Quality of Life Research (2005) 14: 1225–1237
DOI 10.1007/s11136-004-4773-0

Content comparison of health-related quality of life (HRQOL) instruments based on the international classification of functioning, disability and health (ICF)

Alarcos Cieza[1] & Gerold Stucki[1,2]
[1]*ICF Research Branch, WHO FIC Collaborating Center (DIMDI), IMBK, Ludwig-Maximilians-University, Munich, Germany (E-mail: gstucki@med.uni-muenchen.de);* [2]*Department of Physical Medicine and Rehabilitation, Ludwig-Maximilians-University, Munich, Germany*

Accepted in revised form 14 October 2004

Abstract

The increasing recognition of the patient perspective and, more specifically, functioning and health, has led to an impressive effort in research to develop concepts and instruments to measure them. Health-Related Quality of Life (HRQOL) and the International Classification of Functioning Disability and Health (ICF) represent two different perspectives from which to look at functioning and health. Therefore, it is expected that both will often be used concurrently in clinical practice, research and health reporting. The objective of our study was to examine the relationship between six HRQOL instruments (the SF-36, the NHP, the QL-I, the WHOQOL-BREF, the WHODASII and the EQ-5D) and the ICF. All six HRQOL instruments were linked to the ICF separately by two trained health professionals according to ten linking rules developed specifically for this purpose. The degree of agreement between health professionals was calculated by means of the kappa statistic. Bootstrapped confidence intervals were calculated. In the 148 items of the 6 instruments a total of 226 concepts were identified and linked to the ICF. The estimated kappa coefficients range between 0.82 and 0.98. The concepts contained in the items of the HRQOL instruments were linked to 91 different ICF categories, 17 categories of the component *body functions*, 60 categories of the component *activities* and *participation*, and 14 categories of the component *environmental factors*. Twelve concepts could not be linked to the ICF at all. In the component *body functions*, only *emotional functions* are covered by all examined instruments. In the component *activities and participation*, all instruments cover aspects of work, but the half of them scarcely cover aspects of mobility. Only four of the six instruments address *environmental factors*. The ICF proved highly useful for the comparison of HRQOL instruments. The comparison of selected HRQOL instruments may provide clinicians and researchers with new insights when selecting health-status measures for clinical studies.

Key words: Content validity, Generic instruments, Health-related quality of life (HRQOL), Health status measures, ICF

Introduction

The patient perspective is the core of health-care provision and research. From the patient perspective, functioning and health are of utmost importance. Ultimately, any health-care intervention is intended to restore impaired body structures and functions, to overcome activity limitations and participation restrictions, and to prevent new symptoms and disabilities from developing. The increasing recognition of the patient perspective and, more specifically, functioning and health, has led to an impressive effort in research to develop concepts and instruments to measure them.

Health-Related Quality of Life (HRQOL) instruments are increasingly being used to describe and evaluate functioning and health in clinical trials and clinical practice, as well as various other fields of research [1]. The main purpose of HRQOL instruments is to describe the burden of disease of the population studied. HRQOL instruments focus on activities and participation, which are considered to be the components most relevant to patients and society, and are relevant to all health conditions. Such instruments enable functioning and health to be compared across health conditions, populations, and interventions. Generic health profiles, including the Short-Form 36 (SF-36) [2], the Nottingham Health Profile (NHP) [3] or the World Health Organization Disability Assessment Schedule (WHODASII) [4] are best suited for this purpose. HRQOL indices or utility measures, such as the European Quality of Life instrument (EQ-5D) [5], are used to measure the societal burden of disease and the economic benefit of interventions across conditions, populations, and settings.

Many studies compare the psychometric properties of HRQOL instruments, but content comparisons are scarcely represented in the literature. This is probably due to the varying use of concepts, scales, and items in the different HRQOL instruments. With the newly available International Classification of Functioning, Disability and Health (ICF) [6], a universal framework of functioning and health, the representation of items and scales of HRQOL instruments can be compared better [7,8]. The ICF was approved in May 2001 by the World Health Assembly and is the successor of the International Classification of Impairments Disabilities and Handicaps (ICIDH) [9]. Using established linkage rules [8], items from specific instruments can be linked to the best corresponding ICF categories, and the representation of the ICF components *body functions and structures*, *activities and participation* and *contextual factors* can be examined.

