

ORIGINAL ARTICLE

Further methodological development of the Test Instrument for Profile of Physical Ability (TIPPA) designed for patients with long-term musculoskeletal pain

MINAJ MISSAGHI-WEDEFALK¹, MALIN LINDH², CHRISTINA SCHÖN-OHLSSON³ & CARIN WILLEN⁴

¹The Primary Health Care & Rehabilitation in Gothenburg, Sweden, ²Institute of Neuroscience and Physiology/Rehabilitation Medicine, The Sahlgrenska Academy at Gothenburg University, Gothenburg, Sweden, ³Institute of Neuroscience and Physiology/Physiotherapy, The Sahlgrenska Academy at Gothenburg University, Gothenburg, Sweden, and ⁴Institute of Neuroscience and Physiology/Rehabilitation Medicine, The Sahlgrenska Academy at Gothenburg University, Gothenburg, Sweden

Abstract

In a heterogeneous group of patients with long-term musculoskeletal pain (LMP), there is a need for a short test to assess physical ability. The Test Instrument for Profile of Physical Ability (TIPPA) includes observed physical ability and pain behaviour as well as the individual's self-assessment of his or her ability and experience of exertion. The relation between these four dimensions of the individual is made visible in a profile. The aim of this study was to investigate the inter-rater reliability of pain behaviour, to reconstruct the limits for the grades in the nine physical activities and to investigate the test-retest reliability of these reconstructed grades. The results showed that the percentual agreement between the two physiotherapists who observed LMP patients ($n = 7$) was 100% for seven of the activities and 86% for the two remaining activities. The reconstruction of the grading system was based on the results of healthy persons ($n = 48$) and LMP patients ($n = 658$) performing the nine physical activities. For test-retest of the reconstructed grading system in a patient group ($n = 13$), no disagreement was found between occasions but a systematic disagreement was found in one activity. This study confirms that the TIPPA instrument is reliable and the grading system discriminates between healthy persons and LMP patients.

Key words: chronic pain, pain behaviour, physiotherapy, rehabilitation, reliability

Introduction

Demands are being made to assess individuals' functional capacity in medical and work-oriented rehabilitation and to judge work ability. Appropriate tests are needed to form a basis for decisions on rehabilitation measures and for the management of insurance systems. A number of valid and reliable methods have been developed that aim to measure the patient's functional ability and to assess whether the patient has the capacity to meet the demands of the physical part of the job (1–3). Many of the Functional Capacity Evaluations (FCE) that are intended to assess a patient's ability to perform work-related and non-work-related activities of daily living or acts are far too comprehensive and thus time-consuming and costly (4).

There is a need especially for the many patients with long-term musculoskeletal pain (LMP) for a short test instrument to identify and assess physical functional ability that also relates to the patient's self-assessment and pain behaviour. LMP patients constitute a heterogeneous group of individuals with different diagnoses and impairments, disabilities, pain localizations, individual personality traits, ages and sexes (5,6). LMP often leads to physical disability and is one of the main reasons for sick leave (7). Evidence from a recent systematic review confirms the complexity of the LMP in view of biopsychosocial risk prognostic factors for long-term disability (5). The biopsychosocial perspective includes the patients' self-efficacy, which has been proven to be low, and self-rated physical disability, which is often

overestimated (8–10). Moreover, discrepancies have been found between the patients' self-rated assessments and their observed abilities (11). LMP patients often show signs of 'pain behaviour' in audible–visible, affective–behavioural dimensions to communicate pain and suffering (12,13). Many self-administrated questionnaires have been developed in which patients assess their physical and psychosocial abilities (14). There is a need to supplement patients' self-reports with functional physical tests, and vice versa, as patients' own reporting of disability may be only moderately associated with how tests are carried out (15,16).

The above-mentioned factors constitute the background for the development of the Test Instrument for Profile of Physical Ability (TIPPA), which aims at evaluating patients' physical capacity in a rehabilitation clinic. When the instrument was designed, one requirement was that the physical activities included should be easy to carry out in a clinical setting and need only inexpensive materials. Another demand was that not only physical tests but also an assessment of observed pain behaviour and a self-rated measure of ability should be included in the procedure (17).

The TIPPA comprises four variables (A, B, C, D):

- A: The patient's self-rating of current physical ability;
- B: Assessment of physical performance of nine selected physical activities;
- C: The patient's estimation of perceived exertion;
- D: Assessment of observed 'pain behaviour' (Table I).

The Appendix describes the TIPPA variables A–D.

