 and the highest score as the highest possible score on a scale.

The responsiveness is the ability of a questionnaire to detect clinically important changes over time[13]. For each subscale of the USER the responsiveness was determined using two methods, the Standardized Response Mean (SRM) and the Effect Size. The mean change divided by the standard deviation for the mean change was used to calculate the SRM for each subscale. The Effect Size for each subscale was calculated using Cohen's d. The following formula was used to calculate the Effect Size (Cohen's d): $d = (M_0 - M_1)/SD$ in which M_0 and M_1 are the means of the baseline and follow-up measurements and SD is the standard deviation of the baseline measurement[18]. Cohen's d under 0.20 is considered a trivial effect, between 0.20 and 0.49 a small effect, between 0.50 and 0.79 a moderate effect and above 0.80 a large effect[18]. The SRM and Effect Size were calculated for all patients included in the analyses and for a subgroup without patients who had the highest or lowest possible scores. The responsiveness of the USER was compared to the responsiveness of the Barthel Index.

Floor or ceiling effects were considered to be present if more than 15% of the respondents has the lowest or highest possible score[19]. If floor and ceiling effects are present, the reliability is reduced, because patients with the lowest or highest possible score cannot be distinguished from each other. Furthermore, the responsiveness is limited, because changes cannot be measured in these patients[19]. The floor and ceiling effects were determined at admission by calculating frequency distributions.

The correlation between the USER subscale Functional independence and Barthel Index was determined using linear regression analysis.

The proportions of patients who improved, remained the same or deteriorated between admission and discharge were calculated and presented as percentages. Improvement on the subscales Mobility, Self-care, Cognitive functioning, the combined Functional independence scale and the Barthel Index was defined as a higher score at discharge than admission and deterioration as a higher score at admission than discharge. On the subscales Pain, Fatigue and Mood, improvement was defined as a lower score at discharge than admission and deterioration as a higher score at discharge than admission. A patient had a stable score when the scores at admission and discharge were exactly the same. The same was calculated without the patients with the lowest and highest possible scores.

In all statistical analyses, a p value of <0.05 was considered as statistically significant.

Results

In 2014–2015, 254 patients with a stroke were admitted for clinical rehabilitation at Basalt Rehabilitation. Figure 1 shows the flow of participants in the study. At admission, the USER scores of 240 (94.5%) patients were available. Data of 14 (5.5%) patients could not be retrieved from the database for unknown reasons. At discharge, the paired USER scores of 198 (78.0%) patients were available.

