

ORIGINAL REPORT

CONSTRUCT DIMENSIONALITY AND PROPERTIES OF THE CATEGORIES IN THE ICF CORE SET FOR LOW BACK PAIN

Cecilie Røe, MD, PhD^{1,2}, Unni Sveen, PhD¹, Szilvia Geyh, PhD³, Alarcos Cieza, PhD⁴ and Erik Bautz-Holter, MD, PhD^{1,2}

From the ¹Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Ullevål, ²Faculty of Medicine, University of Oslo, Oslo, Norway, ³Swiss Paraplegic Research, Nottwil, Switzerland and ⁴ICF Research Branch of the WHO Collaborating Center for the Family of International Classifications at the German Institute of Medical Documentation and Information (DIMDI), IHRS, Ludwig-Maximilian University, Munich, Germany

Objective: The aim of this study was to explore by Rasch analysis whether the Comprehensive International Classification of Functioning, Disability and Health (ICF) Core Set might represent a future clinical tool for measuring functioning of patients with low back pain.

Material and methods: The Comprehensive ICF Core Set for low back pain was scored by health professionals for 118 patients with low back pain. Qualifier levels, invariance, construct validity and ordering of the categories in the components of Body function, Body structure, Activities and participation and Environmental factors were explored by Rasch analysis.

Results: The number of qualifier levels had to be reduced. Categories within Body functions and within Environmental factors reflected a single underlying construct. The categories within the component of Activities and Participation did not meet the requirements of a single underlying construct in the present population. Few categories covered the problems reported by patients with a relatively high level of function.

Conclusion: Rasch analysis indicated that the Comprehensive ICF Core Set for low back pain may be used with some modification of categories as a common tool for assessing problems within the components Body functions, and Activity and Participation. However, detecting ICF categories that reflect the higher functional levels in patients with low back pain, and revision of the qualifier levels may be necessary.

Key words: ICF, Core Set, lumbosacral, health, rehabilitation.

J Rehabil Med 2009; 41: 429–437

Correspondence address: Cecilie Røe, Department of Physical Medicine and Rehabilitation, Oslo University Hospital Ullevål, NO-0407 Oslo, Norway. E-mail: eller@medisin.uio.no

Submitted May 28, 2008; accepted March 17, 2009

INTRODUCTION

The International Classification of Functioning Disability and Health (ICF) (1) has been developed in order to broaden the perspective on patients' functioning, including activities, participation, environmental and personal factors. To enhance feasibility, the ICF Core Set project was initiated in 2001 (2, 3). The goal was to establish a selection of categories tailored

to represent standards for multi-professional assessment of specific patient groups. The development of the Comprehensive ICF Core Set for low back pain (LBP) included a formal consensus process (4), as well as empirical data collections (5, 6). The concepts contained in the established outcome measurements for LBP were also encompassed (7). The proposed Comprehensive ICF Core Set for LBP comprises 78 categories within the components Body functions, Body structures, Activities and Participation, and Environmental factors (4). The huge reduction in categories in the ICF Core Set as opposed to the whole ICF is a prerequisite regarding clinical use. The Comprehensive ICF Core Set for LBP is now undergoing worldwide testing to determine the extent to which the problems of patients with LBP are captured by the Core Set, and whether there are missing or redundant categories (4).

The ICF Core Set for LBP was developed as a classification tool. In order to achieve interval-scaled measurement from the sum of categories in the different ICF components, the requirement is that they satisfy Rasch model and reflect a unidimensional underlying construct (8). The items are termed categories according to ICF terminology. Applying the data to the Rasch model can also test other important psychometric properties of the ICF Core Set for LBP, including the ordering of the qualifiers used to score the degree of problems in the ICF categories (9, 10). Furthermore, invariance of the Core Set across factors such as age, gender and education ensures the comparability of results across different settings and populations (11), and may be tested. In addition, it is important in clinical practice and research to be able to assess different levels of problems. In this context, Rasch analysis is also useful to study the targeting of the components and the hierarchical ordering of both patients and categories. This provides information about whether the categories match the different levels of problems presented by the actual patient population, and can, along with local dependency among the categories, provide information about possible redundant categories (12).

Hence, if the categories in the Comprehensive ICF Core Set reflect unidimensional constructs of Body functions, Activities and Participation as well as Environmental factors, if they are invariant across populations, well-targeted and non-redundant, then they can represent a future clinical tool for measuring functioning of patients with LBP.

Thus, the aim of this study was to evaluate: (i) whether the qualifier levels of the categories within the components are adequate; (ii) whether the response pattern is uniform across age, gender and levels of education; (iii) the unidimensionality of the components of Body Functions, Activities and Participation, and Environmental factors; (iv) the ordering and difficulty of the categories as well as the targeting of the components and redundancy of categories.

MATERIAL AND METHODS

Study design

The design was a multi-centre cross-sectional study involving 4 study centres. This Norwegian multi-centre study was part of an international validation study coordinated by the ICF Research Branch at the Ludwig-Maximilian University in Munich, Germany.

Data were collected from November 2005 and throughout December 2006 by physicians, physiotherapists and nurses at the participating study centres. The health professionals were trained in the study procedures by researchers at the Ullevål University Hospital, Oslo, Norway. The study was approved by the Regional Committee for Medical Research Ethics in Health Region East.

Inclusion criteria

The study was performed with samples of patients with LBP treated in the participating study centres. Patients were included if they were at least 18 years old and had LBP as their main health problem. They should have sufficient knowledge of the Norwegian language, the purpose of and reason for the study had to be understood, and informed consent was signed.