The ICF is being increasingly applied in clinical research and clinical practice, like rehabilitation medicine [10]. ICF-based instruments such as the WHODASII [4], The Burden of Stroke Scale (The BOSS) [11], and the London handicap scale (LHS) [12] will be used in the future in clinical and epidemiological trials, as well as in health reports. These instruments have been developed guided by the theoretical World Health Organization (WHO) framework of functioning and disability, which is the basis of the ICF. Since HRQOL and ICF represent two different perspectives from which to look at functioning and health, it is expected that both will often be used concurrently in clinical practice, research and health reporting.

The objective of our study was to examine the relationship between the most frequently used short HRQOL instruments and the ICF.

The specific aims were to: (1) examine whether the content of the HRQOL instruments is represented by the ICF categories and whether the ICF can serve as the common framework when comparing HRQOL instruments; (2) to identify areas of the ICF not covered at all or not covered in enough detail or precision to be considered in a future revision of the ICF, and, finally; (3) to examine the differences in the contents covered by the HRQOL instruments based on the linkage of their content to the ICF.

Methods

Instruments

Since it would be neither feasible nor useful to compare all HRQOL instruments developed in the last decades, we decided to include the two currently most widely used, short generic health profiles in our comparison. These are the SF-36 [2] and the NHP [3]. We also examined Spitzer's Quality of Life Index (QL-I) [13], one of the earliest scales to be designated a 'quality-of-life' measurement, as well as the two instruments developed by the WHO, the 26-items version of the World Health Organization Quality of Life Assessment (WHO-QOL-BREF) [14] and the WHODASII [4]. We also included the EQ-5D [5] which is probably the currently most used HRQOL index. We did not include the new HRQOL index based on the SF-36, the SF-6D, since they contain the same items. All instruments have fewer than 40 items. Since all instruments are well known and extended descriptions are available, the descriptions presented here are brief. The dimensions covered, the mode of administration, length, and time needed to complete are summarized in Table 1.

Table 1. Generic health status measures

Instrument	Dimensions covered	Mode of administration	Number of Items	Time to complete
European Quality of life Instrument (EQ-5D) [5]	Mobility Self-care Usual activities Pain/discomfort Anxiety/depression	Self-administered	5 + VAS–General Health	~5 min
Medical Outcome Study Short Form 36 (SF-36) [2]	Physical health, Mental health Social functioning Role functioning General health Vitality Pain	Self-administered Interview telephone interview	36	~10 min
Nottingham Health Profile (NHP) [3]	Physical mobility Pain Emotional reaction Energy level Sleep Social interaction	Self-administered	38	~10–15 min
Quality of life Index (Ql-I) [13]	Activity Daily living Health Support Outlook	5 +VAS-QoL	6	~5 min
WHO Disability Assessment Schedule (WHODASII) [4]	Understanding and communicating Getting around Self care Getting along with people Life activities Participation in society	Self-administered, interview proxy	36	~10 min
WHOQOL-BREF [14]	Physical health Psychological Social relationships Environment	Interview Self-administered	26	~10 min

The QL-I was developed for use by physicians in patients with cancer and chronically-ill patients. It is widely used by cancer specialists interested in the health status and quality of life of their patients before and after therapy. QL-I is one of the earliest scales to be designated a 'quality-of-life' measurement [13]. The QL-I consists of five items with three options for replies. The item responses are scored 0, 1 or 2, giving an overall score of 0 to 10. The scale can be summed into a single score, or each item can be calculated separately. Each item represents a different domain of life functioning. The respondents choose the item in terms of the statement's applicability to themselves: activity, daily living, health, support, and outlook. The scale also comprises a visual-analogue rating scale to rate quality of life.

The WHOQOL-BREF was developed by the World Health Organization Quality of Life (WHOQOL) Group and represents an abbreviated version of the WHOQOL-100, which was developed in collaboration with 15 international field centers simultaneously in an attempt to develop a quality-of-life assessment that would be applicable cross-culturally [14]. The WHOQOL-BREF contains a total of 26 questions. It includes one item

for general quality of life, one item for health-related quality of life, and 24 items belonging to four domains (physical capacity, psychological, social relationships, and environment). In constructing the WHOQOL-BREF, data were taken from 20 field stations located in 18 countries [14].

The WHODASII is a new 36-item instrument developed to assess activity limitations and participation restrictions in six domains (understanding and communicating, getting around, self care, getting along with people, life activities, and participation in society). The WHODASII is different from the other measures of health status in that it is based on an international classification system, i.e. the ICF [6].