In an earlier study aiming to establish the reliability of TIPPA, the content validity of the nine activities in variable B was evaluated by a group of experts ($n = 7$) that appraised eight of activities as being representative of everyday life activities (17). The test–retest reliability of the nine activities has

been proven to be satisfactory for patients with LMP when it was tested by different physiotherapists (PTs) in several studies (6,17,18). Further test–retest of the whole instrument (variables A–D) was done in an LMP patient group ($n = 13$) and showed high reliability for variables B + D ($r = 0.860–0.987$) and moderate reliability for variable A ($r = 0.601$) (17). A modified form of the University of Alabama UAB Pain Behaviour Scale was constructed to assess pain behaviour (variable D) (19,20) (Table I). The preliminary TIPPA grading system was constructed on the basis of a comparison between a healthy group ($n = 48$) and a patient group ($n = 45$) performing the nine physical activities. The grading system has five grades for each variable with separate values for men and women, as it is shown that men outperform women irrespective of patient or non-patient status (21). Grade I represents a very low level and grade V a very high level. The grading system allows the test results to be illustrated graphically as an individual profile (see example in Figure 1), and two assessments can easily be compared. However, as the grades were based on a small sample, the limits were not accurate enough to discriminate all patients from the sample of healthy persons and needed to be reconstructed.

This study aims to develop further the TIPPA instrument to:

- (1) investigate the inter-rater reliability of the scoring of pain behaviour (variable D);
- (2) reconstruct the limits for the grades in nine physical activities (variable B);
- (3) investigate the test–retest reliability of the reconstructed grades I–V in nine physical activities (variable B).

Material and methods

(1) Inter-rater reliability of the scoring of pain behaviour (variable D)

The first author and a PT from the same clinic experienced in the area of working with LMP patients observed the patients ($n = 7$) performing the nine physical activities in a clinical setting.

Eight criteria for pain behaviour in TIPPA were chosen from the UAB Pain Behaviour Scale (19,20) (Table I). Two additional criteria were constructed for TIPPA (Table I). Physical impairments that would influence the performance of an activity are also counted as pain behaviour. The scoring in TIPPA is different from that in UAB Pain Behavior Scale. One or more signs of pain behaviour of any criteria may occur in each activity, and it is the number of activities with pain behaviour that make up the grades I–V. The instructions for scoring pain behaviour are specially standardized for the position

Table I. Criteria for pain behaviour and function during performance of nine physical activities of the Test Instrument for Profile of Physical Ability (TIPPA).

1. Vocal complaints: verbal
2. Vocal complaints: non-verbal
3. Facial grimace
4. Standing posture
5. Mobility
6. Body language
7. Use of visible supportive equipment
8. Stationary movement
9. Interruption or total avoidance of an activity
10. Changing the standard position of the test

The first eight criteria are as in the UAB Pain Behavior Scale (19,20) and the last two criteria are constructed for TIPPA. Note: The scoring model differs from the UAB Pain Behavior Scale.

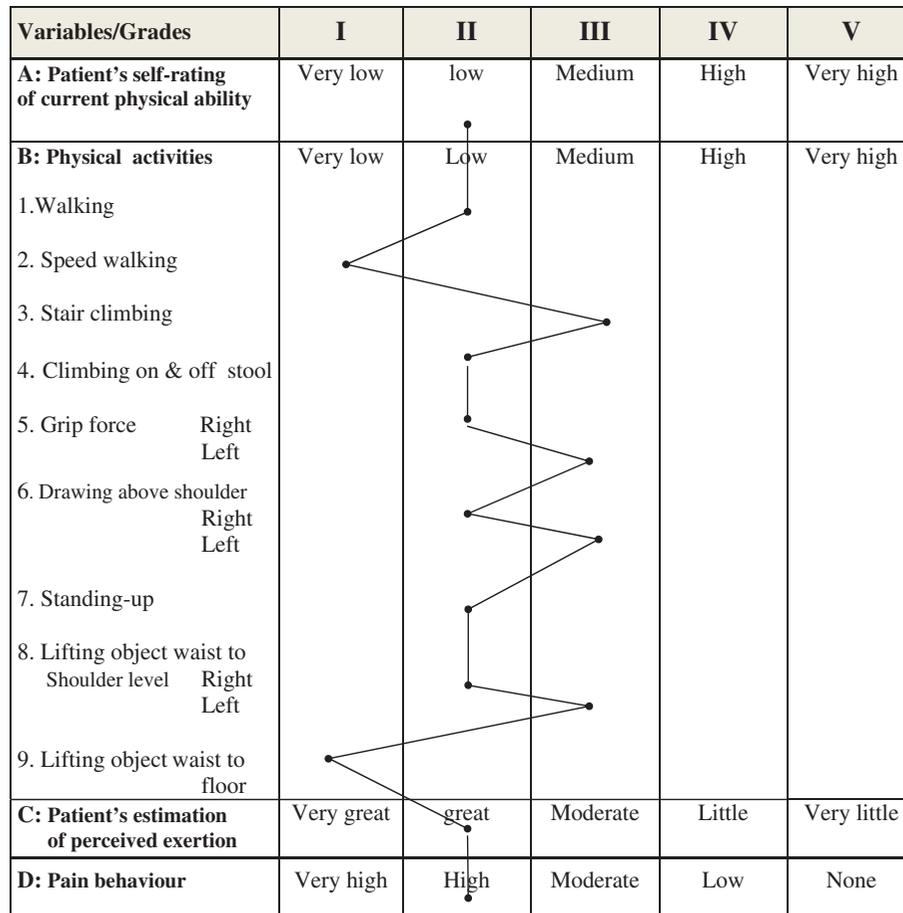


Figure 1. An example of a Test Instrument for Profile of Physical Ability (TIPPA) graphic profile.

of each activity. Each activity in which any kind of pain behaviour is demonstrated counts as 1 point and 0 if no pain behaviour is demonstrated. The scale consists of: I = very high (pain behaviour shows in seven to nine activities), II = high (pain behaviours show in five or six activities), III = moderate (pain behaviours show in three or four activities), IV = low (pain behaviours show in one or two activities) and V = no pain behaviour.

