DISABILITY AND REHABILITATION , 2000; VOL. 22, NO. 8, 363–371

Polytomous Rasch analysis as a tool for revision of the severity of disability code of the ICIDH

MARIJKE HOPMAN-ROCK*, STEF VAN BUUREN and MARIJKE DE KLEIJN-DE VRANKRIJKER

TNO Prevention and Health, Division Public Health, P.O. Box 2215, 2301 CE Leiden, The Netherlands

Accepted for publication: March 1999

Abstract

Purpose: To develop a preliminary proposal for the revision of the severity code of the ICIDH-D 1980.
Method: Quantitative analysis (polytomous Rasch analysis) of linked existing data sources including items about walking and dressing disability.
Results: The Rasch analysis provided estimates of threshold parameters for walking and dressing item categories. Factor analysis showed that more than one dimension was present, but that the first factor could definitely be interpreted as 'disability'. The reliability of the solutions was satisfactory (0.88 for walking and 0.91 for dressing).
Conclusion: Based on the results, tentatively a new severity code is proposed with more distinct categories of 'difficulty' than the existing code, and the use of technical aids and personal assistance in the higher ranks. The Rasch method could be a useful tool for calibrating and measuring disability, as well as for converting existing disability data into a new uniform severity of disability code.

Introduction

While revising the International Classification of Impairments Disabilities and Handicaps,[1–3] the Dutch WHO collaborating centre for the ICIDH was asked to develop proposals for the revision of the classification of disabilities of the ICIDH 1980 (ICIDH-D), including the severity code* of the ICIDH-D (table 1).

This revision occurred in successive phases. In the first phase, 96 well-known existing disability measures were reviewed, and the number of items per measure that are related to the ICIDH-D in the domains of personal care, locomotion, body disposition and dexterity were determined. Then criteria were established to select reliable and valid disability measures with 10 or more items that are related to the ICIDH-D classification in the selected domains. The selected measures differed from each other in the way the severity of disability was operationalized. These first two phases of the revision have been described in detail by Hopman-Rock and Miedema.[4] In the next phases a method using polytomous Rasch analysis was developed for quantitative analysis of a variety of existing data sources that included two or more of the previously selected disability measures. This culminated in a preliminary proposal for a revised code for the severity of disability.[5,6]

The definitions of 'impairment' and 'disability' are often subject to discussion[7] and will probably be changed, as indicated by the draft version ICIDH-2.[8] In our article, we use the term 'disability' as defined in the 1980 edition of the ICIDH and the corresponding severity code. We describe the methods used to develop a preliminary proposal for revision of the severity of disability code. The aim of the project was to develop a working method that facilitates and enhances development of such a severity code on the basis of quantitative analyses of (existing) data.

Methods

SELECTING DISABILITY MEASURES

In the last decade, several methods have been used to assess the presence and severity of disabilities. Some investigators have used (parts) of the ICIDH-D, with or without its severity code, while others have included questions about disability in population surveys or patient questionnaires. A popular method for indicating

* Usually, we speak of the Severity of Disabilities Scale (SDS), however, to avoid misunderstanding and confusion with existing disability measures we use in this paper the term 'severity code'.

Table 1 Existing severity coding of the ICIDH 1980 disability section[3]

0	Not disabled Includes: no disability present (the individual can perform the activity or sustain the behaviour unaided and on his own without difficulty).
1	Difficulty in performance Includes: difficulty present (the individual can perform the activity or sustain the behaviour unaided and on his own but only with difficulty).
2	Aided performance includes: aid and appliance necessary (the individual can perform the activity only with a physical aid or appliance).
3	Assisted performance Includes: the need for a helping hand (the individual can perform the activity or sustain the behaviour, whether augmented by aids or not, only with some assistance from another person).
4	Dependent performance Includes: complete dependence on the presence of another person (the individual can perform the activity or sustain the behaviour, but only when someone is with him most of the time). Excludes: inability.
5	Augmented inability Includes: activity impossible to achieve other than with the help of another person, the latter needing an aid or appliance to enable him or her to provide this help (for example, the individual cannot get out of bed other than by the use of a hoist); behaviour can be sustained only in the presence of another person and in a protected environment.
6	Complete inability Includes: activity or behaviour impossible to achieve or sustain (for example, an individual who is bed-bound is also unable to transfer).
8	Not applicable.
9	Severity unspecified.