The NHP has two parts. Part I contains 38 items grouped into six sections: physical abilities, pain, sleep, social isolation, emotional reactions, and energy level. Part II provides a brief indicator of handicap and contains seven items that record the effect of health problems on occupation, jobs around the house, personal relationships, social life, sex life, hobbies, and holidays. Part II is optional, as some items, like work and sex life, may not be applicable. Yes/no responses are used throughout [3]. The NHP is one of the more frequently used measures, especially in Europe and continues to be used and tested.

The SF-36 derives from a larger battery of questions administered in the Medical Outcomes Study [2]. The SF-36 includes eight multi-item scales containing 2–10 items each and a single item to assess health transition. The scales cover the dimensions of physical health, mental health, social functioning, role functioning, general health, pain, and vitality. Two summary scales can also be obtained – the Physical Component Summary Score (PCS) and the Mental Component Summary Score (MCS) [2]. The SF-36 is the most widely-used general health-status instrument.

The EQ-5D is a brief, self-administered, two-page questionnaire. The first page contains five items describing health status across five dimensions (mobility, self-care, usual activity, pain/distress, and depression/anxiety) (the EQ-5D). The second page has a visual analogue rating scale on which the respondent marks an assessment of his/her overall health [5]. Each dimension is divided into three levels which, when taken together, define a total of 243 (3^5) unique health states. The responses to the five items in the EQ-5D can be scored using a utility-weighted algorithm [15], which has been recommended for use in economic evaluation. The EQ-5D, therefore, provides two single-index measures of health, the rating scale, and the EQ-5D index, ranging from 0 to 100 [16].

The ICF is a multipurpose classification belonging to the WHO family of international classifications. The ICF was designed to record and organize a wide range of information about health and health-related states in a standardized, common language, thereby facilitating communication about health and health care in various disciplines and scientific fields worldwide. It makes the comparison of data across countries, healthcare disciplines, services, and time possible.

The ICF has two parts, each containing two separate components. Part 1 covers FUNCTIONING and DISABILITY and includes the components: *body functions* (b) and *structures* (s) and *activities and participation* (d). Part 2 covers CONTEXTUAL FACTORS and includes the components: *environmental factors* (e) and *personal factors*.

Body functions are the physiological and psychological functions of body systems. *Body structures* represent the anatomical parts of the body, such as organs, limbs and their components. *Activities and participation* are given in the ICF in a single list that covers the full range of life areas, from basic learning or watching, to composite areas such as interpersonal interaction or employment. *Environmental factors* consist of the physical, social and attitudinal environment in which people live and conduct their lives. *Personal factors* are the particular background of an individual's life and living, and comprise features of the individual that are not part of the health condition or health states. These factors may include gender, race, age, fitness, lifestyle and habits. The *personal factors* are not as yet classified in the ICF.

In the ICF classification, the letters b, s, d, and e, which refer to the components of the classification, are followed by a numeric code starting with the chapter number (one digit), followed by the second level (two digits) and the third and fourth levels (one digit each). The component letter with the suffix of two, three, or four digits corresponds

to the code of the so-called categories. Categories are the units of the ICF classification. Within each chapter, there are individual two-, three- or four-level categories. An example selected from the component *body functions* is presented in the following:

b2 Sensory functions and pain (first level)
b280 Sensation of pain (second level)
b2801 Pain in body part (third level)
b28013 Pain in back (fourth level).

Within each component, the categories are arranged in a stem/branch/leaf scheme. Consequently, a lower-level category shares the attributes of the higher-level categories to which it belongs, i.e., the use of a lower-level (more detailed level) category automatically implies that the higher-level category is applicable, but not the other way around.

At the end of each embedded set of third- or fourth-level categories and at the end of each chapter, there are 'other specified' categories (uniquely identified by the final code 8). These categories allow the coding of aspects that are not included in any other specific categories. For example, at the end of the set of the fourth-level categories embedded in the third-level category *b2801 Pain in body part*, one finds the category *b28018 Pain in body part, other specified.*

Linkage of items to the ICF

We applied linking rules which were developed to link the concepts contained in health-status measures to the ICF [7]. According to these rules, experts trained in the ICF are advised to link concepts contained in a health-status measure to the ICF category representing this concept most precisely. If an item of a measure contains more than one concept, each concept has to be linked separately. For example, in item 8 of the SF-36 'During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)', the concept 'pain', 'work outside the home' and 'housework' are linked to *b280 sensation of pain, d859 work and employment, other specified and unspecified* and *d640 doing house work*, respectively. In item 3 of the SF-36 '… Does your health now limit you in these activities?…,' which contains 10 different questions referring to different activities, the concept 'health' and all other concepts contained in the respective 10 questions are identified and linked to the ICF.