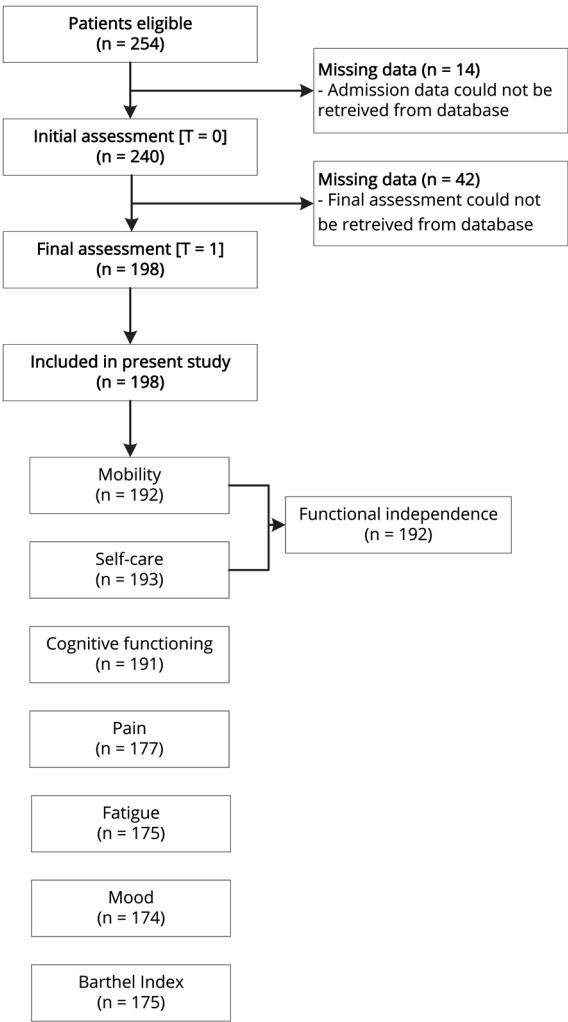**3**

Figure 1. Flow chart of stroke patients admitted for clinical rehabilitation in one rehabilitation centre in 2014-2015 and of paired measurements for each subscale of the Utrecht Scale for Evaluation of Rehabilitation and Barthel Index.

Table 1 presents the characteristics of patients who were and who were not included in the study. There was no significant difference between the patients who were and who were not included in the study in age, percentage of males, localization of stroke, the time between stroke and admission or USER and Barthel Index scores at admission (*P*-values not shown). However, patients who were included in the present study had a significantly longer length of stay than patients who were not included in the study (*p* value 0.005).

Table 1. Characteristics of stroke patients admitted for clinical rehabilitation in one rehabilitation centre between 2014 and 2015 included and not included in the study.

	Stroke patients included in the study n = 198	Stroke patients not included in the study n = 42
Age, years (mean, SD)	61.5 (11.8)	58.9 (14.0)
Male (n, %)	125 (63%)	25 (60%)
Stroke localization		
Left (n, %)	91 (46%)	14 (41%)
Right (n, %)	67 (34%)	13 (31%)
Other (n, %)	26 (13%)	9 (21%)
Unknown (n, %)	14 (7%)	3 (7%)
Stroke type, ischemic (n, %)	135 (68%)	
Barthel Index (mean, SD)	13.6 (5.4) n = 192	14.0 (5.8)
Time between stroke and admission, days (median, (25-75 percentiles))	11 (7-20 n = 194	11.5 (7-19)
Length of stay, days (median, (25-75 percentiles))	52 (31-78)	35 (25-56)

Table 2 shows the USER and Barthel Index scores at admission and discharge and the Effect Size and SRM. The Effect Size for the USER subscale Pain was trivial; for the USER subscales Cognitive functioning, Fatigue and Mood the Effect Sizes were small and for the USER subscale Self-care moderate. The Effect Sizes of the Mobility and Functional independence scales and the Barthel Index were large. The calculation of the SRM yielded similar results, and a similar ranking.

The presence of floor and ceiling effects in each subscale at admission is also shown in Table 2.

Table 2. Admission and discharge scores of stroke patients admitted for clinical inpatient rehabilitation and numbers and percentages of stroke patients with the lowest and highest scores on the USER and Barthel Index at admission.