Table 2 Selected measures of severity of disability[4]

Type of scale	*Selected measure*	*Number of relevant ICIDH topics*[3]
Ordinal	Health Assessment Questionnaire (HAQ)	15
	Functional Status Index (FSI)	12
	OECD Long-Term Disability Questionnaire	11
	Pediatric Evaluation of Disability Inventory (PEDI), part I	14
	Rehabilitation Activity Profile	15
	Physical Abilities Scale	18
	Osteoporosis Functional Disability Scale	16
	Barthel ADL Index	11
	Northwick Park Index of Independence in ADL	14
	Arthritis Impact Measurement Scale (AIMS)	16
	Groningen Activity Restriction Scale (GARS)	12
	Childrens HAQ	15?
	Rivermead ADL Scales	15
	OARS Activities of Daily Living	11
	Functional Independence Measure (FIM)	12
	ADL Rating Scale	12
	Pediatric Evaluation of Disability Inventory (PEDI), part II*	12
Weighted items	OPCS Disability Scales	20
	Sickness Impact Profile (SIP)	17
Time scores	Physical Performance Test (PPT)	8

See references 9–27.

the severity of disabilities is the use of standardized measures (often patient questionnaires), in which the degree of restriction in performing certain activities is assessed by using ordinal scales, weighted items, or time scores. Often, an overall (severity of) disability score can be calculated by summing item scores. Although many

measures have been developed, especially in the field of activities of daily living (ADL), their use is often restricted to certain research projects, research groups or fields of research. In most cases there is no direct relationship between the disability score (individual item or sum score of items) obtained with these measures and the existing severity code of the ICIDH-D. As a first step in our revision project we made an inventory of 96 measures that are currently used in different (research) fields such as population surveys and statistics, rehabilitation, vocational assessment, nursing homes, etc. The severity of disability could be evaluated with 77 of these measures.

For each measure we assessed the way in which the severity of disability was defined and its reliability, validity, and relation to the ICIDH (on the 2-digit level). We then asked researchers and clinicians to give their opinion about the selected measures and the current severity coding of the ICIDH 1980.[4] In this way, we selected 19 popular, valid and reliable disability measures (see table 2)[9–27]

We then classified each of the relevant items in the selected measures into the most appropriate category of the ICIDH-D. Disability categories that were measured most frequently were walking, dressing, disability in transfer to the toilet, bathing, other personal hygiene activities, feeding, climbing stairs, transfer, and subsistence.

LINKAGE

Our next step was to combine existing, overlapping sources of data on disability for different groups of respondents as follows:

(1) choose relevant disability measures;
(2) obtain data sources that contain responses for at least two of these measures;
(3) select all items that belong to the same ICIDH-D category;
(4) construct a linked data set of items;
(5) estimate the severity per item category by polytomous Rasch analysis;
(6) order item categories according to this severity; and
(7) construct a severity code.

The main criterion for inclusion was that a data set had to contain data about sex and age, and at least two of the selected disability measures for each respondent (see table 2) so that the overlap between different sources could be exploited. We wrote to several authors of articles which described a relevant data set, and asked for their co-operation.

DESCRIPTION OF INDIVIDUAL DATA SOURCES

Liang

This data set included five health status instruments that were administered, in random order, to 50 patients with arthritis before and after total joint arthroplasty.[28] The patients were aged 50 to 80 years and they had a diagnosis of rheumatoid arthritis or osteoarthritis. Subjects with cognitive impairments, language barriers or visual or hearing deficits were excluded. Four of the five health status instruments (the FSI (Difficulty, Pain and Help section), the HAQ, the AIMS and the SIP) used were on our list of selected disability measures. After 1 year, a follow-up study was carried out[29] with the same group (response n = 38). Only data for the last mentioned group were available at an item level. The mean age was 67.4 years and 58% of the subjects were women. Eighty-seven per cent of the sample suffered from osteoarthritis and 13% from rheumatoid arthritis.

ERGOPLUS

This data set included data from the Rotterdam study[30] (in Dutch, the ERGO study; ERGO= Erasmus Rotterdam Health Research on Elderly People). The HAQ was used (n= 2,895)[31] and the SIP was used in a sub-sample of 306 persons.[32] All subjects were aged 55 to 75 years and lived independently. The mean age of the sub-sample was 64.8 years (SD= 5.6), and 68% of the subjects were women.

EURIDISS

EURIDISS (European Research on Incapacitating Diseases and Social Support) is an international longitudinal study of patients with recently diagnosed rheumatoid arthritis.[33,34] This data set included data for 242 patients, and the HAQ and the GARS were used in combination. The mean age of the sample was 53.9 (SD = 11.8) and 64% of the subjects were women.