For the sake of comprehensibility of the results of this study rule 5, 8, 9, and 10 require special annotation. Rule 5 points out that the response options of an item are linked to the ICF if they refer to other than the concepts contained in the corresponding item. For example in the item 'self care' of the EQ-5D, which contains the response categories 'I have no problems with self-care', 'I have some problems washing or dressing myself' and 'I am unable to wash or dress myself', not only 'self-care' but also the concepts 'washing myself' and 'dressing myself' are linked.

According to rule 8, if the content of an item/concept is specified by examples, the latter will be linked to the so-called 'unspecified'-ICF categories, which are uniquely identified by the final code 9. For example, in the item 'usual activities (e.g. work, study, housework family or leisure activities)' of the EQ-5D the concepts 'work', 'study', 'housework' 'family' or 'leisure activities' are linked to the corresponding 'unspecified'-ICF categories in addition to the linkage of the concept 'usual activities'. E.g. work is linked to the ICF category *d8509 remunerative employment, unspecified.*

Rule 9 states that if the information provided by the item/concept is not sufficient to decide which ICF category should be chosen, this item/concept should be labelled nd (not definable). The abbreviation nd-gh (not definable-general health) is used for items/concepts concerning health in general, and the abbreviation nd-qol (not definable-quality of life) is used for items/concepts concerning the quality of life of patients in general. According to rule 10, if an item/concept is not contained in the ICF classification, then this item/concept is labelled 'nc' (not covered by the ICF). 'nc' does not differentiate between concepts relating to functioning not covered by the ICF, concepts relating to personal factors for which no categories currently exist, and other concepts relating e.g. to time and space.

Consensus between two health professionals was used to decide which ICF category should be linked to each item/concept of the questionnaires. To resolve disagreements between them concerning the selected categories, a third person trained in the

linking rules was consulted. In a discussion led by the third person, the two health professionals that linked the item stated their pros and cons for the linking of the concept under consideration to a specific ICF category. Based on these statements, the third person made an informed decision.

The degree of agreement between the two health professionals at the component, 1st, 2nd and 3rd ICF's levels was calculated by means of the kappa statistic, which is a measure of the agreement that exists beyond the amount of that expected by chance alone [17]. Values of kappa generally range from 0 to 1, whereas 1 indicates perfect agreement and 0 indicates no additional agreement beyond what is expected by chance alone. A negative value of kappa indicates agreement, which is less than that expected by chance alone.

Kappa by definition is bounded by 1, i.e. its sampling distribution becomes progressively skewed to the left as kappa approaches 1. Since the asymptotic confidence interval does not take this skewness into account, especially with small sample sizes, and can produce upper confidence limits that exceed 1, bootstrapped intervals, which are produced by percentiles of samples based on the observed data, were calculated [18].

The data analysis was performed with SAS for windows V8.

Results

Linkage process

In the 148 items of the 6 instruments a total of 226 concepts were identified and linked to the ICF. The results of the kappa statistic as well as the bootstrapped confidence intervals at the component, 1st, 2nd and 3rd ICF's levels are presented in Table 2.

Table 2. Estimated kappa coefficient and the bootstrapped confidence intervals at the component, 1st, 2nd and 3rd ICF's levels

	Estimated Kappa Coefficient	95% Bootstrapped Confidence Intervals
Component	0.89	(0.84, 0.93)
Chapter 1st Level	0.98	(0.95, 0.99)
2nd Level	0.87	(0.82, 0.91)
3rd Level	0.82	(0.59, 0.90)

The estimated kappa coefficients range between 0.82 at the 3rd ICF's level and 0.98 at the chapter or first ICF's level.

The width of the 95% bootstrapped confidence interval, which indicates the precision of the estimated kappa coefficient is narrowest at the chapter level and widest at the 3rd ICF's level.