(2) *Reconstruction of the limits for the grades in nine physical activities (variable B)*

The test results of 658 patients (Table II) were collected from 34 PTs who worked with LMP patients in 19 different Swedish primary health care centres, hospitals and private clinics. The LMP diagnoses included both localized and widespread pain > 4 months. All PTs were skilled in using TIPPA and followed the standardized test instructions. The patients could not be identified as test results were not coded and only age, sex and diagnoses were reported.

Furthermore, the test results from a group of 48 healthy persons (Table II), 20–60 years old, were

included in the adjustment of the grading system. Inclusion criteria for this group were voluntarily participation, self-report of being healthy and presently employed and working. The grading system was reconstructed on the basis of measurement values in patients ($n = 658$) and healthy persons ($n = 48$) by distributing the measurement results into 25th, 50th and 75th percentiles. Women's and men's values were treated separately, as were the values in the patient group and the group of healthy persons.

Table II. Descriptive data of age for the patients ($n = 658$) and healthy persons ($n = 48$) women and men respectively.

Age group	Patients		Healthy persons	
	Women	Men	Women	Men
20–29	30	12	2	2
30–39	105	43	9	4
40–49	172	44	11	8
50–59	134	42	9	3
60–65	26	11	0	0
Missing	23	16	0	0
Median	45	43	41.5	42.5
Total	490	168	31	17

(3) Test-retest reliability of the reconstructed grades I–V in nine physical activities (variable B)

The nine physical activities (variable B) were tested twice by another PT than the author, within 7–10 days, in a group of LMP patients with >4 months pain (n = 13, seven women, six men).

Ethical approval

The part of the study involving patients was approved by the Ethics Committee of Gothenburg University.

Statistics

The analysis was made using the PASW 18.0. Median and minimum to maximum values were used to describe data. Percentual agreement (PA) was used for the test of agreement between the observers. The new grading system was calculated from the 25th, 50th and 75th percentiles among LMP patient group and the healthy group. A rank-transformable statistical approach for ordinal data was used to investigate test-retest (22,23). The relative rank variance (RV) is a measure of occasional disagreement. Possible values of RV range from 0 to 1. Relative position (RP) is a measure of systematic disagreement and can be demonstrated in a receiver operating characteristic (ROC) curve. The possible values of RP are -1 to 1, and zero means a lack of systematic disagreement. A p-value of ≤ 0.05 was considered statistically significant.

Results

(1) Inter-rater reliability of the scoring of pain behaviour (variable D)

The PA between the two observers was 100% for seven of the nine activities and 86% for the two remaining

activities: stair climbing and drawing above shoulder level (Table III). The two PTs agreed on all types of pain behaviour with the exception of patients 3 and 6.

(2) Reconstruction of the limits for the grades in nine physical activities (variable B)

The reconstructed grading system based on the percentiles is outlined in Figure 2:

- Grade I – the upper limit is the value of the 25th percentile in the patient group and the lower limit is the lowest registered value in the patient group.
- Grade II – the upper limit is the median value in the patient group corresponding to the lowest value in the healthy group. The lower limit is the value of the 25th percentile in the patient group.
- Grade III – the upper limit is the value of the 25th percentile in the healthy group. The lower limit is the lowest value in the healthy group corresponding to the median of the patient group.
- Grade IV – the upper limit is the value of the 75th percentile in the LMP patient group. The lower level is the value of the 25th percentile in the healthy group.
- Grade V – the upper limit is the highest value (100th percentile) in the healthy group and the lower limit is the value of the 75th percentile in the same group.

The value of the grades and the frequency of the patients and healthy persons for each grade among women and men are shown in Tables IVa and IVb.

(3) Test-retest reliability of the reconstructed grades I–V in nine physical activities (variable B)

The raw data for the nine physical activities expressed in median and minimum to maximum values are

Table III. Inter-rater reliability between two observers in assessment of pain behaviour among long-term musculoskeletal pain patients (n = 7).