CBS

This file is a public microdata file of health survey interviews conducted in 1994 by Statistics Netherlands

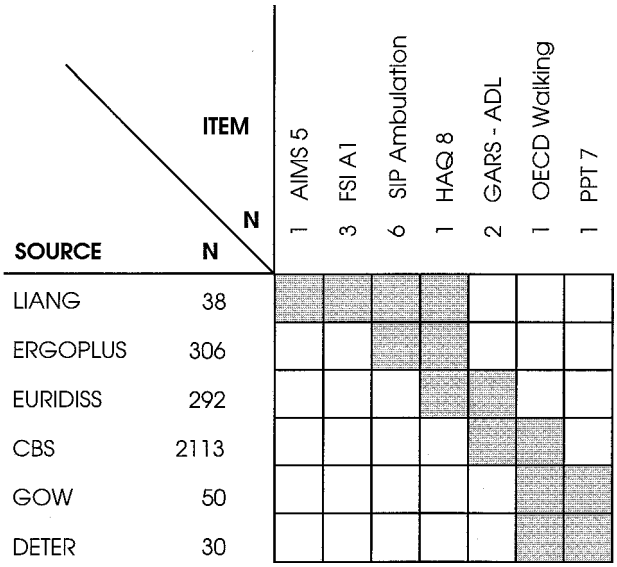

Figure 1 Linkage diagram of walking items

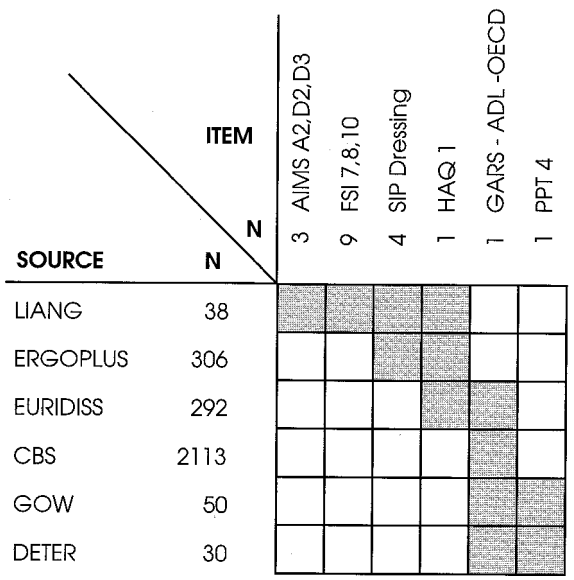

Figure 2 Linkage diagram of dressing items

(CBS = Centraal Bureau voor de Statistiek) among people aged 16 years and older. Items on long-term disability (parts of the OECD disability indicator) and items on Activities of Daily Living linked this data set with those of EURIDISS and GOW (see next section). These parts of the health interview were only completed by people aged 55 years and older (N = 2,113). The ADL questions related to eating and drinking, getting in and out of a chair, getting in and out of bed, dressing and undressing, moving to another room on the same floor, walking up and down the stairs, leaving and entering the house, moving outside the house, washing the face and hands, washing the entire body. The answer categories were: without difficulty, with some difficulty, with great difficulty and only with help from others. The questions about 'moving to another room on the same floor', 'moving outside the house' (both ICIDH 40 walking) and 'dressing and undressing' were used. These questions were similar to some of the items of the GARS. The mean age of the sample was 66.6 years (SD 8.8), and 55% of the subjects were women.

GOW (*goed oud worden = ageing well*)

This data set contains information about disability collected by means of the OECD disability indicator and the Physical Performance Test (PPT). Fifty healthy people aged 75 to 85 years (mean age 78.7 years, SD 3.0) took part, 58% of whom were women.[35] All people lived independently in the city of Leiden and were participants of a course called 'ageing well'. The pre-test measures of disability were used.

DETER

This data set (determinants of immobility and physical activity) contains information about 30 older individuals (age 75 years and over) living in the city of Leiden who were on a waiting list for home care. In a small pilot study[36] these individuals were asked about their diseases/disorders, their complaints and their physical activity. One of the functional tests was the PPT. The OECD indicator was also used. The mean age of the sample was 78.6 years, (SD 3.2) and 70% of the subjects were women.

We used the items about walking and dressing disability because these items were common to most of the data sets. Figure 1 and figure 2 are linkage diagrams showing which disability items were used in each source of data. '1 AIMS 5' means that one item (the fifth question of the AIMS about walking disability) was used in the data set of Liang. One question of the HAQ was used in the data sets of Liang, ERGOPLUS, and EURODISS. The CBS data set contained items of the GARS as well as items of the OECD indicator. The GOW data set contained items of the OECD and the PPT. The data set DETER contained the OECD and the PPT. All the data sets with walking items and the data set with dressing items were linked[37] in two separate data sets. These two linked data sets contained a lot of

missing data (about 76%, see figures 1 and 2). The final linked data set of walking items contained information from 2484 individuals on 14 linked items (one SIP item had to be removed because none of the respondents scored the 'yes' answer on the statement 'I do not walk at all'). The linked dataset of dressing items contained information from 2670 individuals on 17 items (the items AIMS A2, and FSH2 were removed because of zero frequencies).