The concepts contained in the items of the HRQOL instruments were linked to 91 different ICF categories. The concepts were linked to 17 categories of the component *body functions*, to 60 categories of the component *activities and participation*, and to 14 categories of the component *environmental factors*. No concepts were linked to the component *body structures*. Twelve concepts could not be linked to the ICF and were coded nc (see Table 3). These concepts seem to be *personal factors* encompassing individual characteristics, such as self-perception, perception of others (i.e., 'I seem to get sick a little easier than other people' and 'I am a burden to people') and perception of life (i.e., '…life is not worth living' and 'How much do you enjoy life?'). Also, terms referring to a time, such as 'In the past 30 days, how many days were these difficulties present?' could not be linked to the ICF.

It is important to emphasize that 14 of the 91 ICF categories are 'unspecified-ICF categories' indicating that the concepts that were linked to them were examples of more general and unspecific concepts.

Table 3. Items/concepts assigned to nc (not covered by the ICF)

SF-36
I seem to get sick a little easier than other people.
I am as healthy as anybody I know.
I expect my health to get worse.
NHP
… life is not worth living.
… if I'm losing control
… I am a burden to people.
QLI
…accepting and in control of personal circumstances, including surroundings
WHODASII
Overall, in the past 30 days, how many days were these difficulties present?
Overall, how much did these difficulties interfere with your life?
WHOQOL-BREF
How much do you enjoy life?
How safe do you feel in your daily life?
How satisfied are you with …?

For 19 concepts nd, nd-gh and nd-qol was used (see Table 4).

Linkage results

Tables 5–7 show the comparison of the concepts contained in the items of the HRQOL instruments using the ICF categories as a reference and ordered by component, i.e., *body functions, activities and participation*, and *environmental factors*. The numbers contained in the tables represent the frequency with which the ICF categories were addressed in the different instruments. Generally, the ICF categories were linked to just one item/concept of a HRQOL instrument as it is indicated by a '1' in the tables. A higher number indicates that the ICF did not differentiate in greater detail, and, therefore, several items or concepts of items from a specific health-status instrument had to be linked to the same ICF category. For example, for the SF-36, the ICF category *b152 Emotional functions* was chosen to link a number of different feelings: 'feeling depressed or anxious', 'emotional problems', 'very nervous', 'I felt so down in the dumps nothing could cheer me up', 'I felt calm and peaceful', 'I felt downhearted and blue', and 'Have you been a happy person?'. If there were different categories for different feelings, the named items would have been linked to different categories.

Representation of body functions (Table 5)

Emotional functions are covered by all examined instruments. They are covered in more detail in the NHP and the SF-36 than, for example, in the WHODASII. *Energy level* (b1300), which may include positive and negative aspects, such as fatigue or energy, is represented in the QL-I, the WHOQOL-BREF, the NHP and the SF-36, but not in the WHODASII or the EQ-5D. *Sleep functions* are only covered by the WHOQOL-BREF and the NHP. The NHP differentiates more than the WHOQOL-BREF. The NHP distinguishes among onset of sleep, maintenance of sleep, and quality of sleep. *Pain* is covered in more detail in the NHP than in the SF-36 and the

Table 4. Items/concepts assigned to nd (not definable), nd-gh (not definable-general health) or to nd-qol (not definable-quality of life)

nd	nd-gh	nd-qol
EQ-5D		
	Compared with my general level of health over the past 12 months, my health state today is	
SF-36		
Your health limits you a little in moderate activities	In general, would you say your health is: (excellent, very good, good, fair, poor)	
Your health does not limit you in vigorous activities	Compared to one week ago, how would you rate your health in general now?	
	Does your health now limit you in these activities?	
	I expect my health to get worse.	
	My health is excellent.	
	... as a result of your physical health?	
	Has your physical health ..	
QLI		
	– has been appearing to feel well or reporting feeling 'great' most of the time	
	– has been feeling very ill or 'lousy'...	
WHODASII		
Staying by yourself for a few days?		
Getting all the work done that you need to do?		
WHOQOL-BREF		
How available to you is the information that you need your day-to-day life?	How satisfied are you with your health?	How would you rate your quality of life?
How healthy is your physical environment?		
… prevents you from doing what you need to do?		