Selected physical activities	Patients							Agreement
	1	2	3	4	5	6	7	
1. Walking	○◇	○◇	●◆	○◇	○◇	●◆	●◆	100%
2. Speed walking	○◇	○◇	●◆	○◇	●◆	●◆	●◆	100%
3. Stair climbing	○◇	○◇	●◆	○◇	●◆	●◇	●◆	86%
4. Climbing on and off a stool	○◇	○◇	●◆	○◇	○◇	●◆	●◆	100%
5. Grip	○◇	○◇	○◇	●◆	○◇	○◇	●◆	100%
6. Drawing above shoulder level	○◆	●◆	●◆	●◆	●◆	○◇	●◆	86%
7. Standing-up	●◆	○◇	●◆	●◆	○◇	○◇	●◆	100%
8. Lifting object waist to shoulder	●◆	●◆	●◆	●◆	○◇	●◆	●◆	100%
9. Lifting object waist to floor	●◆	●◆	●◆	●◆	○◇	●◆	●◆	100%
Total								96.6%

The figures ○◇ indicate no observed pain behaviour and ●◆ indicate observed pain behaviour among two observers.

Adv Physiother Downloaded from informahealthcare.com by 83.248.39.103 on 09/03/12 For personal use only.

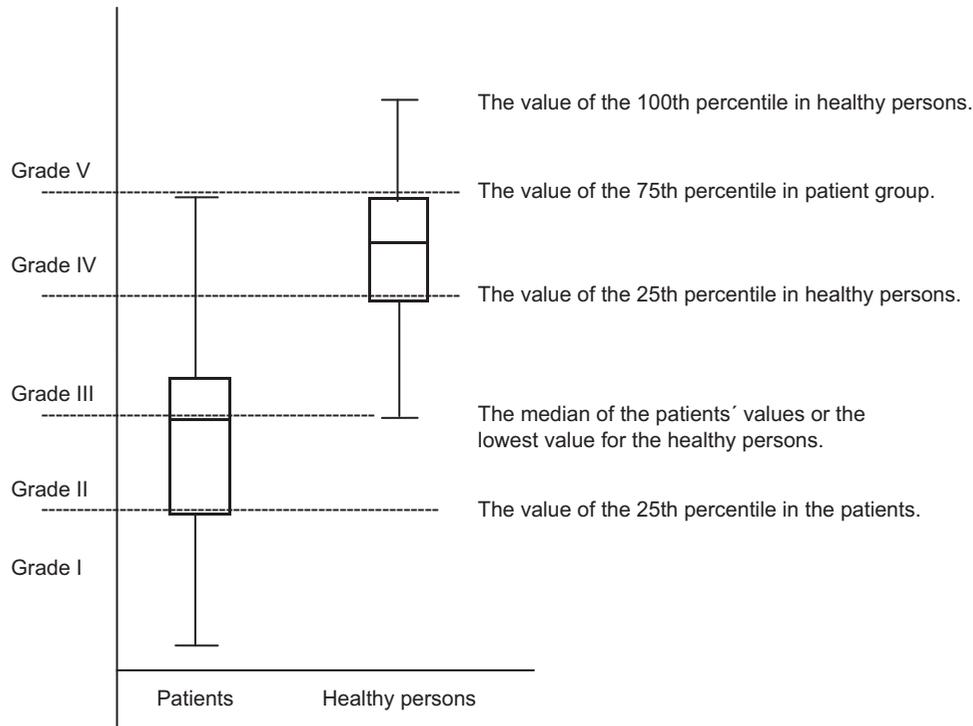


Figure 2. The grading system for each activity in variable B is based on both patients' and healthy persons' results and is different for females and males. This graphic shows how the grades are chosen.

given in Table V. The grading between the two occasions was in agreement for all activities except one, *climbing on and off a stool*, where the median changed from grade II on the first occasion to grade III on the second (Table V). There was no statistical difference in occasional disagreement, and the values for RV varied between 0.00 and 0.05 (Table VI). For systematic disagreement, there was a statistically significant difference for the activity *drawing above shoulder level left*, $RP -0.2012$ ($-0.3751; -0.0272$), demonstrating a 20% chance to perform lower at the second test (Table VI).

Discussion

The TIPPA instrument is primarily a test of activity level and includes everyday activities according to the International Classification of Functioning, Disability and Health (ICF) (24). In addition to assess physical ability, TIPPA includes the individual's self-rating of his or her ability, pain behaviour, impact on pain and experience of exertion, which gives a further dimension in the test. The TIPPA instrument is suitable especially in pain and work-oriented rehabilitation but is also used as a part of a more complex investigation of work ability. Besides being a valid and reliable instrument, TIPPA meets the requirements presented in "Guidelines for functional capacity evaluation of people with medical conditions",

concerning safety, practicality and utility (25). The physical activities included in TIPPA are everyday activities that involve the whole body, which addresses the demand for utility. The demand of practicality is fulfilled inasmuch as the test materials are inexpensive and simple to gather, and the patients do not need to change clothes. Safety aspects in TIPPA are that the patients themselves can choose weight and tempo, instruction is given for the lifting technique before the lift tests, and the standard position of the tests can be changed when necessary.

The variables in the instrument were chosen on the basis of the following:

Variable A. A discrepancy is often found between a patient's self-rating (variable A) and the observed physical ability (variable B), and this confirms that assessment of physical ability cannot be made only on the basis of questionnaires (16). The self-assessment can reflect the individual's self-efficacy, which has been shown to have significant association with self-report measures of pain intensity, pain disability and pain behaviours among LMP patients and as a predictor of persistent disability in patients with whiplash-associated disorders (26,27).