Because a lot of data were missing from the two linked data sets, we used multiple imputation[38] to facilitate factor analysis. This imputation was programmed in SAS/IML[39] as a Gibbs sampler.[40]

POLYTOMOUS RASCH ANALYSIS

A common severity of disability dimension for all items can be constructed and validated by applying so-called Item Response models. The Rasch Model[41] is the basic item response model, and is particularly suited for constructing scales and for equating information from different sources. The Graded Response (GR) model [42] is a generalization to handle items with more than two categories (polytomous). The GR model stipulates that the response probability for each of K item categories varies only with the true level of disability. The GR model uses logistic functions ('S' curves) to describe the relationship between the 'true' (virtual) disability and the chance of a certain response for a given individual item. The point on the 'true' disability scale where the chance of a response in category k or higher is 50% (the bend in the S-curve) is called the *threshold parameter*. Relevant computations for the estimation of the threshold parameters for each item were made using MULTILOG 6.03.[43] With the polytomous Rasch model all item categories are ordered such that item categories reflecting less disability come before item categories reflecting greater disability. Individuals and groups of individuals can be ordered according to the 'true' disability dimension.

VALIDITY AND RELIABILITY

A standard approach to test if the 'true' disability is in fact one dimensional is to apply factor analysis. SAS PROC FACTOR was used on the imputed data sets (with two replications) for a maximum likelihood factor analysis. The eigenvalue-larger-than 1 criterion and the elbow-rule were used to determine the number of factors. To test reliability, Cronbach's alpha was estimated from item-item correlations that were observed as alpha = $mr/(1+(m-r)r)$, where r is the average correlation and m is the number of items.

Results

WALKING DISABILITY

Table 3 lists the estimates, obtained by polytomous Rasch analysis, of the threshold parameters for all walking item categories. The item category with the highest threshold parameter (6.83) was category 4 of the 7th item of the GARS: Can you, fully independently, get around in the house (if necessary with a cane)? answer category 4 = 'No I can not do this independently, only with help from others'. According to the model, people that answered this category will also have problems with the activities described in all other items included in the linked data set. The item category with the lowest threshold parameter (−.22) was category 2 of the 7th item of the PPT: walking 15 meters in 15 to 20 seconds. The item categories were then ordered according to their

Table 3 Estimated threshold parameters by polytomous Rasch analysis of categories of walking items

Estimated threshold	*n*	*Item*	*Category*	*Description*
6.83	4	GARS7	4	Inside: only with help
6.41	2	HAQ8	4	Outdoors: unable
5.41	2	SIP8	2	Only walk with help
4.72	26	GARS7	3	Inside: much difficulty
4.22	49	GARS9	4	Outdoors: only with help
3.96	1	FSI-H	2	Inside: used cane, etc.
3.84	9	SIP11	2	Use frame, crutches, etc.
3.27	15	SIP7	2	Limp, wobble, etc.
2.91	6	PPT7	5	Cannot walk 15 m
2.87	106	GARS9	3	Outdoors: much difficulty
2.74	3	FSI-D	3	Inside: moderate difficulty
2.73	3	FSI-P	3	Inside: moderate pain
2.73	57	HAQ8	3	Outdoors: much difficulty
2.67	3	AIMS5	2	Unable unless assisted
2.51	174	OECD	4	Cannot walk 400 m
2.44	178	GARS7	2	Inside: some difficulty
2.19	37	SIP1	2	Shorter distances
2.03	72	OECD	3	400 m: much difficulty
1.59	10	PPT7	4	15 m: > 25 sec.
1.38	4	FSI-D	2	Inside: mild difficulty
1.29	312	GARS9	2	Outside: some difficulty
1.28	71	SIP12	2	More slowly
1.09	209	OECD	2	400 m: some difficulty
1.09	116	HAQ8	2	Outdoors: some difficulty
1.06	6	PPT7	3	15 m: 20–25 sec.
0.93	6	FSI-P	2	Inside: mild pain
−0.22	20	PPT7	2	15 m: 15–20 sec.

n = number of respondents; item = item of a selected disability measure; category = item category; description = description of item category.