|	Admission	Discharge	Mean Change (SD, CI)	SRM	ES	Lowest score (n,%)	Highest score (n,%)
All patients included in the present study							
USER							
Functional independence (0-70) (mean, SD) n = 192	43.1 (18.9)	59.3 (13.8)	16.3 (13.8)	1.18	0.86	1 (0.4) n = 240	21 (8.8) n = 240
Mobility (0-35) (mean, SD) n = 192	19.1 (10.5)	28.0 (8.2)	8.9 (8.2)	1.09	0.85	3 (1.3) n = 240	27 (11.3) n = 240
Self-care (0-35) (mean, SD) n = 193	23.9 (9.6)	31.3 (6.3)	7.4 (7.3)	1.01	0.77	1 (0.4) n = 240	44 (18.3)[a] n = 240
Cognitive functioning (0-50) (mean, SD) n = 191	37.5 (10.9)	42.7 (9.3)	5.1 (8.4)	0.61	0.48	1 (0.4) n = 237	29 (12.2) n = 237
Pain (0-100) (mean, SD) n = 177	17.9 (23.8)	13.4 (23.0)	-4.5 (24.5)	0.18	0.19	126 (55.0)[b] n = 229	1 (0.4) n = 237
Fatigue (0-100) (mean, SD) n = 175	40.9 (25.9)	30.5 (25.8)	-10.4 (28.5)	0.36	0.40	42 (18.4)[b] n = 228	3 (1.3) n = 228
Mood (0-400) (mean, SD) n = 174	66.7 (68.8)	47.5 (74.6)	-19.2 (79.8)	0.24	0.28	64 (28.3)[b] n = 226	0 (0.0) n = 226
Barthel Index (0-20) (mean, SD) n = 175	13.3 (5.4)	18.4 (3.3)	5.1 (4.3)	1.19	0.94	1 (0.4) n = 232	38 (16.4)[a] n = 232
Without patients with the lowest or highest scores on the subscales Self-care, Pain, Fatigue and Mood and on the Barthel index							
USER Self-care (0-35) (mean, SD) n = 160	21.6 (8.9)	30.6 (6.6)	9.0 (7.1)	1.26	1.01		
Pain (0-100) (mean, SD) n = 80	39.6 (19.7)	21.1 (28.1)	-18.5 (26.4)	0.70	0.94		
Fatigue (0-100) (mean, SD) n = 147	48.7 (20.5)	33.8 (25.1)	-14.8 (27.5)	0.54	0.73		
Mood (0-400) (mean, SD) n = 130	89.3 (65.7)	57.7 (79.5)	-31.6 (84.7)	0.37	0.48		
Barthel Index (0-20) (mean, SD) n = 149	12.2 (5.0)	18.1 (3.5)	6.0 (4.1)	1.47	1.18		

USER: Utrecht Scale for Evaluation of Rehabilitation; CI: Confidence Interval; SRM: Standardized Response Mean; ES: Effect Size; SD: Standard Deviation.
[a]Ceiling effect presents.
[b]Floor effect presents.

Table 2 also shows the Effect Size and SRM calculated without the patients who had the lowest or highest scores. The Effect Size of the Barthel Index remained large and of the USER subscale Mood remained small. The Effect Sizes of the USER subscales Self-care, Pain and Fatigue increased from moderate to large, trivial to large and small to moderate, respectively.

Figure 2 shows a scatterplot relating the results of the Barthel Index to the USER subscale Functional independence. R^2 equals 0.86.

Figure 2. Scatterplot comparing the Barthel Index with the USER subscale Functional independence.
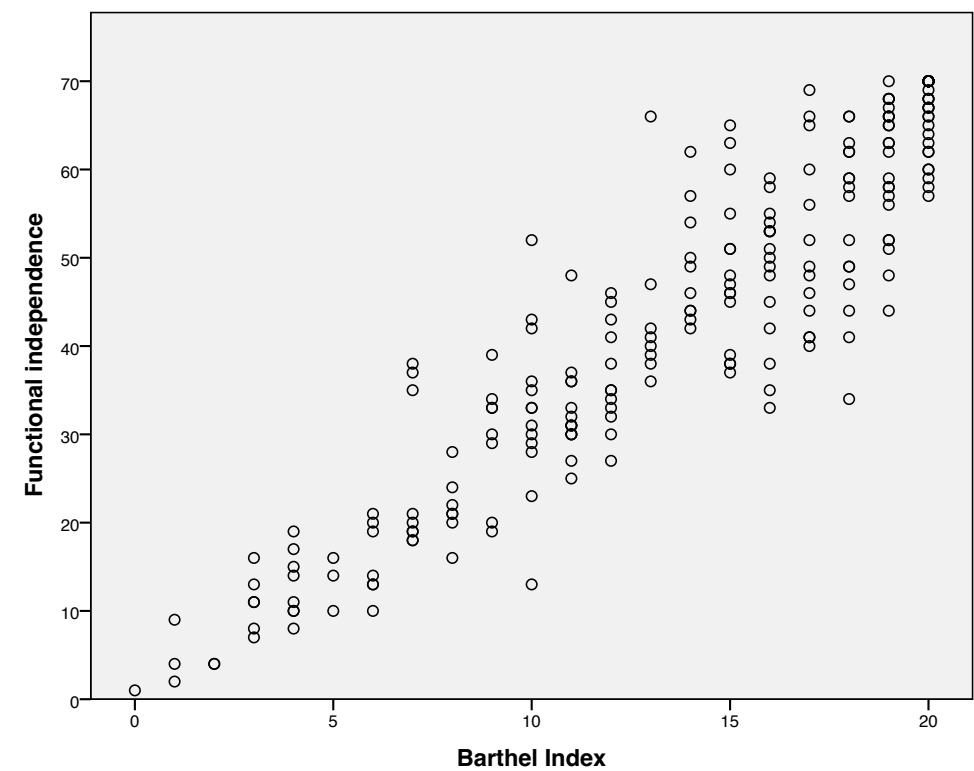
USER: Utrecht Scale for Evaluation of Rehabilitation.