Variable B. In order to make the TIPPA test useful for patients irrespective of disability and pain

Table IVa. The value of the grades for nine physical activities in variable B of the Test Instrument for Profile of Physical Ability (TIPPA) and the frequency of the patients and healthy persons for each grades among women.

Outcome units Physical activities	Grade I P + H	Grade II P + H	Grade III P + H	Grade IV P + H	Grade V P + H	Total <i>n</i>
M	< 300	305–355	360–420	425–490	> 495	
^a Walking	P 130 H 0	110 1	147 7	77 16	11 7	475 31
Second Speed walking ^b	> 15 P 77 H 0	14–13 67 0	12–11 148 1	10–8 176 25	< 7 12 5	480 31
Number of steps Stair climbing	< 77 P 173 H 0	78–98 106 0	99–120 107 9	121–133 51 8	> 134 53 14	490 31
Number of repetitions Climbing up and off a stool	< 12 P 134 H 1	13–16 130 1	17–20 128 9	21–27 80 13	> 28 18 7	490 31
Newton Gripe Force	< 88 P 293 H 0	89–152 84 1	153–224 61 8	225–296 35 15	> 297 17 7	490 31
Second Drawing above shoulder level	< 45 P 126 H 0	46–121 243 1	122–186 75 7	187–300 43 16	> 301 3 7	490 31
Number of repetitions Standing-up	< 12 P 144 H 1	13–15 86 1	16–19 139 8	20–25 90 14	> 26 31 7	490 31
Kg×repetitions Lifting object waist to shoulder	< 16 P 138 H 0	17–40 192 1	41–65 90 8	66–106 56 16	> 107 13 6	489 31
Kg×repetitions Lifting object waist to floor	< 34 P 140 H 0	35–60 109 0	61–105 162 6	106–194 71 18	> 195 8 7	490 31

P, number of the patients; H, number of the healthy persons. ^aWalking distance is adjusted by 5. ^bSpeed walking has a lower grade when it takes a longer time.

localization, nine physical activities were selected for variable B (6,17,18). Seven of the nine activities are uncomplicated activities whereas lifting activities can be performed in different ways; the two lifting in TIPPA will thus be discussed. Lifting offers a possibility to analyse different body functions in activities in daily life and work (16). Two lifting activities in TIPPA have the character of a dynamic and psychophysical test, which implies that the patient chooses the load himself (28,29). The patient's choice of load is influenced by cognition, motivation and perceptual stimuli (30). In the TIPPA-grading system, consideration is taken to sex. This is based on the fact that men are on average taller, heavier and stronger than women. If the grading system had also been based on each individual's weight, height and age, it would have been too complicated to use and not lucid. The lifting, activity 9 in TIPPA, is done laterally from table to floor at 90° and requires that the feet are moved, as this is the type of lift that occurs most often in everyday life. This technique is more demanding compared with other similar tests, e.g. the PILE (Progressive Isoinertial Lifting Evaluation) (3,31). Lifting above

the waist, activity 8, is done with one hand at a time since a patient can have one-sided shoulder/arm pain that negatively affects lifting done with both arms (Appendix). TIPPA is not intended to measure the maximum lifting ability but rather the weight that the individual subjectively chooses in the situation. Judging from the literature and our own clinical experience, a majority of LMP patients stop the test when they feel uncomfortable (3).

Variable C. To gain information about how the patient has experienced carrying out the nine physical activities in terms of exertion, he or she is asked to estimate his or her degree of exertion. Grades I–V for this variable, i.e. “very great” to “very little”, correspond to the reversed interval 7–1 on Borg's CR 10 scale, which is a general scale for measuring the intensities of most kinds of sensory perceptions, experiences and feelings (32).

Variable D. The reason for modifying the UAB Pain Behaviour Scale was that there was no need in TIPPA to assess the two UAB Pain Behaviour Scale criteria of “Down-time” and “Medication”

Table IVb. The value of the grades for nine physical activities in variable B of the Test Instrument for Profile of Physical Ability (TIPPA) and the frequency of the patients and healthy persons for each grades among men.

Outcome units Physical activities	Grade I P+H	Grade II P+H	Grade III P+H	Grade IV P+H	Grade V P+H	Total n
M	< 300	305–355	360–430	435–500	> 505	
^a Walking	P 41 H 0	48 0	47 5	24 8	3 4	163 17
Second	> 14	13–12	11–10	9–8	< 7	
Speed walking ^b	P 36 H 0	23 0	63 4	35 10	10 3	167 17
Number of steps	< 77	78–103	104–122	123–153	> 154	
Stair climbing	P 53 H 0	37 2	44 4	29 7	5 4	168 17
Number of repetitions	< 12	13–16	17–20	21–27	> 28	
Climbing up and off a stool	P 54 H 0	37 0	49 3	21 12	7 2	168 17
Newton	< 150	151–280	281–380	381–490	> 491	
Gripe Force	P 112 H 0	17 0	19 5	10 7	10 5	168 17
Second	< 45	46–132	133–240	241–383	> 384	
Drawing above shoulder level	P 44 H 0	70 0	31 6	18 3	5 8	168 17
Number of repetitions	< 12	13–15	16–20	21–25	> 26	
Standing-up	P 48 H 0	45 0	42 6	26 7	7 4	168 17
Kg×repetitions	< 20	20–60	61–110	111–190	> 191	
Lifting object waist to shoulder	P 42 H 1	62 0	44 4	16 8	4 4	168 17
Kg×repetitions	< 35	36–80	81–170	171–310	> 311	
Lifting object waist to floor	P 47 H 0	34 0	58 5	25 7	4 5	168 17