Table 4 Factor analyses (2-factor solutions, 2 replications) on linked data set walking items

2-*Factor solution* (*unrotated*)					
	FACTOR1	*FACTOR2*		*FACTOR1*	*FACTOR2*
AIMS5	0.50	– 0.24	AIMS5	0.44	0.06
FSI-P	0.79	– 0.43	FSI-P	0.69	– 0.31
FSI-D	0.88	– 0.17	FSI-D	0.61	– 0.58
SIP1	0.31	– 0.13	SIP1	0.29	– 0.36
SIP7	0.30	– 0.29	SIP7	0.26	– 0.36
SIP12	0.44	– 0.24	SIP12	0.56	– 0.40
HAQ8	0.91	– 0.07	HAQ8	0.82	– 0.02
GARS7	0.48	0.39	GARS7	0.46	0.40
GARS9	0.63	0.56	GARS9	0.52	0.48
OECD	0.42	0.74	OECD	0.56	0.71
PPT7	0.08	0.76	PPT7	0.41	0.37
Eigenvalue	3.70	2.06	EV	3.20	1.93

Factor solution for two different imputated data sets.

threshold parameters, such that the higher the estimated threshold parameter the greater the disability measured. When the disability is less severe, then item categories with 'with some or mild difficulty' dominate.

Factor analysis suggested a two-factor solution (see table 4), because higher factors had a eigenvalue lower than 1. The first factor (eigenvalue 3.70) could be interpreted as the expected walking disability factor. Interpretation of the second factor (eigenvalue 2.06) was not straightforward and the solution was less stable. Cronbach's alpha for this solution was 0.88, so the reliability was satisfactory.

DRESSING DISABILITY

Table 5 gives the estimated threshold parameters for the dressing items. In this case, again a PPT item category (coat on and off in less than 10 seconds) had the lowest threshold parameter, – 0.82 (according to the model, people who could complete this task have no problems with the activities associated with all the other items). Higher threshold parameters were found for item categories concerning dressing performed with some or mild difficulty, and with moderate or much difficulty. Item categories with estimated threshold parameters higher than 4 involved the use of special devices. The highest threshold parameter (6.24) was found for dressing with the help of someone else (SIP35).

Table 5 Estimated threshold parameters by polytomous Rasch analysis of categories of dressing items

Estimated threshold	*n*	*Item*	*Category*	*Description*
6.24	1	SIP35	2	Dress only with someone's help
4.96	28	GAO	4	Dress and undress: only with help
4.58	1	FSI-D 8	4	Buttoning clothes: severe difficulty
4.57	5	SIP31	2	Require help with buttons, zips etc.
4.50	1	FSI-H7	2	Underpants: special device
4.50	1	FSI-H10	2	Shoes: special device
4.39	16	HAQ1	4	Dress myself: cannot
4.11	35	GAO	3	Dress and undress: much difficulty
3.69	1	FSI-D7	3	Underpants: moderate difficulty
3.68	2	FSI-D10	4	Shoes: severe difficulty
3.67	3	PPT4	5	Coat on and off: > 20 sec.
3.44	14	SIP34	2	Dress myself very slowly
3.03	3	AIMS2	2	Cannot button articles
3.01	37	HAQ1	3	Dress myself: much difficulty
2.67	27	SIP29	2	Trouble with shoes
2.53	2	FSI-D10	3	Shoes: moderate difficulty
2.21	5	AIMS3	2	Cannot easily tie shoes
2.14	4	FSI-D8	3	Buttoning clothes: moderate difficulty
2.12	277	GAO	2	Dress and undress: some difficulty
1.79	1	FSI-D8	2	Buttoning clothes: mild difficulty
1.64	13	PPT4	4	Coat on and off: 15–20 sec.
1.37	110	HAQ1	2	Dress myself: some difficulty
0.83	8	FSI-D7	2	Under pants: mild difficulty
0.73	12	PPT4	3	Coat on and off: 10–15 sec.
0.29	8	FSI-D10	2	Shoes: mild difficulty
– 0.82	24	PPT4	2	Coat on and off: < 10 sec.

GAO is a compromise of linking GARS and OECD with the ADL questions of the CBS data base, there were slight differences in formulation of the questions.