Table 3. Numbers and percentages of stroke patients improving, being stable or deteriorating according to the Barthel Index or Utrecht Scale for Evaluation of Rehabilitation (USER) between admission and discharge in inpatient rehabilitation.

	All patients included in the present study			Without patients with the lowest or highest possible score		
	Deterioration (n, %)	Unchanged (n, %)	Improvement (n, %)	Deterioration (n, %)	Unchanged (n, %)	Improvement (n, %)
USER						
Functional independence	7 (3.6) (n = 192)	15 (7.8) (n = 192)	170 (88.5) (n = 192)			
Mobility	9 (4.7) (n = 192)	23 (12.0) (n = 192)	160 (83.3) (n = 192)			
Self-care	8 (4.1) (n = 193)	34 (17.6) (n = 193)	151 (78.2) (n = 193)	7 (4.4) (n = 160)	2 (1.3) (n = 160)	151 (94.4) (n = 160)
Cognitive functioning	22 (11.5) (n = 191)	36 (18.8) (n = 191)	133 (69.6) (n = 191)			
Pain	36 (20.3) (n = 177)	76 (42.9) (n = 177)	65 (36.7) (n = 177)	12 (15.0) (n = 80)	3 (3.8) (n = 80)	65 (81.3) (n = 80)
Fatigue	37 (21.1) (n = 175)	38 (21.7) (n = 175)	100 (57.1) (n = 175)	27 (18.4) (n = 147)	20 (13.6) (n = 147)	100 (68.0) (n = 147)
Mood	38 (21.8) (n = 174)	40 (23.0) (n = 174)	96 (55.2) (n = 174)	27 (20.8) (n = 130)	7 (5.4) (n = 130)	96 (73.8) (n = 130)
Barthel Index	2 (1.1) (n = 175)	27 (15.4) (n = 175)	146 (83.4) (n = 175)	2 (1.3) (n = 149)	1 (0.7) (n = 149)	146 (98.0) (nf = 149)

USER: Utrecht Scale for Evaluation for of Rehabilitation; improvement: a higher score at discharge than admission; deterioration: a higher score at admission than discharge; unchanged: scores at admission and discharge were exactly the same.

Table 3 shows the numbers and percentages of patients who improved, remained the same or deteriorated on the subscales of the USER and on the Barthel Index. The largest proportions of patients (>70%) showing an improvement were seen for the USER subscales Mobility, Self-care, the combined Functional independence scale and the Cognitive functioning subscale. For the USER subscales Pain, Fatigue and Mood not only the proportions of patients who improved were lower than for the other subscales but also the proportions of patients who showed a deterioration were higher.

Table 3 also shows the numbers and percentages of patients who improved, remained the same or deteriorated on the subscales of the USER and on the Barthel Index without patients with the lowest or highest possible scores. Without patients with the highest or lowest possible score, the proportion of patients who remained the same decreased and the proportion of patients who improved increased. The proportion of patients who deteriorated was comparable between the two groups.