Measures

The health professionals documented the International Classification of Diseases-10 diagnosis, the number of surgical interventions and hospitalizations due to LBP, as well as whether the pain was acute (<4 weeks), subacute (4–12 weeks) or chronic (>12 weeks) and the employment status of the subjects.

The Comprehensive ICF Core Set for LBP was scored by the health professionals based on all available clinical information. The Comprehensive ICF Core Set consists of 19 (24%) categories from the component Body functions, 5 (6%) from the component Body structures, 29 (37%) from the component Activities and Participation, and 25 (33%) from the component Environmental Factors. To evaluate the extent of the patient's problem in each of the ICF categories contained in the Comprehensive ICF Core Set for LBP, the qualifier scale was used. The qualifier scale of the components Body functions, Body structures and Activities and Participation has 5 response levels, ranging from 0 to 4: no/mild/moderate/severe/complete impairment or difficulty. The qualifier scale of the component Environmental factors has 9 response levels, ranging from -4 to +4. A specific environmental factor can be a barrier (-1 to -4), a facilitator (1–4), or can have no influence (0) on the patient's functioning. If the factor has an influence, the extent of the influence (either positive or negative) can be coded as mild/moderate/severe/complete. In addition, the qualifiers "8 – not specified" and "9 – not applicable" can be used.

In addition to the data collected by the health professionals, the patients also filled in the Oswestry Disability Index (ODI) (13). ODI assesses pain and problems with functioning related to LBP in 10 items scored on a 6-point Likert scale. The items represent pain intensity, personal activities of daily living, lifting, walking, sitting, standing, sleeping, sexual activity, social activity and travelling. A validated Norwegian version was used (14).

Data analysis and statistics

The qualifier "not applicable" (9) was deemed not to represent a problem and was coded as 0 (no problem). The qualifier not "specified"

(8) was deleted from the database and considered a missing value. The internal construct validity of the components of Body functions, Activities and Participation, and Environmental factors was explored by Rasch analysis for polytomous categories (9).

The partial credit model was applied, as this model is valid without the assumption of equidistance between thresholds across items (ICF categories) (15). The qualifier ordering for each category was examined, and qualifiers with overlapping thresholds collapsed.

Whether the response pattern is uniform across age, gender and levels of education was examined by differential item functioning analyses (DIF). DIF is assessed by analysis of variance for each category, comparing scores across each level of age, gender and education (11). Significant main effects of age, gender and education (uniform DIF), and interaction (non-uniform DIF) between age, gender and education and subgroups of the patients (class intervals according to the level of functioning) were both evaluated. For these analyses, age was dichotomized, below and above the median age of 45 years. Education was divided into below and above 12 years, according to the cut-off for higher education in Norway. The F ratio (F) for the group difference and probability (*p*) were given for the DIF analysis.

Local dependency in each component was evaluated by correlation of the residuals of the categories, with a coefficient of 0.3 as threshold value.

To study the unidimensionality of the components of Body function, Activities and Participation, and Environmental factors, χ^2 item trait interaction statistics were applied (16). A non-significant probability value is an indication of a unidimensional underlying construct. An additional verification of unidimensionality was undertaken by creating 2 subsets of categories, representing the categories with the most positive and most negative residuals according to a principal component analysis. Person estimates for each of the 2 subsets were calculated, and independent *t*-tests comparing the 2 estimates in each person were performed (17). Similar estimates indicate unidimensionality of the underlying construct. The number of *t*-tests with *p*-values below 0.05 and the corresponding 95% confidence interval (CI) were reported. The recommendation for a unidimensional construct is that CI should include 0.05.

The difficulty of the ICF categories was evaluated by examining the hierarchical distribution of the categories and their qualifiers within each component. The Person Separation Reliability Index is reported, providing an indication of the power of the measure to discriminate among persons with different levels of the trait. A value above 0.8 was deemed to differentiate across at least 3 patient groups (16). The fit of individual persons and categories were reported as a mean logit with standard deviation (SD), a mean logit of 0 and a SD of 1 representing an optimal fit of the categories. The fit of the categories was statistically evaluated by standardized residuals and χ^2 statistics. Item residuals ± 2.5 and a non-significant χ^2 probability value were considered to indicate adequate fit to the Rasch Model (16).

Table I. Categories with distorted thresholds and the subsequent combination of qualifier levels (×: combined qualifiers)

ICF categories	Combining qualifier 0 and 1	Combining qualifier 1 and 2
b134 Sleep functions		×
b180 Experience of self and time functions		×
b280 Sensation of pain	×	
b620 Urinating functions		×
b640 Sexual functions		×
b720 Mobility of bone functions		×
b750 Motor reflex functions		×
d570 Looking after one's health		×
d760 Family relationship		×
d770 Intimate relationship		×
d859 Work and employment		×

ICF: International Classification of Functioning, Disability and Health.

Table II. Fit of the Components in the Comprehensive ICF Core Set for low back pain, based on Rasch analysis ($n = 118$). The categories in the component activity and participation had to be divided in 2 sets in order to meet the requirements of unidimensionality

ICF component	χ^2 item trait interaction statistics	p -value	Person fit Mean (SD)	Person Separation Reliability Index	Item fit Mean (SD)
Body functions	47.16	0.10	-0.21 (0.88)	0.74	-0.08 (0.88)
Body structure	28.73	0.001			
Activity and Participation I	46.46	0.08	-0.22 (0.88)	0.90	-0.19 (1.15)
Activity and Participation II	20.59	0.30	-0.40 (0.9)	0.85	-0.57 (0.78)
Environmental	44.30	0.05	-1.11 (2.11)	0.85	-0.08 (0.88)

ICF: International Classification of Functioning, Disability and Health; SD: standard deviation.