Table 6 Factor analyses (2-factor solutions, 2 replications) on linked data set dressing items

2-Factor solution (*unrotated*)					
	FACTOR1	*FACTOR2*		*FACTOR1*	*FACTOR2*
AIMS2	0.35	−0.46	AIMS2	0.37	0.67
AIMS3	−0.09	0.77	AIMS3	0.61	−0.48
FSI-D7	−0.29	0.84	FSI-D7	0.12	0.45
FSI-D10	0.48	0.19	FSI-D10	0.59	0.71
SIP29	0.52	0.33	SIP29	0.67	−0.19
SIP31	0.64	0.10	SIP31	0.70	−0.04
SIP34	0.72	0.42	SIP34	0.65	−0.08
HAQ1	0.81	0.18	HAQ1	0.78	−0.40
GAO	0.45	0.12	GAO	0.22	−0.42
PPT4	0.25	−0.48	PPT4	−0.29	−0.41
Eigenvalue	3.07	2.25	EV	2.98	2.24

Note: factor solution for two different imputated data sets.

The factor analysis indicated a three-factor solution, but was very difficult to interpret. The two-factor solution with two replications is reported here (see table 6). The eigen value of the first factor was 3.07 , and of the second 2.25. Factor 1 could be interpreted as a dressing disability factor. Cronbach's alpha for the solution for dressing items was 0.91.

The results of the Rasch analyses on walking and dressing disability were then compared with the existing severity code of the ICIDH 1980 (Table 7). There were substantial empirical differences in severity between the levels of the current severity code 'difficulty in performance' (code 1). In our tentatively proposed severity code, which is based on the results of Rasch analysis, only code 6 (complete inability) remained the same as in the existing severity code.

Discussion

One of the frequently discussed but as yet unsolved problems in the field of disability severity coding is whether 'difficulty' and 'assistance required' (personal or non-personal) reflect the same level of disability. It does not seem illogical to classify the severity of disability as being greater if someone starts to use equipment or gets personal assistance (implying a greater dependency). Because the results of our analyses did not show a good interpretable second dimension of 'dependency', we recommend using only one dimension of severity of disability. However, it is possible that if someone uses some form of aid, such as a cane, then he or she might have an 'improved' score on the so-called 'dependency handicap' scale (see ICIDH 1980).

The Rasch analysis makes it possible to convert the current ICIDH severity code into the proposed severity code, and also to convert the severity, as measured with existing disability items, into the proposed severity code, or to convert existing items into other (existing or novel) items of disability measures by using the estimated threshold parameters.

Because this article describes the first application of the Rasch model to this field, it is obvious that it will have some limitations. First, we investigated two types of disability (walking and dressing) only. Second, we only used 7 of the 19 selected disability measures. Third, there were very few persons with severe disabilities in the sample. Fourth, because the data were incomplete we could only apply rigid models with strong assumptions. These limitations could affect the outcome. The factor analyses showed that besides the 'disability' dimension other dimensions were present but we were unable to

Table 7 Combined results of the polytomous Rasch analyses (n combined = 2500) for walking and dressing items, compared with the existing severity code of the ICIDH 1980

Level of threshold parameter	*Proposed* severity code*	*Examples of item-categories in disability measures*	*Existing severity code ICIDH-D 1980*
0	Not disabled	Are you able to walk outdoors on flat ground? (HAQ): without any difficulty Can put a coat on and off in less than 10 seconds (PPT)	0
1	mild difficulty	I walk more slowly (SIP):yes. Buttoning clothes (FSI): mild difficulty	1
2	Moderate difficulty	Can you walk 400 m without resting? (OECD): no I can't Can you easily tie a pair of shoes? (AIMS): no	1
3	Severe difficulty	Cannot walk 15 meters (PPT). Can you button articles or clothing? (HAQ):no	1
4	Technical aids	I get about only by using a walking frame, crutches, stick, walls, or hold on to furniture (SIP):yes. Putting on underpants (FSI): used a cane, special equipment or other device	2
5	Personal assistance	I only walk with help of someone else (SIP): yes. Can you fully independent dress and undress yourself? (GARS): no, only with help of someone else.	3,4,5

* Proposed by the authors of this paper. Proposal by WHO see reference number 8.

interpret them. This aspect may need to be addressed in future analyses.

If the revision process is continued by applying Rasch analyses to new representative data sets, then it will be possible to validate our tentatively proposed severity code and to make standard tables for the verification and conversion of widely-used disability items. The potential of the Rasch model for the calibration and measurement of disability has already been recognized by other investigators such as Tennant and McKenna,[44] Revicki and Cella[45] and Martin and Elliot.[46] The techniques described here could be used to re-analyse old data sets and to convert existing measures into a uniform severity of disability code. This would greatly increase the possibility for comparing populations with disabilities at a national and international level.

Acknowledgments

We thank Dr M. H. Liang (Department of Medical Rheumatology, Brigham and Women's Hospital, Boston MA, USA), Dr T. P. B. M. Suurmeyer (Northern Centre for Health Care Research, University of Groningen, The Netherlands), and Dr E. Odding, Erasmus Medical School, Rotterdam, The Netherlands) for making their data bases available for our analyses. We also like to thank our colleague Dr Harald Miedema for his stimulating role in the initial stage of the project.