Discussion

This study embedded in daily practice found that overall the Barthel Index is more responsive than the USER between admission and discharge of inpatient rehabilitation in patients with stroke. The Barthel Index-related USER subscales Mobility, Self-care and Functional independence were the most responsive. The changes seen for the subscales Cognitive functioning, Fatigue, Mood and Pain were moderate to small. The lack of responsiveness of some USER subscales may in part be related to the observed ceiling and floor effects. When calculations of the Effect Size and SRM were repeated without patients with the lowest or highest possible score, the Effect Sizes of the USER subscales Self-care, Pain and Fatigue increased from moderate to large, trivial to large and small to moderate, respectively. This indicates that the floor and ceiling effects limit the responsiveness. However, the Effect Size of the USER subscale Mood still remained small and the Barthel Index still was more responsive than the most responsive subscale of the USER. Without patients with the lowest or highest possible scores, the proportion of patients who remained the same decreased and the proportion of patients who improved increased. This shows that the floor and ceiling effects observed in the USER subscales Self-care, Pain, Fatigue and Mood limit the responsiveness, because changes cannot be measured in these patients[19].

Our findings regarding USER scores are in line with results from a previous study in stroke patients executed in five different rehabilitation centres in the Netherlands. The average admission scores of 42.9, 39.1, 16.8, and 37.9 for Functional independence, Cognitive functioning, Pain and Fatigue in that study[10] were in the same range of the results of our patient group. The average admission score for the subscale Mood was higher (84.3) than we found in the present study (66.7). In that study no effect sizes or other measures of responsiveness were calculated and no comparison was made with another measure commonly used in inpatient rehabilitation, for example, the Barthel Index.

So far the responsiveness of the USER subscales has only been calculated in a population of patients with several diagnoses together, including stroke, chronic pain, spinal cord injury, amputation and other[8]. Similar to our results, in that study the subscales Mobility, Self-care and Functional independence showed large effects. The Cognitive functioning, Pain, Fatigue and Mood subscales showed a small effect. In an attempt to increase responsiveness the subscales Pain, Fatigue and Mood were changed in the process of developing the USER by the developers from a 0 to 3 scale to an 11-level numerical rating scale[8]. No further calculations of responsiveness have been done until the present study.

At admission, a floor effect was present for the subscale Pain, as 126 (55%) patients reported that they had no pain. For 35 patients (20%), the pain they experienced increased during the clinical rehabilitation period. In the literature, the reported prevalence of post-stroke pain

varies[20]. However, there is a general consensus that post-stroke pain is an underreported and under-detected phenomenon[21,22].

A limitation of this study is that it concerns a selection of patients, with patients included and not included in the study, having a significantly different length of stay. The 42 (16.5%) patients who were not included in the study and for whom no discharge data were available had significantly shorter length of stay in the rehabilitation centre than the patients for whom data on admission and discharge was available. For patients discharged from the rehabilitation centre within 6 weeks no final assessment of the USER was available, because nurses were initially trained to repeat the USER and Barthel Index every 6 weeks. Another reason might be that patients with missing data were physically independent, as was reported in a previous study concerning the main reason for the exclusion of patients missing USER data[10].

Our results may have been distorted by the fact that about 30 nurses were involved in gathering the USER and Barthel Index data, while the intra-rater reliability could not be determined. Although the relatively large number of nurses involved in data gathering reflects the reality of every day practice, a recommendation for future research is to involve a limited number of nurses in data gathering and determine the intra-rater reliability.

A way to test responsiveness is to relate the Smallest Detectable Change to the minimal important change[19]. The minimal important change is 'the smallest difference in score in the domain of interest which patients perceive as beneficial and would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management'[23]. An anchor-based method, which uses an external criterion to determine an important change, is recommended to determine the minimal important change[19]. For future research, we recommend this method to determine the responsiveness.