The Rasch analysis was performed in RUMM 2020 (RUMM laboratory, Perth, Australia). Other analyses were performed by SPSS for windows version 13.0. A significance level of 0.05 was adopted. This significance level was Bonferroni corrected according to the number of categories and groups tested (18).

RESULTS

All 118 participants were Norwegian residents, mean age 47 (SD 12) years, and 52% ($n = 63$) were women. The mean duration of LBP was 41 months (SD 85 months). Six percent had radiculopathy (mean 54.1), 4% sciatica (mean 54.3), 18% lumbago with sciatica (mean 54.4), 64% lumbago (mean 54.5) and 8% spinal stenosis (mean 48.0). Twenty-three percent had had at least one surgical intervention due to LBP and 62% had been hospitalized once due to LBP. Sixty-two subjects were employed (56%); however, 43 of these subjects were on sick leave. Twenty percent of patients were unemployed, the majority for health reasons. Eleven percent had retired and 13% were pensioned due to LBP. The mean ODI score (13, 14) was 33% (SD 14%).

Body functions and Body structures

For several of the categories of Body functions and Body structures, the qualifier levels 3 and 4 were not used. Thus, these qualifiers were combined with qualifier level 2, leaving

3 levels of qualifiers for the analysis. The qualifier levels of some of the categories were not distinct and had to be combined (Table I). The combination of qualifiers 1 and 2 was generally the best solution, with the exception of *b280 Sensation of pain*, where combining the qualifier level 0 and 1 seemed to be the best solution judged by the category probability curves. Body structures were omitted in the further analysis, as this component did not meet the assumptions in the Rasch model (Table II).

Invariance across age, gender and level of education was found for all categories in the component of Body functions. The correlation matrix of the residuals in the component of Body functions revealed one correlation exceeding the preset level of +0.3, indicating local dependency between *b180 Experience of self and time functions* and *b630 Urinary functions* ($r = 0.44$).

The Person Separation Reliability Index was 0.74 (Table II), indicating that 2 groups of subjects could be separated for body functions. The categories had a logit distribution of -3.23 to 2.35 (Table III), where 1 logit difference between categories represents a difference in odds of 2.7 for a person to report problems between the categories. The mean location for the persons was -0.95 (SD 0.93), indicating that the present patients had less problems with body functions than was reflected by the selection of categories (Fig. 1). Positive location of the categories

Table III. Fit of the categories in the component Body functions of the Comprehensive Core Set for low back pain, based on Rasch analysis ($n = 118$)

ICF code	ICF category title	Location	SE	Residual	χ^2	p -value
b180	Experience of self and time functions	2.35	0.43	-1.14	2.66	0.26
b260	Proprioceptive function	1.66	0.24	-0.94	1.93	0.38
b620	Urination functions	1.29	0.29	1.19	6.70	0.04
b640	Sexual functions	1.12	0.28	-0.38	0.46	0.80
b780	Sensation related to muscle and movement	0.80	0.18	-1.39	0.51	0.77
b715	Stability of joint functions	0.74	0.18	0.95	0.55	0.76
b720	Mobility of bone functions	0.44	0.23	-0.33	0.65	0.72
b126	Temperament and personality functions	0.01	0.14	0	1.63	0.44
b750	Motor reflex functions	0.01	0.22	0.56	1.30	0.52
b152	Emotional functions	-0.13	0.14	-0.37	0.83	0.66
b735	Muscle tone functions	-0.14	0.14	0.47	1.66	0.44
b770	Gait pattern functions	-0.19	0.15	-1.26	2.55	0.28
b730	Muscle power function	-0.21	0.15	-1.10	6.70	0.04
b710	Mobility of joint functions	-0.63	0.13	0.90	0.38	0.83
b740	Muscle endurance functions	-0.90	0.14	0.73	2.44	0.30
b455	Exercise tolerance function	-1.30	0.13	-0.85	2.62	0.27
b134	Sleep functions	-1.72	0.21	0.76	5.11	0.08
b280	Sensation of pain	-3.23	0.30	0.77	8.48	0.01

ICF: International Classification of Functioning, Disability and Health; SE: standard error.

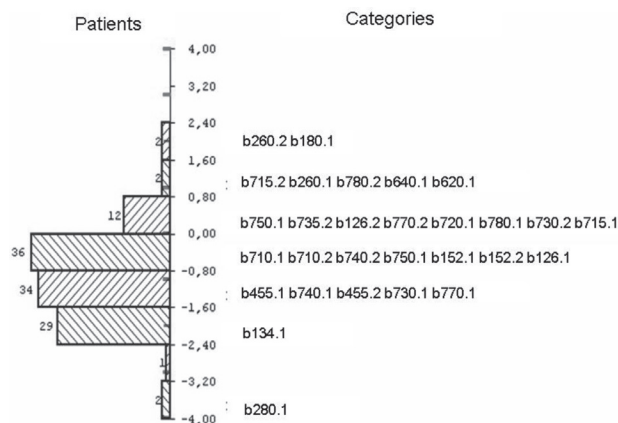


Fig. 1. Distribution of categories and patients ($n=118$) along the Rasch-calibrated metric scale of the component Body function. The right-hand panel shows the location of the categories with the threshold between no and mild problems (.1) and between mild and the more severe problems (.2), with exception of the categories with combined qualifiers (Table VI), which have only 1 threshold (.1). The left-hand panel shows the distribution of patients.

indicates that they are responded to by patients with lower level of function/higher impairment than the patients responding to categories with negative location. The lowest qualifier level of the category *b280 Sensation of pain* was easiest to approve and marked a high level of functioning, and was endorsed by all subjects (Fig. 1). There was a gap in the lower range of the logit scale, between *b280 Sensation of pain*, and *b134 Sleep functions*. The highest qualifier levels of the categories *b180 Experience of self and time functions* and *b260 Proprioceptive functions* (their highest qualifier levels) represented the most challenging functions. In this upper part of the logit scale there were no gaps between the categories and they seemed better matched to

patients' problems at this end of the scale (Fig. 1). None of the subjects were misfits or extreme in this component.