This project was financially supported by a grant from the WHO Collaborating Centre for the ICIDH in the Netherlands.

Appendix

In the draft version of the ICIDH-2 (issued for field trials in June 1997) two qualifiers for activity are suggested: 'difficulty' and 'personal or non-personal assistance'.

Difficulty is to be rated as follows:

0 = no difficulty
1 = slight difficulty
2 = moderate difficulty
3 = severe difficulty
4 = unable to carry out the activity
9 = level of difficulty unknown

Assistance is to be rated as follows:

0 = no assistance used
1 = non-personal assistance
2 = personal assistance
3 = both non-personal and personal assistance
9 = level of assistance unknown.

References

1 Heerkens YF, Brandsma JW, Ravensberg van CD. A survey of criticism about the classification of impairments and the classification of disabilities of the International Classification of Impairment, Disabilities and Handicaps (ICIDH). Nationale Raad voor de Volksgezondheid, Zoetermeer, 1993. (93WCC43.3).
2 Hirs W. Rules for application of the ICIDH. Meeting on the initial revision of the ICIDH, Strasbourg, November 1990.
3 WHO. *International Classification of Impairments, Disabilities and Handicaps*. Geneva: World Health Organization, 1980, 1993.
4 Hopman-Rock M, Miedema HS. The development of a proposal for revision of the Severity of Disabilities Scale of the ICIDH. Personal care, body disposition, locomotor and dexterity disabilities: phase 1 and phase 2. Leiden: TNO Prevention and Health. Publ.nr. 95.043, 1995.
5 Buuren van S, Hopman-Rock M, Miedema HS. The development of a proposal for revision of the severity of disabilities scale of the ICIDH. Personal care, body disposition, locomotor and dexterity disabilities: phase 3. Leiden: TNO Prevention and Health. Publ. Nr. 96.067, 1996.
6 Hopman-Rock M, Miedema HS, Buuren van S, Kleijn-de Vrankrijker MW de. WHO collaborating centre for the ICIDH in the Netherlands. Towards revision of the severity of disability scale of the ICIDH. *Newsletter CSIZ* 1997; **3**: 3–4.
7 Brandsma JW, Lakerveld-Heyl K, van Ravensberg CD, Heerkens YF. Reflection on the definition of impairment and disability as defined by the World Health Organization. *Disability and Rehabilitation* 1995; **17**: 119–127.
8 WHO. International Classification of Impairments, Activities and Participation. A manual of dimensions of disablement and functioning. Beta 1-draft for field trials. Geneva: WHO, 1997.
9 Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire (HAQ) disability and pain scales. *Journal of Rheumatology* 1982; **9**: 789–93.
10 Jette AM. Functional status index: reliability of a chronic disease evaluation instrument. *Archives of Physical Medicine and Rehabilitation* 1980; **61**: 395–401.
11 McDowell I. Screening for disability. An examination of the OECD survey questions in a Canadian study. *Review Epidémiology Santé Publique* 1981; **29**: 421–429.
12 Haley SM, Coster WJ, Ludlow LH, Haltiwanger JT, Andrellos PJ. Pediatric Evaluation of disability Inventory (PEDI). Boston (MA): New England Medical Centre and PEDI Research Group, 1992.
13 Bennekom van CAM, Jelles F, Lankhorst GJ. Rehabilitation activity profile: the ICIDH as a framework for a problem-oriented assessment method in rehabilitation medicine. *Disability and Rehabilitation* 1995; **17**: 169–175.
14 Helmes E, Instructions for using the Physical Abilities Scale (PAS). Psychology Department, St. Josephs Health Centre, 35 Grosvenor St., London, Ontario Canada N6A 1Y6.
15 Helmes E, Hodsman A, Lazowski D. *et al*. Reliability, validity and sensitivity to intervention of a questionnaire to evaluate disability in osteoporotic patients with vertebral compression fractures. *Journal of Gerontology* 1995; **50**: M91–98.
16 Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Maryland State Medical Journal* 1965; **14**: 61–65.
17 Benjamin J. The Nortwick Park ADL index. *British Journal of Occupational Therapy* 1976; **39**: 301–306.
18 Meenan RF. New approaches to outcome assessment: the AIMS questionnaire for arthritis. *Advances in Internal Medicine* 1986; **31**: 167–185.
19 Kempen GIJM, Doeglas DM, Suurmeijer TPBM, Het meten van problemen met zelfredzaamheid op verzorgend en huishoudelijk gebied met de Groningen Activiteiten Restrictieschaal (GARS). Een handleiding[Groningen Activity Restriction Scale]. Groningen, The Netherlands:University of Groningen, Northern Centre for Health Care Research, 1993.
20 Anderson-Gäre B, Fasth A, Wiklund I. Measurement of functional status in juvenile chronic arthritis: evaluation of a Swedish version of the Childhood Health Assessment Questionnaire. *Clinical and Experimental Rheumatology* ,1993; **11**: 569–576.
21 Whiting S, Lincoln N. An ADL assessment for stroke patients. *British Journal of Occupational Therapy* 1980; **43**, 44–46.