In ordinal-based outcome measures, the distances between scores are separated by unknown quantities of the measured variable. Therefore, the unit distance between adjacent categories can vary in meaning across the scale[24]. To overcome this, Rasch analysis can be used[25]. This is based on log-odd transformation that determines the extent to which the observed responses fit the pattern formalized by the model. Emerging evidence shows that, compared with standard scores of ordinal scales, outcome measurement subjected to Rasch analysis shows a higher magnitude of meaningful changes over time[26,27]. For future research, we recommend that this measure is used in order to determine responsiveness.

The findings of our study may have implications for clinical practice. As indicated in the introduction, in the Netherlands the use of both the Barthel Index and the USER is advocated. Yet our study clearly demonstrated that the Barthel Index is more responsive to changes over time than the USER. As it is also more easy to administer than the USER in total, suggesting that

the use of only the Barthel Index to measure general functioning should be recommended. However, as apart from general functioning, specific functions such as speech or cognitive functioning may be affected by stroke; the use of additional outcome measures should be considered.

The Barthel Index was more responsive than the most responsive subscales of the USER in inpatient rehabilitation after stroke. A potential advantage of the USER over the Barthel Index is that it comprises dimensions of functioning other than physical functioning, such as Pain, Fatigue and Mood. However, exactly these dimensions were found to be relatively insensitive to changes over time, probably due to observed floor effects. Therefore, in clinical practice, it could be considered to use only the Barthel Index as a measure of independence in activities of daily living in stroke patients. In order to measure other common problems in stroke patients besides limitations in independence in activities of daily living, such as aphasia an cognitive functioning, other measures should be used.

Clinical messages

- The Barthel Index was more responsive than the most responsive subscales of the Utrecht Scale for Evaluation of Rehabilitation (USER) in inpatient rehabilitation after stroke.

- In clinical practice, it could be considered to use only the Barthel Index as a measure of the effect of clinical rehabilitation in stroke patients.

References

1.   Katan M, Luft A. Global Burden of Stroke. Semin Neurol. 2018;38(2):208-11.

2.   Feigin VL, Mensah GA, Norrving B, Murray CJ, Roth GA; GBD 2013 Stroke Panel Experts Group. Atlas of the Global Burden of Stroke (1990-2013): The GBD 2013 Study. Neuroepidemiology. 2015;45(3):230-6.

3.   Erikson A, Karlsson G, Tham K. Living with the long-term consequences 11-13 years after stroke: A phenomenological study. J Rehabil Med. 2016;48(10):847-52.

4.   Nieboer A, Pepels R, van der Have L, Kool T, Huijsman R. Stroke services gespiegeld. Publieksversie van de haalbaarheidsstudie naar een landelijke benchmark van CVA-ketenzorg. Den Haag: ZonMw; 2005.

5.   Warlow C, van Gijn J, Dennis M, Wardlaw J, Bamford J, Hankey G, et al. Stroke: practical management. 3rd ed. Oxford: Blackwell Publishing, 2008.

6.   Berdenis van Berlekom S, Koppe P. To be continued, ontwikkeling van prestatie-indicatoren in de revalidatie. [To be continued, development of perfromance indicators in rehabilitation]. Revalidata 2009;31:12-5.

7.   Quinn TJ, Langhorne P, Stott DJ. Barthel index for stroke trials: development, properties, and application. Stroke. 2011;42(4):1146-51.

8.   Post MW, van de Port IG, Kap B, Berdenis van Berlekom SH. Development and validation of the Utrecht Scale for Evaluation of Clinical Rehabilitation (USER). Clin Rehabil. 2009;23(10):909-17.

9.   Willems M, Berdenis van Berlekom S, van Asbeck F, Post M. The continuing story: USER als uitkomstmeting van klinische revalidatie. [The continuing story: USER as outcome measure of clinical rehabilitation]. Revalidata. 2010;156:6-9.