The component Body functions reflected a unidimensional construct when the category *b130 Energy and drive functions* was excluded. Also, the additional independent *t*-test supported unidimensionality with 5.9% of the tests outside the range of ± 1.96 , and the 95% CI for this probability was 0.04–0.08.

Activities and Participation

Also in this component 3 levels of the qualifiers were used for analyses. The qualifier levels of 4 categories were not distinct and had to be combined (Table I). The categories in the components of Activities and Participation, except *d630 Preparing meals*, responded in the same way across age, gender and level of education; hence no uniform DIF was found. No interactions between these factors and the class intervals were found, also indicating an absence of non-uniform DIF. However, *d630 Preparing meals* showed an interaction between age and class interval, reflecting that in the class interval with subjects having the greatest problems with activities, the subjects aged less than 45 years reported considerably greater problems than the older subjects.

In the Activities and Participation sets the category *d455 Moving around* was positively correlated with *d430 Lifting and carrying objects* and *d410 Changing body positions* ($r > 0.32$), indicating local dependency.

For the component of Activities and Participation the Person Separation Reliability Index was 0.85 and 0.90 (Table II) for the 2 sets emerging in this component (see below). Thus, for both sets 3 groups of subjects could be separated. The categories in set 1 had a logit distribution of -3.55 to 2.70 (Table III). The mean location for the persons was -1.26 (SD 1.65), indicating that the present patients had fewer problems than reflected by the categories in this set. The distribution of categories in this set is shown in Fig. 2A.

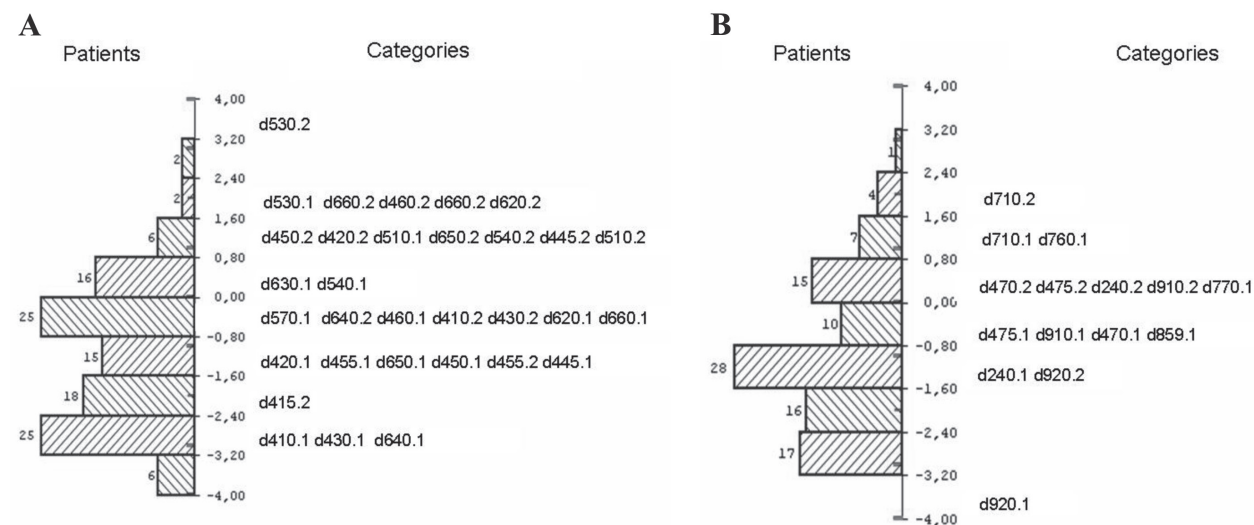


Fig. 2. Distribution of categories and patients in the Activity set 1 (A; $n=118$) and set 2 (B; $n=118$) of the component Activities and Participation, along the Rasch-calibrated metric scale. The right-hand panel shows the location of the categories with the threshold between no and mild problems (.1) and between mild and the more severe problems (.2), with exception of the categories with combined qualifiers (Table VI), which only have one threshold (.1). The left-hand panel shows the distribution of patients.

Table IV. Fit of the categories in set one of the component Activities and Participation of the Comprehensive Core Set for low back pain. Rasch analysis ($n=118$)

ICF code	ICF category title	Location	SE	Residual	χ^2	p -value
d530	Toileting	2.70	0.31	-1.45	4.83	0.09
d630	Preparing meals	2.34	0.24	-1.80	7.65	0.02
d510	Washing oneself	1.21	0.21	-0.57	1.17	0.56
d620	Acquisition of goods and services	1.15	0.21	-0.96	2.05	0.36
d660	Assisting others	0.93	0.20	0.50	0.28	0.87
d540	Dressing	0.80	0.19	0.51	4.26	0.12
d460	Moving around in different locations	0.79	0.19	-0.96	2.28	0.32
d445	Hand and arm use	0.21	0.18	1.47	1.42	0.49
d650	Caring for household objects	0.01	0.18	-1.65	3.93	0.14
d450	Walking	-0.04	0.17	0.76	0.28	0.87
d420	Transferring oneself	-0.22	0.18	-1.89	6.16	0.05
d570	Looking after one's health	-0.76	0.23	0.13	0.60	0.74
d455	Moving around	-1.10	0.15	-0.32	0.18	0.91
d430	Lifting and carrying objects	-1.44	0.17	1.77	0.88	0.24
d640	Doing housework	-1.45	0.16	0.17	2.60	0.27
d410	Changing basic body position	-1.58	0.17	-0.04	2.33	0.31
d415	Maintaining a body position	-3.55	0.21	1.14	3.58	0.17

ICF: International Classification of Functioning, Disability and Health; SE: standard error.