P, number of the patients; H, number of the healthy persons. ^aWalking distance is adjusted by 5. ^bSpeed walking has a lower grade when it takes a longer time.

(19). Instead, two criteria were included that have to do with performance, i.e. “interruption or total avoidance of an activity” and “changing the standard position for the test” (Table I). Another difference between UAB Pain Behaviour Scale and TIPPA is the scoring. In UAB, point is given for different pain

behaviour regardless of the type of activity, unlike TIPPA, where observation of pain behaviour is always related to one or more of the nine activities. It is therefore possible to read off whether the pain behaviour is related to one or more parts of the body or whether it is general.

Table V. The median values and their grades for the nine physical activities at two test occasions.

Physical activities	1st test occasion				2nd test occasion			
	Min	Max	Median	Grade	Min	Max	Median	Grade
Walking, m	315	445	390	III	300	460	405	III
Speed walking, s	16 ^a	7	11	III	16 ^a	7	10	III
Stair climbing, steps	72	177	117	III	74	152	114	III
Climbing on and off stool, numbers	1	28	16	II	0	27	18	III
Grip force, right, N	42	391	187	II	35	296	115	II
Grip force, left, N	62	374	136	II	44	348	176	II
Drawing above shoulder level, right, s	8	482	118	II	7	482	65	II
Drawing above shoulder level, left, s	16	482	66	II	15	482	46	II
Standing-up, numbers	10	22	17	III	11	24	16	III
Lifting object waist to shoulder, right kg×repetitions	0	240	48	II	0	245	54	II
Lifting object waist to shoulder, left kg×repetitions	0	184	42	II	0	225	45	II
Lifting object waist to floor, kg×repetition.	10	250	78	III	27	274	78	III

^aSpeed walking has a lower grade when it takes a longer time.

Table VI. Values of occasional and systematic disagreement in the test–retest reliability study ($n = 13$).

Items	Systematic disagreement			
	RV		RP	
Physical activities	CI		CI	
Walking	0.0000	0.0000 to 0.0000	0.2012	-0.0072 to 0.4096
Speed walking	0.0210	0.0210 to 0.0210	0.0222	-0.2432 to 0.2875
Stair climbing	0.0055	0.0000 to 0.0200	-0.0178	-0.0928 to 0.1279
Climbing on and off a stool	0.0164	0.0000 to 0.0569	0.0473	-0.1180 to 0.2127
Grip force, right	0.0055	0.0000 to 0.0200	-0.1538	-0.3168 to 0.0091
Grip force, left	0.0546	0.0000 to 0.1725	0.0828	-0.1374 to 0.3031
Drawing above shoulder level, right	0.0382	0.0000 to 0.1111	-0.1065	-0.3096 to 0.0966
Drawing above shoulder level, left	0.0000	0.0000 to 0.0000	-0.2012	-0.3751 to -0.0272
Standing-up	0.0328	0.0000 to 0.0812	-0.0947	-0.3005 to 0.1112
Lifting object waist to shoulder, right	0.0000	0.0000 to 0.0000	-0.0059	-0.1369 to 0.1250
Lifting object waist to shoulder, left	0.0000	0.0000 to 0.0000	0.0355	-0.0980 to 0.1690
Lifting object waist to floor	0.0000	0.0000 to 0.0000	0.0000	0.0000 to 0.0000

RV, relative rank variance; RP, relative position; CI, confidence interval. Possible RV ranges 0 to 1 and RP ranges -1 to 1.

The agreement in observations of pain behaviour in the LMP patients between the two PTs was found to be high. The high agreement (96.6%) despite the small sample size speaks for well defined observation criteria. This can be compared with the Swedish version of UAB Pain behaviour Scale, where the agreement was 80% (20).