10.   Ten Brink AF, Hajos TR, van Bennekom C, Nachtegaal J, Meulenbelt HE, Fleuren JF, et al. Predictors of physical independence at discharge after stroke rehabilitation in a Dutch population. Int J Rehabil Res. 2017;40(1):37-45.

11.   Ten Brink AF, Hajos TRS, van Bennekom CAM, Nachtegaal J, Meulenbelt HEJ, Fleuren JFM, et al. De Utrechtse Schaal voor de Evaluatie van klinische Revalidatie (USER): implementatie en normgegevens. [The Utrecht Scale for Evalutation of clinical rehabilitation (USER): implementation and norms]. Nederlands Tijdschrift voor Revalidatiegeneeskunde. 2017;3:105-10.

12.   Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chronic Dis. 1985;38(1):27-36.

13.   Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: a clarification. J Clin Epidemiol. 1989;42(5):403-8.

14.   Good Research Practice & Integrity. [Internet]. Available from: https://www.lumc.nl/research/grp-and-integrity/. [Accessed 10th August 2017].

15.   General Data Protection Regulation (GDPR). [Internet]. Available from: https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/algemene-informatie-avg/algemene-informatie-avg. [Accessed 16th May 2019].

16.   Hsueh IP, Lee MM, Hsieh CL. Psychometric characteristics of the Barthel activities of daily living index in stroke patients. J Formos Med Assoc. 2001;100(8):526-32.

17.   Green J, Forster A, Young J. A test-retest reliability study of the Barthel Index, the Rivermead Mobility Index, the Nottingham Extended Activities of Daily Living Scale and the Frenchay Activities Index in stroke patients. Disabil Rehabil. 2001;23(15):670-6.

18.   Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1977.

19.   Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42.

20.   Paolucci S, Martinuzzi A, Scivoletto G, Smania N, Solaro C, Aprile I, et al. Assessing and treating pain associated with stroke, multiple sclerosis, cerebral palsy, spinal cord injury and spasticity. Evidence and recommendations from the Italian Consensus Conference on Pain in Neurorehabilitation. Eur J Phys Rehabil Med. 2016;52(6):827-40.

21.   Harrison RA, Field TS. Post stroke pain: identification, assessment, and therapy. Cerebrovasc Dis. 2015;39(3-4):190-201.

22.   Henon H. Pain after stroke: a neglected issue. J Neurol Neurosurg Psychiatry. 2006;77(5):569.

23.   Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407-15.

24.   Piscitelli D, Pellicciari L. Responsiveness: is it time to move beyond ordinal scores and approach interval measurements? Clin Rehabil. 2018;32(10):1426-7.

25.   Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum. 2007;57(8):1358-62.

26.   Draak TH, Vanhoutte EK, van Nes SI, Gorson KC, Van der Pol WL, Notermans NC, et al. Changing outcome in inflammatory neuropathies: Rasch-comparative responsiveness. Neurology. 2014;83(23):2124-32.

27.   Doganay Erdogan B, Leung YY, Pohl C, Tennant A, Conaghan PG. Minimal clinically important difference as applied in rheumatology: an OMERACT rasch working group systematic review and critique. J Rheumatol. 2016;43(1):194-202.

Appendix 1. Utrecht Scale for Evaluation of Rehabilitation (USER) scales and items[10]

Domains	Scales	Items
Functional independence	Mobility	Sitting
		Standing
		Transfers
		Walking indoors
		Walking longer distances
		Climbing stairs
		Wheelchair mobility
	Self-care	Eating and drinking
		Grooming
		Bathing/showering
		Dressing/undressing
		Toileting
		Incontinence bladder
		Incontinence bowels
Cognitive functioning		Expressing
		Understanding
		Visual perception
		Orientation in place and time
		Attention and concentration
		Memory
		Task execution
		Initiative
		Behaviour control
		Social behaviour
Subjective complaints	Pain	
	Fatigue	
	Mood	Depressed mood
		Grief
		Anxiety
		Anger

3