The categories in set 2 had a logit distribution from -2.14 to 1.49 (see Table V). The mean location for the persons was -1.56 (SD 1.73), again indicating that the present patients had fewer problems than reflected by the categories in this set. The distribution of the individual categories can be seen in Fig. 2B. The categories had a narrow logit range and did not match with the distribution of the patients. A total of 23 subjects were extremes and located at the very lower end of the logit scale, with values of -5.73 regarding Activities and Participation. The interpretation of this result is that these subjects have small problems with Activities and Participation, at least regarding the categories included in the present ICF Core Set.

The categories in the component of Activities and Participation did not meet the requirements of a unidimensional construct. Based on a principal component analysis, the categories with the highest and lowest residuals were grouped into separate sets. New analyses of the 2 subsets were performed. Seventeen categories representing daily activities could be fitted to the Rasch model (Table IV). However, the p -value was rather low, and the independent t -test showed that 10% of the tests were outside the range of ± 1.96 , (95% CI 0.060–0.139), indicating that the categories were close to, but not completely fulfilling the requirements for reflecting an underlying unidimensional

construct. Another 9 categories (Table V) could be fitted to the Rasch model, with 4% of the independent t -tests outside the range of ± 1.96 (95% CI 0.000–0.079), supporting an underlying unidimensional construct.

Environmental factors

For the Environmental factors, barriers were re-scored to 0, neither barrier nor facilitator was scored to 1 and facilitators scored 2. For the categories in the Environmental component, the thresholds for the qualifiers distinguished the levels for all 15 categories. No DIF nor any local dependency was found among the environmental factors. The Environmental categories had a logit distribution from -0.80 to 0.73 (Table VI). The mean location for the persons was 0.006 (SD 1.44), indicating that barriers and facilitators outweighed each other in the present population (Fig. 3). The hierarchy of barrier and facilitating categories are presented in Fig. 3, showing that *e550 Legal services* and *e155 Design and construction*, for both private and public use, represented the barriers also reported by subjects with high ability, who otherwise have few environmental restrictions. For the Environmental factors 27 subjects had fit residuals outside the range of ± 2.5 . Most of all subjects were located within ± 2.5 logits on the scale, hence none were deemed to be extreme persons.

Table V. Fit of the categories in set 2 of the component Activities and Participation of the Comprehensive Core Set for low back pain. Rasch analysis ($n = 118$)

ICF code	ICF category title	Location	SE	Residual	χ^2	p -value
d710	Basic interpersonal interactions	1.49	0.24	-1.24	2.82	0.24
d760	Family relationship	1.23	0.32	-1.01	1.16	0.56
d770	Intimate relationship	0.75	0.29	-0.53	0.44	0.80
d910	Community life	-0.07	0.18	-2.09	6.51	0.04
d859	Work and employment, other	-0.23	0.25	-0.45	1.46	0.48
d470	Using transportation	-0.27	0.18	-0.31	4.01	0.14
d475	Driving	-0.33	0.17	-0.14	0.69	0.71
d240	Handling stress and other psychological demands	-0.43	0.18	0.58	1.58	0.45
d920	Recreation and leisure	-2.14	0.18	0.05	1.92	0.38

ICF: International Classification of Functioning, Disability and Health; SE: standard error.

Table VI. Fit of the categories in the Environmental component of the Comprehensive Core Set for low back pain. Rasch analysis (n = 118)

ICF code	ICF category title	Location	SE	Residual	χ^2	p-value
e540	Transportation services	0.73	0.24	-0.70	2.84	0.24
e465	Social norms	0.61	0.29	-1.64	1.67	0.43
e110	Products and substances for personal consumption	0.54	0.24	0.92	1.03	0.60
e460	Societal attitudes	0.37	0.17	-0.21	4.12	0.13
e155	Design, construction and building products	0.12	0.29	-0.76	0.48	0.79
e330	People in position of authority	0.11	0.18	-0.57	0.87	0.65
e590	Labour and employment services	0.10	0.17	-0.42	6.38	0.04
e135	Products and technology for employment	0.07	0.19	-1.42	2.74	0.25
e150	Design, construction and building products	-0.05	0.29	-1.22	0.92	0.63
e575	General social support services	-0.07	0.18	-0.55	0.49	0.78
e120	Products/technology for personal indoor/outdoor mobility	-0.13	0.25	-0.21	0.35	0.84
e550	Legal services	-0.47	0.25	-1.52	1.88	0.39
e580	Health services	-0.49	0.16	-0.42	3.53	0.17
e325	Acquaintances, peers, etc.	-0.73	0.17	-1.77	8.36	0.02
e355	Health professionals	-0.80	0.16	-1.13	8.64	0.01

ICF: International Classification of Functioning, Disability and Health; SE: standard error.

The 15 Environmental categories presented in Table VI seemed to fit a unidimensional construct and the independent *t*-test shows that 9% of the person estimates were outside the range of ± 1.96 , CI 0.049–0.131.