Grading system

The grading system for physical activities in TIPPA is built on test results from groups of healthy persons and patients. The previously constructed grading system showed significant differences between the patient group and the healthy group for both categories of men and women (17). The system showed weaknesses in some of the activities, however, such as walking and lifting above the waist, where it was difficult to discriminate between the scores. The reconstruction of the scores in this study is built on a large number of patients with LMP, both women and men ($n = 658$) with varying localizations of pain. The number of persons in the healthy reference group was unchanged ($n = 48$). The unchanged reference group can be motivated since there were limits for the patients' grades that needed to be adjusted. When the individuals' values in each of the nine activities (variable B) were distributed according to the 25th, 50th and 75th percentiles, the majority of patients were found in grades I–III and the majority of the healthy persons in grades IV–V (Tables IVa and IVb). The grading system was reconstructed in line with what we had hypothesized based on the previously construction and our clinical experience. One of the activities in men and two in women differed slightly from this model, according to experiences from clinical practice. Others who also

used percentile distribution are Soer et al. (3), who defined normative values for evaluation of functional capacity according to 1–99th percentiles among healthy persons, 20–60 years old, men and women. Here, the grades were classified into four categories of workload following the physical job demands used in the Dictionary of Occupational Titles (DOT) as sedentary, light, medium and heavy/very heavy (3,33). TIPPA's scores are not related to any specific occupational group or physical job demands but, if a patient has the same grades as healthy persons, i.e. grades III–V, this is interpreted such that physical conditions exist for some form of physical work. The broad range of different pain diagnoses and conditions makes the grading system very adequate for the LMP patient group.

The results of test–retest reliability of the reconstructed grades I–V in nine physical activities (variable B) showed very good RP and RV values, indicating good test–retest reliability. Only one significant systematic change was found in the activity of *drawing above shoulder level left*. It is difficult to explain why this was shown for this activity and it may simply be a result of chance. The activity of *climbing on and off a stool* changed in grading, from II on the first occasion to III on the second, based on the median values. There may perhaps be a training effect in this activity, which is not an ordinary daily activity. The results of the test–retest study should be interpreted in light of its limitations. A larger sample size would have strengthened the test of reliability for the nine activities in TIPPA.

Clinical considerations

TIPPA is intended for use in evaluations of patients' physical ability and can as such be a part of an

overall assessment of work ability that is more complex and multifaceted. TIPPA's standardized grading system defines concepts such as "low", "medium", "high" and so forth (Figure 1). The grading system makes it possible to formulate an individual profile. The profile makes it easy to see the relation between the patient's self-assessment, the observed physical ability and possible pain behaviour. The profile is easy to use, to provide feedback to the patient/individual, the doctor, the members of the team and other stakeholders. The profile of physical ability makes it easy to compare test results before and after rehabilitation. If the patient's self-rating of his or her physical ability (variable A) and/or experience of the weight of the activities (variable C) differs substantially from the observed physical ability (variable B), it is the responsibility of the examiner to try to clarify the reason for the discrepancy. In the same way, it is important to judge whether the pain behaviour shown (variable D) is an adequate expression of pain and/or functional impairment or whether it is an expression of suffering. Clinical experience of using TIPPA by physiotherapists in Sweden confirms that this instrument discriminates between patients and healthy persons.

The TIPPA profile, like FCE, is not a solitary evaluation. The profile should be interpreted in its context against the background of medical history, interview, physical examination, observations during the conduct of the test, estimation of pain by the visual analogue scale (VAS) and assessment of pain behaviour to give finally an evaluation of the patient's functional ability in interdisciplinary cooperation. As a part of the judgement of work ability, the demands of the job must also be related to the patient's capacity (1,16). Assurance of a reliable assessment requires trained and experienced examiners (16). It is therefore suggested that users of the TIPPA be trained in the procedure.