22 Fillenbaum GG, Smyer MA. The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire. *Journal of Gerontology* 1981; **36**: 428–434.

23 Kidd D, Stewart G, Baldry J *et al*. The Functional Independence Measure: a comparative validity and reliability study. *Disability and Rehabilitation* 1995; **17**: 10–14.

24 Dinnerstein AJ, Lowenthal M, Dexter M. Evaluation of a rating scale of ability in activities of daily living. *Archives of Physical Medical Rehabilitation* 1965; **46**: 579–584.

25 Martin J, Meltzer H, Elliot D. The prevalence of disability among adults. Office of population censuses and surveys. London: HSMO, 1988.

26 Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care* 1981; **19**: 787–805.

27 Reuben DB, Siu AL. An objective measure of physical function of elderly patients: the physical performance test. *Journal of the American Geriatric Society* 1990; **38**: 1105–12.

28 Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis and Rheumatism* 1985; **28**: 542–7.

29 Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Medical Care* 1990; **28**: 632–42.

30 Hofman A, Grobbee DE, De Jong PTVM, van den Ouweland FA. Determinants of disease and disability: the Rotterdam elderly study. *European Journal of Epidemiology* 1991; **7**: 403–22.

31 Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Association of locomotor complaints and disability in the Rotterdam study. *Annals of the Rheumatic Diseases* 1995; **54**: 721–5.

32 Hopman-Rock M, Odding E, Hofman A, Kraaimaat FW, Bijlsma JWJ. Physical and psychological disability in elderly subjects in relation to pain in the hip and/or knee. *Journal of Rheumatology* 1996; **23**: 1037–44.

33 Suurmeyer TPBM, Doeglas DM, Moum T, Briancon S. *et al*. The Groningen activity restriction scale for measuring disability: its utility in international comparisons. *American Journal of Public Health*, 1994; **84**: 1270–73.

34 Kempen GIJM, Miedema I, Ormel J, Molenaar W. The assessment of disability with the Groningen activity restriction scale. Conceptual framework and psychometric properties. *Social Science & Medicine* 1996; **43**: 1601–1610.

35 Hell van L, Hopman-Rock M. Ontwikkeling en evaluatie van het programma 'goed oud worden': de testfase. Leiden: TNO PG publ.nr 95.040, 1995.

36 Hopman-Rock M. Determinanten van immobiliteit en fysieke activiteit. Een pilotstudie onder zelfstandig wonende ouderen die op de wachtlijst voor thuiszorg staan. Leiden: TNO NIPG. Publ. nr. 94.007, 1994.

37 Vale CD. Linking item parameters onto a common scale. *Applied Psychological measurement* 1986; **10**: 333–44.

38 Schafer JL. *Analysis of Incomplete Multivariate Data*. London: Chapman and Hall, 1997.

39 SAS Institute Inc. *SAS/STAT User's Guide*, Version 6, Fourth Ed. Cary, NC: 1990.

40 Van Buuren S, Brand JPL. TRIM: A multiple imputation program in SAS/IML. Department of Statistics, TNO Prevention and Health, 1996.

41 Rasch G. *Probabilistic Models for Some Intelligence and Attainments Tests*. Chicago: Meta press, 1980

42 Samejima F. Estimation of latent ability using a response pattern of graded responses. *Psychometrika* 1969; **34**: Monograph 17.

43 Thissen D. *MULTILOG User's Guide*. Chicago: Scientific Software, Inc., 1991.

44 Tennant A, McKenna SP. Conceptualizing and defining outcome. *British Journal of Rheumatology* 1995; **34**: 899–900.

45 Revicki DA, Calla DF. Health status assessment for the twenty-first century: item response theory, item banking and computer adaptive testing. *Quality of Life Research* 1997; **6**: 595–600.

46 Martin J, Elliot D. Creating an overall measure of severity of disability for the office of population censuses and surveys disability survey. *Journal of the Royal Statistical Society* 1992; **155**:121–140.