DISCUSSION

The present results indicate that the qualifier levels should be reduced, and for several categories rescored, in order to discriminate. There was invariance of measurements across age, gender and education levels except for one category. Both with respect to body functions, activities and participation, relatively few categories were suited to the subjects with high level of function and mild activity restrictions. Many of the selected categories in the Comprehensive ICF Core Set for LBP reflected the underlying constructs of Body function, Activities and participation and Environmental factors in a unidimensional way.

We chose Rasch analysis and a partial credit model in order to evaluate whether the summed categories within the ICF components could be eligible as measurements. This model was used as we did not assume equidistance between thresholds of the qualifiers across the ICF categories (15). Sample sizes affect the interpretation of the Rasch analysis in several aspects. For the χ^2 based fit statistics, larger sample sizes increase the chance of misfit to the Rasch model. This should be taken into consideration regarding the relatively small sample size in the present study. The parameter estimates are also influenced by the sample size, with the DIF estimates being the most critical estimates (19). However, based on the calculations of Elasoﬀ (20), differences of 0.1 logits could be detected in DIF analysis in groups down to 25 subjects, given a power of 80%. Thus, the present sample size is sufficient in this respect. Another important factor for the results is the characteristics of the participating subjects. Three-quarters of the patients had chronic LBP. The mean score of functioning in ODI was 33%. This indicates worse functioning than a Norwegian population with acute LBP, and better functioning than those with chronic LBP (21, 22), in agreement with the mixed duration of LBP in the present patients. Hence, the subjects in the present study were fairly representative with respect to functioning of the LBP population seeking medical care (21, 23, 24). However, the diagnostic profile of the sample may influence the results. Unspeciﬁc LBP is the dominating label for LBP (25), which is also reflected in the present sample.

The qualifier levels of the ICF classiﬁcation have been criticised (26). In the present study very few patients had complete or severe problems in functioning according to the categories contained in the Comprehensive ICF Core Set. Subpopulations with more severe problems exist. However, most patients with LBP do not have complete problems in daily function and activities, and the present study indicates that a revision of the qualifier levels may be needed in order to use ICF as a measurement in this special population.

An important aspect to consider is to what extent the Comprehensive ICF Core Set works in the same way irrespective of which group is being assessed. Age and gender are important

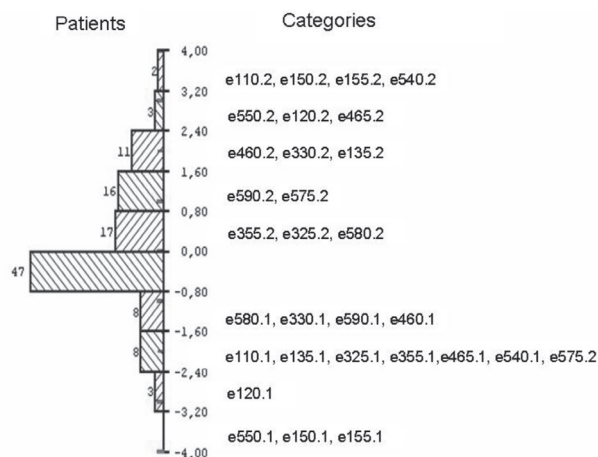


Fig. 3. Distribution of categories and patients in the Environmental component, (n = 118) along the Rasch-calibrated metric scale. The right hand panel shows the location of the threshold between barriers and no influence (1) and the threshold between no influence and facilitators (2) for the categories. The left-hand panel shows the distribution of patients.

factors influencing painful LBP conditions (27). Educational level is shown to be an important factor predicting the consequences for functioning in LBP (28). However, differences were found only for *d630 Preparing meals*. This activity was more difficult for subjects with poor functioning when they were less than 45 years of age, compared with over 45 years of age. The reason for this result is not obvious, but responsibilities for family and children in the younger group may be of relevance. Otherwise this set also responded similarly across age, gender and educational level.

The Rasch approach offers a linear transformation of the ordinal raw score of the Comprehensive ICF Core Set, and an ordering of categories and subjects along the same hierarchical continuum (9, 29). These attributes of Rasch imply the possibilities of exploring the underlying construct of the categories, the function of the ICF Core Set components across subgroups of subjects or patients, as well as the discriminating ability of the qualifiers. There is an ongoing discussion about the choice between the rating scale and the partial credit model in Rasch measurement. The rating scale model specifies that a set of items (ICF categories) shares the same rating scale structure, whereas the partial credit model specifies that each item has its own rating scale structure (30). By using the partial credit model we could prove that none of the components of the ICF comprehensive core set met the requirements of the same qualifier level structure across categories.

Clinical measurements rely on tools that can evaluate patients across functional levels. The hierarchical ordering of categories and patients can be used to evaluate this. In the component of Body functions, *b280 Sensation of pain* was the category easiest to endorse. This agrees with clinical experience that pain is the key feature in LBP and is affected in every patient to some extent. Sleep disturbances and problems with muscle endurance and exercise tolerance may also be reported, even when the subjects have few other problems with functioning. At the other end of the hierarchy of categories, *b180 Experience of self and time functions* and *b260 Proprioceptive functions* represent severe overall functional problems when reported. It is not surprising that the category *b620 Urination functions* also represents a high degree of functional impairment when reported. It is perhaps more surprising that *b640 Sexual functions* does. It may be that minor problems with sexual functions are under-reported, and that those reported represent problems in patients with more severe sciatica and radiculopathy (31). The patients had fewer problems with activities than reflected by the ICF categories, only *d415 Maintaining a body position* and *d410 Changing body position* represent problems even when the overall burden of LBP is slight. This problem was even larger for the participation set that did not match the distribution of the patients' problems. Only the category *d920 Recreation and leisure* was reported to be a problem by many patients with a relatively high degree of participation. Problems with relations are retained until the problems become severe. As the present study was undertaken in specialist practice, and the mismatch between the functional level of persons and categories may be even more pronounced when persons with LBP in primary care or in the general popu-

lation are evaluated. On the other hand, one may expect that the present categories would match more affected subgroups than included in the present study, for example, patients with radiculopathy and surgical candidates. The narrow logit scale for the environmental factors indicates a low difference in odds of reporting problems between the lowest and highest category in this component. Hence, the environmental factors do not discriminate well regarding different levels of functioning.