Based on the results of this study, the TIPPA is found to be valid and reliable enough to be used in rehabilitation settings and as a part of an overall assessment of work ability in patients with LMP. Validation of TIPPA's variables A–D related to other instruments is recommended for further study designs. The grades III–V for healthy persons are based on the test results of a limited group of healthy persons as compared with the patient group. A larger number of tests in healthy persons of different ages and sexes would give a more certain distribution for these grades of grading systems.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. King PM, Tuckwell N, Barrett TE. A critical review of functional capacity evaluations. *Phys Ther.* 1998;78:852-66.
2. Fishbain, DA. Functional Capacity Evaluation. *Phys Ther.* 2000;80:110-2.
3. Soer R. Functional Capacity Evaluation; Measurement qualities and normative values. Chap. 9. Thesis, University of Groningen; 2009. Available from: <http://irs.ub.rug.nl/ppn/318136333>
4. Genovese E, Galper J, editors. Guide to the evaluation of functional ability; How to request, interpret and apply Functional Capacity Evaluations. Chicago, IL: American Medical Association; 2009. DOI: 10.1007/s10926-010-9279-8. [cited 2011 Dec 2]; available from <http://wdn.ipublishcentral.net/impelsys549/viewinside/115601268817107>
5. Lakke SE, Soer R, Takken T, Reneman M. Risk and prognostic factors for non-specific musculoskeletal pain: A synthesis of evidence from systematic reviews classified into ICF dimensions. *Pain.* 2009;147:153-64.
6. Harding VR, Williams AC, Richardson PH, Nicolas MK, Jackson JL, Richardson IR, et al. The development of a battery of measures for assessing physical functioning of chronic pain patients. *Pain.* 1994;58:367-75.
7. Ijzelenberg W, Burdorf A. Risk factors for musculoskeletal symptoms and ensuing health care use and sick leave. *Spine.* 2005;30:1550-6.
8. Arnstein P. The medication of disability by self efficacy in different samples of chronic pain patients. *Disabil Rehabil.* 2000;22:794-801.
9. Busch H, Göransson S, Melin B. Self-efficacy beliefs predict sustained long-term sick-absenteeism in individuals with chronic musculoskeletal pain. *Pain Pract.* 2007;7:234-40.
10. Gatchel RJ, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: Scientific advances and future directions. *Psychol Bull.* 2007;133:581-624.
11. Linton SJ. The relationship between activity and chronic back pain. *Pain.* 1985;21:289-94.
12. Fordyce WE. Behavioral factors in pain. *Neurosurg Clin N Am*1991;2:749-59.
13. Turk DC, Wack JT, Kems RD. An empirical examination of the "pain-behavior" construct. *J Behav Med.* 1985;8:119-30.
14. Deyo RA. Measuring the functional status of patients with low back pain. *Arch Phys Med Rehabil.* 1988;69:1044-53.
15. Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. *Spine.* 1998;23:2412-21.
16. Oliveri M. Functional capacity evaluation (FCE). In: Gobelet C, Franchignoni F, editors. Vocational rehabilitation. Series of European Academy of Rehabilitation Medicine. Berlin: Springer-Verlag; 2006. p. 69-93. Available from <http://springerling.com/content/978-2-287-22609-0#section=406065&page=1&locul=0>
17. Missaghi-Wedefalk M, Lindh M, Nordholm L, Schön-Ohlsson C. Test Instrument for Profile of Physical Ability = TIPPA – ett testbatteri för kartläggning och bedömning av fysisk förmåga avsett för patienter med långvarig muskuloskeletal värk. *Nordisk Fysioterapi.* 2004;8:74-83. In Swedish. Abstract in English. Available from <http://www.tippa-info.se>
18. Nordenskiöld UM, Grimby G. Grip force in patients with rheumatoid arthritis and fibromyalgia and in healthy subjects: A study with the Grippit instrument. *Scand J Rheumatol.* 1993;22:14-19.
19. Richards Js, Nepomuceno C, Riles M, Suer Z. Assessing pain behavior: The UAB pain behavior scale. *Pain.* 1982;14:393-8.

20. Dottori M, Eriksson M. Interrater reliability test of the University of Alabama at Birmingham (UAB) Pain Behavior Scale: A pilot study. Research project in physiotherapy. 10 credits within Physiotherapy programme; Autumn 2001.
21. Novy DM, Simonds MJ, Olson SL, Lee CE, Jones SC. Physical performance: Differences in men and women with and without low back pain. *Arch Phys Med Rehabil.* 1999; 80:195–8.
22. Svensson E. Ordinal invariant measures for individual and group changes in ordered categorical data. *Stat Med.* 1998;17:2923–36.
23. Svensson E. Analysis of systematic and random differences between paired ordinal categorical data. Dissertation, University of Gothenburg, Almqvist & Wiksell International; 1993.
24. WHO. ICF: International Classification of Functioning, Disability and Health. Geneva: World Health Organization; 2001. Available from <http://www3.who.int/icf/>
25. Hart DL, Isernhagen SJ, Matheson LN. Guideline for functional capacity evaluation of people with medical conditions. *J Orthop Sports Phys Ther.* 1993;18:682–6.
26. McCahon, S, Strong J, Sharry R, Cramond T. Self-report and pain behavior among patients with chronic pain. *Clin J Pain.* 2005;21:223–31.
27. Bunkertorp L, Lindh M, Carlsson J, Stener-Victorin E. The perception of pain and pain-related cognitions in subacute whiplash-associated disorders: Its influence on prolonged disability. *Disabil Rehab.* 2006;28:271–9.
28. Griffin AB, Tourp JD. Test of lifting and handling capacity: Their repeatability and relationship to back symptoms. *Ergonomics.* 1984;7:305–20.
29. Troup JDG, Foreman TK, Baxter CE, Brown D. The perception of back pain and the role of psychophysical tests of lifting capacity. *Spine.* 1987;12:645–57.
30. Gamberale F. Perception of effort in manual materials handling. *Scand J Work Environment Health.* 1990;16 Suppl 1:59–66.
31. Mayer TG, Barnes D, Kishino ND, Nicholas G, Gatchel RJ, Meyer H, et al. Progressive isoinertial lifting evaluation. I: A standardized protocol and normative database. *Spine.* 1988;13:993–7. Erratum in *Spine.* 1990;15:5.
32. Borg G. Psychophysical scaling with applications in physical work and perception of exertion. *Scand J Work Environ Health.* 1990;16 Suppl 1:55–8.
33. U.S. Department of Labor, Employment and Training Administration. Dictionary of Occupational Titles. 4 ed.; 1991.

Supplementary material available online

Supplementary Appendix to be found online at <http://www.informahealthcare.com/doi/abs/10.3109/14038196.2012.694907>