Table 5. General Health Status Instruments – frequencies showing how often the *body-function* categories were addressed in the different instruments

ICF Category	Ql-I Spitzer	WHOQOL-CHEF	WHODASII	NHP	SF-36	EQ-5D
b110 Consciousness functions	1					
b1100 State of consciousness	3					
b1263 Psychic stability	1					
b1266 Confidence	1					
b1300 Energy level	2	1		3	4	
b134 Sleep functions		1		1		
b1341 Onset of sleep				1		
b1342 Maintenance of sleep				3		
b1343 Quality of sleep				1		
b1400 Sustaining attention		1	1			
b144 Memory functions			1			
b152 Emotional functions	4	1	1	8	7	2
b1529 Emotional functions, unspecified		1			1	
b1801 Body image		1				
b280 Sensation of pain		1		8	2	1
b289 Sensation of pain, other specified and unspecified						1
b450 Additional respiratory functions				1		

Table 6. General Health Status Instruments – frequencies showing how often the *activities-and-participation* categories were addressed in the different instruments

ICF Category	Ql-I Spitzer	WHOQOL-CHEF	WHODASII	NHP	SF-36	EQ-5D
d1 Learning and applying knowledge			1			
d175 Solving problems			1			
d2100 Undertaking a simple task	2					
d230 Carrying out daily routine		2	2		2	1
d310 Communicating with – receiving – spoken messages			1			
d3500 Starting a conversation			1			
d3501 Sustaining a conversation			1			
d4 Mobility						1
d410 Changing basic body position				1		
d4102 Kneeling					1	
d4104 Standing			1			
d4105 Bending				1	1	
d4153 Maintaining a sitting position				1		
d4154 Maintaining a standing position			1	2		
d430 Lifting and carrying objects					1	
d4309 Lifting and carrying, unspecified					1	
d4452 Reaching				1		
d4459 Hand and arm use, unspecified					1	
d450 Walking				1		1
d4500 Walking short distances					1	
d4501 Walking long distances			1		2	
d455 Moving around				2		
d4551 Climbing					2	
d4559 Moving around, unspecified					1	
d460 Moving around in different locations				1		
d4600 Moving around within the home			1			
d4702 Using public motorized transportation	1					
d4751 Driving motorized vehicles	1					
d498 Mobility, other specified	1		1			1
d499 Mobility, unspecified		1				

Table 6. *Continued*

ICF Category	Ql-I Spitzer	WHOQOL-CHEF	WHODASII	NHP	SF-36	EQ-5D
d5 Chapter 5 self-care	1					1
d510 Washing oneself	1					1
d5101 Washing whole body			1		1	
d530 Toileting	1					
d540 Dressing	1		1	1	1	1
d550 Eating	1		1			
d6309 Preparing meals, unspecified				1		
d640 Doing housework			1			
d6409 Doing housework, unspecified	1			1	2	1
d649 Household tasks, other specified and unspecified			2			
d6509 Caring for household objects, unspecified				1		
d7 Chapter 7 interpersonal interactions and relationships	1	1	1	2		
d7500 Informal relationships with friends			2			
d7609 Family relationships, unspecified						1
d770 Intimate relationships		1	1			
d7702 Sexual relationships			1	1		
d779 Particular interpersonal relationships, other specified and unspecified				1		
d830 Higher education	1					
d839 Education, other specified and unspecified			2			1
d850 Remunerative employment		1		1	1	
d8502 Full-time employment	1					
d855 Non-remunerative employment	1					
d859 Work and employment, other specified and unspecified	3		2		2	
d8509 Remunerative employment, unspecified						1
d870 Economic self-sufficiency			1			
d9109 Community life, unspecified			1			
d920 Recreation and leisure		1	1	1		
d9204 Hobbies				1		
d9205 Socializing					2	
d9209 Recreation and leisure, unspecified				2	2	1

EQ-5D. It is not included in the WHODASII or in the QL-I.

Representation of activities and participation (Table 6)

Carrying out daily routine is not contained in the QL-I, the NHP, or in the EQ-5D while it is covered by the other instruments. Aspects of *mobility* are scarcely represented with the exception of the NHP and the SF-36. The EQ-5D covers *mobility* in general and *walking* and the WHODASII only covers *walking long distances*. With respect to *self-care*, all instruments, with the exception of the WHOQOL-BREF, cover *dressing*. The QL-I and the WHODASII more broadly cover aspects of self care. At least one of the categories included in the chapter *domestic life* (d6) is covered in all instruments with the exception of the WHOQOL-BREF. However, the WHODASII, the NHP, and the SF-36 cover this area into more detail than the other instruments. *Housework* is contained in the WHODASII, the NHP, and the SF-36. 'Housework' is named as an example in Item 3 of the EQ-5D 'Usual activities (e.g. work, study, housework, family or leisure activities)'. According to linking rule 8, that concept was linked to *d6409 Doing housework, unspecified.*