In the present study we evaluated unidimensionality according to the independent *t*-test, which is a rather strict approach (12, 17). The activity set did not pass the *t*-test, but was close.

There is debate about the activity and participation constructs of the ICF (32). WHO suggests that the different domains of this component can be grouped as either activities or participation, some domains represent both activities and participation and other either activities or participation or that each domain represents both activity and participation (1). In addition, recent studies point to 2 or more sub-domains (33, 34). A model including 4 components labelled Body functions and structures, Acts, Tasks, and Societal involvement has recently been suggested by Badley (35). In the present study the component Activity and participation failed to meet the Rasch demands for unidimensionality. However, most categories could be retained in 2 unidimensional sets. Each set contained categories from different domains in the component of Activities and Participation. For example, *d450 walking* and *d475 driving* classified in the ICF as mobility components were represented in different sets, none of which directly comparable to the concepts introduced by Badley (35). One of the main challenges is that the categories in the component of Activities and Participation can be scored either from the perspective of capacity or from the perspective of actual performance. Both concepts were integrated in the qualifier scoring of the present Core Set and may have influenced the dimensionality of the categories. In addition, the analysis revealed some local dependency between categories that may have influenced the results. The distribution of categories in the 2 sets neither reflected a distinction between activities and participation nor the dimensions of disability and handicap addressed by the former International Classification of Impairments, Disabilities and Handicaps (ICIDH) model (36).

Being engaged in working life is a very important part of participation (37). However, the category *d850 Remunerative employment* did not fit into any of the 2 subsets. Work may represent a separate dimension, but it is also possible that very different factors influence employment across individuals, and hence reduces the fit of this category. Further studies should explore this, for example within white- and blue-collar occupations.

According to the present analysis, several of the environmental categories also reflected a single underlying construct. Furthermore, categorizing the qualifiers into barriers, no influence, and facilitators seemed to discriminate the environmental influence in this component. From a clinical point of view it may be important to be able to measure the environmental factors that are rarely implemented in traditional measurements

for LBP. These factors may contribute to the person's level of functioning and should be focused on in rehabilitation.

The clinical implementation of the ICF will rely on developing practical tools, which allow classification and measurements (38). In order to improve clinical feasibility of the Comprehensive ICF Core Set one should evaluate a further reduction in categories. However, one would not want to lose categories containing key problems across different levels of functioning. In addition, it is important for a clinical measurement that the items are independent (lack of local dependency). Based on the present Rasch analysis, one can illustrate a strategy to choose categories for a future Comprehensive ICF Core Set. For example, *b280 Sensation of pain* and either *b180 Experience of self and time functions* or *b260 Proprioceptive functions* should be included in order to represent the borders of capability in functioning. *b260 Proprioceptive functions* is probably an important category with respect to body functions in LBP (39) and would be preferred to *b180 Experience of self and time functions*. The local dependency between *b180 Experience of self and time functions* and *b630 Preparing meals* would be another argument for omitting *b180*. *d410 Changing body position*, *d430 Lifting and carrying objects* and *d640 Doing housework* represented common problems representing the same level of difficulty. Hence, *d410* and *d430* showing local dependency may be omitted. *d530 Toileting* is a possible candidate for identifying subjects with severe limitations in activities. *d920 Recreation and leisure* was the category located at the bottom of the hierarchy, reflecting that problems within this category were reported by many patients with a relatively high degree of activities and participation. We suggest that it is important to include this category, as well as the category *d760 Family relationships* characterizing problems in subjects with a lower level of activities and participation. The category *e540 Transportation services* was most difficult to endorse, thus representing a barrier for the subjects with several environmental barriers, and may be important to include in the ICF Core Set.

In conclusion, the present Rasch analysis indicated that the Comprehensive ICF Core Set for LBP may be used with some modification of categories as a common tool for assessing problems within body functions, activities and participation. However, detecting ICF categories that better reflect the higher functional levels in patients with LBP, and a revision of the qualifier levels may be needed. These results lend support to the development of the Comprehensive ICF Core Set as a future clinical tool for measuring the functioning of patients with LBP.

REFERENCES

1. WHO. International Classification of Functioning, Disability and Health: ICF. Geneva: World Health Organization; 2001.
2. Stucki G, Grimby G. Applying the ICF in medicine. *J Rehabil Med* 2004; 36 Suppl 44: 5–6.
3. Cieza A, Ewert T, Ustun TB, Chatterji S, Kostanjsek N, Stucki G. Development of ICF Core Sets for patients with chronic conditions. *J Rehabil Med* 2004; 36 Suppl 44: 9–11.
4. Cieza A, Stucki G, Weigl M, Disler P, Jackel W, van der LS, et al. ICF Core Sets for low back pain. *J Rehabil Med* 2004; 36 Suppl 44: 69–74.
5. Ewert T, Fuessl M, Cieza A, Andersen C, Chatterji S, Kostanjsek N, et al. Identification of the most common patient problems in patients with chronic conditions using the ICF checklist. *J Rehabil Med* 2004; 36 Suppl 44: 22–29.
6. Weigl M, Cieza A, Andersen C, Kollerits B, Amann E, Stucki G. Identification of relevant ICF categories in patients with chronic health conditions: a Delphi exercise. *J Rehabil Med* 2004; 36 Suppl 44: 12–21.
7. Brockow T, Cieza A, Kuhlow H, Sigl T, Franke T, Harder M, et al. Identifying the concepts contained in outcome measures of clinical trials on musculoskeletal disorders and chronic widespread pain using the International Classification of Functioning, Disability and Health as a reference. *J Rehabil Med* 2004; 36 Suppl 44: 30–36.
8. Okochi J, Utsunomiya S, Takahashi T. Health measurement using the ICF: test-retest reliability study of ICF codes and qualifiers in geriatric care. *Health Qual Life Outcomes* 2005; 3: 46.
9. Andrich D. Relationship between the Thurstone and Rasch approaches to item scaling. *Appl Psychol Meas* 1978; 2: 449–460.
10. Svensson E. Guidelines to statistical evaluation of data from rating scales and questionnaires. *J Rehabil Med* 2001; 33: 47–48.
11. Angoff WH. Perspectives on differential item functioning methodology. In: Holland PW, Weiner H, editors. *Differential item functioning*. Hillsdale: Lawrence Erlbaum; 1993, p. 3–23.
12. Pallant JF, Tennant A. An Introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Brit J Clin Psychol* 2007; 46: 1–18.
13. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine* 2000; 25: 3115–3124.
14. Grotle M, Brox JI, Vollestad NK. Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. *J Rehabil Med* 2003; 35: 241–247.
15. Masters G. A Rasch model for partial credit scoring. *Psychometrika* 1982; 47: 149–174.
16. Andrich D, editor. *Rasch models for measurements. Quantitative Application in Social Sciences*, 68. London: Sage Publications; 1988.
17. Smith EV Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 2002; 3: 205–231.
18. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ* 1995; 310: 170.
19. Linacre JM. Sample size and item calibration stability. *Rasch Measure Tran* 1994; 7: 28.
20. Elashoff JD, editor. *NQuery Advisor version 4.0. User's guide*. Boston: Statistical Solutions; 2000.
21. Grotle M, Vollestad NK, Brox JI. Clinical course and impact of fear-avoidance beliefs in low back pain: prospective cohort study of acute and chronic low back pain: II. *Spine* 2006; 31: 1038–1046.
22. Grotle M, Vollestad NK, Veierod MB, Brox JI. Fear-avoidance beliefs and distress in relation to disability in acute and chronic low back pain. *Pain* 2004; 112: 343–352.
23. Holm I, Friis A, Storheim K, Brox JI. Measuring self-reported functional status and pain in patients with chronic low back pain by postal questionnaires: a reliability study. *Spine* 2003; 28: 828–833.
24. Storheim K, Brox JI, Holm I, Koller AK, Bo K. Intensive group training versus cognitive intervention in sub-acute low back pain: short-term results of a single-blind randomized controlled trial. *J Rehabil Med* 2003; 35: 132–140.
25. van TM, Koes B, Bombardier C. Low back pain. *Best Pract Res Clin Rheumatol* 2002; 16: 761–775.
26. Uhlig T, Lillemo S, Moe RH, Stamm T, Cieza A, Boonen A, et al. Reliability of the ICF Core Set for rheumatoid arthritis. *Ann Rheum Dis* 2007; 66: 1078–1084.

27. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van EH. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 1995; 62: 363–372.
28. Hagen KB, Tambs K, Bjerkedal T. What mediates the inverse association between education and occupational disability from back pain? – a prospective cohort study from the Nord-Trøndelag health study in Norway. *Soc Sci Med* 2006; 63: 1267–1275.
29. Rasch G, editor. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press; 1960.
30. Wright BD. Model selection: rating scale or partial credit? *Rasch Measurement Transactions* 1999; 12: 641–642.
31. McCarthy MJ, Aylott CE, Grevitt MP, Hegarty J. Cauda equina syndrome: factors affecting long-term functional and sphincteric outcome. *Spine* 2007; 32: 207–216.
32. Jette AM, Haley SM, Kooyoomjian JT. Are the ICF Activity and Participation dimensions distinct? *J Rehabil Med* 2003; 35: 145–149.
33. Eide AH, Jelsma J, Loeb M, Maart S, Toni MK. Exploring ICF components in a survey among Xhosa speakers in Eastern & Western Cape, South Africa. *Disabil Rehabil* 2008; 30: 819–829.
34. Jette AM, Tao W, Haley SM. Blending activity and participation sub-domains of the ICF. *Disabil Rehabil* 2007; 29: 1742–1750.
35. Badley EM. Enhancing the conceptual clarity of the activity and participation components of the International Classification of Functioning, Disability, and Health. *Soc Sci Med* 2008; 66: 2335–2345.
36. Masala C, Petretto DR. From disablement to enablement: conceptual models of disability in the 20th century. *Disabil Rehabil* 2008; 30: 1233–1244.
37. Keeley P, Creed F, Tomenson B, Todd C, Borglin G, Dickens C. Psychosocial predictors of health-related quality of life and health service utilisation in people with chronic low back pain. *Pain* 2008; 135: 142–150.
38. Stucki G, Kostanjsek N, Ustun B, Cieza A. ICF-based classification and measurement of functioning. *Eur J Phys Rehabil Med* 2008; 44: 315–328.
39. Renkawitz T, Boluki D, Grifka J. The association of low back pain, neuromuscular imbalance, and trunk extension strength in athletes. *Spine J* 2006; 6: 673–683.