The QL-I, WHOQOL-BREF, WHODASII, and NHP contain *interpersonal interactions and relationships* to some degree. Accordingly, the

Table 7. General Health Status Instruments – frequencies showing how often the *environmental-factors* categories were addressed in the different instruments

ICF Category	Ql-I	Spitzer	WHOQOL-CHEF	WHODASII	NHP	SF-36	EQ-5D
e1101 Drugs			1		1		
e115 Products and technology for personal use in daily living	1						
e120 Products and technology for personal indoor and outdoor mobility and transportation	1						
e1209 Products and technology for personal indoor and outdoor mobility and transportation, unspecified					1		
e155 Design, construction and building products and technology of buildings for private use			1				
e1650 Financial assets			1				
e3 Chapter 3 support and relationships	1		1				
e310 Immediate family	1						
e315 Extended family	1						
e320 Friends	1		1				
e399 Support and relationships, unspecified					1		
e4 Chapter 4 attitudes				1			
e5400 Transportation services			1				
e5800 Health services			1				

corresponding items/concepts of the four instruments were linked at the chapter level to the ICF. The WHODASII additionally addresses *relationships with friends* and *intimate relationships*. The WHODASII and WHOQOL-BREF also cover *sexual relationships*. The concept 'family' is contained in the item 3 of the EQ-5D as an example and was therefore linked to the category *d7609 family relationships, unspecified*. In this context it is important to mention that item 6 and 10 of the SF-36 'social activities with family, friends, neighbours, or groups', were linked to *d9205 socializing* which is defined in the ICF as 'engaging in informal or casual gatherings with others, such as visiting friends or relatives or meeting informally in public places'.

All instruments cover aspects of *work*, but they are covered in more detail by the QL-I. With the exception of the QL-I *recreation and leisure* (d920 or d9209) is addressed in all instruments. Within the chapter *community, social and civic life* (d9) the SF-36 additionally covers, as mentioned before, the category *socializing* (d9250) and the NHP covers *hobbies* (d9204).

Aspects only covered by the WHODASII include *solving problems, understanding and communication* e.g. 'starting a conversation' and 'sustaining a conversation'.

Representation of environmental factors (Table 7)

Four instruments, the QL-I, the WHOQOL-BREF, WHODASII and the NHP address *environmental factors*, whereby the QL-I and the WHOQOL-BREF cover *environmental factors* more in detail than the NHP and the WHODASII. The WHODASII addresses the chapter *attitudes* at a very general level. The WHOQOL-BREF and the NHP cover the ICF category *drugs* (e1101). The QL-I especially covers categories within the chapter *support and relationships* (e3), i.e. the support and relationships of immediate and extended family and friends. The WHOQOL-BREF also contains *support and relationships* at a general level (e3), as well as at a more specific level, by addressing the *support and relationships provided by friends* (e320).

Discussion

The ICF proved highly useful for the comparison of HRQOL instruments. With few exceptions, the content of the HRQOL instruments was represented by the ICF categories and therefore the ICF can serve as the common framework when comparing HRQOL instruments.

Based on the linkage, it was possible to study the heterogeneity of HRQOL instruments regarding their representation of *body functions, activities, participations* and *environmental factors*. The comparison of selected HRQOL instruments may provide clinicians and researchers with new insights when selecting health-status measures for clinical studies [19].

Nevertheless, it is important to emphasize that the results of the content comparison of HRQOL measures based on the ICF can never substitute a thorough study of the instruments under consideration. The tables summarising the results of the content comparison, like the tables presented in this paper, do not provide information about the number of items contained in the instruments, the kind of response categories applied or the psychometric properties of the instruments.

The first question when selecting HRQOL instruments is to decide what should be measured in consideration of the study endpoints, the population studied, and the intervention. The second question is to decide which instrument to use amongst all the possible instruments available. Since the ICF-based comparison provides information about the contents addressed in the different instruments, it can be a very useful tool at this stage of the selection process. Further considerations, such as practicability and length of the instrument, response categories and psychometric characteristics have to be then taken into account.

The comparison based on the ICF not only provides insight on the bandwidth of the different generic instruments, i.e., the breadth of health dimensions measured, but also on the precision of the instruments, i.e., the thoroughness and depth of measurement. In this context, precision is conceptualized as a property of a measure that encompasses both the range or depth of measurement and the number of distinct levels enumerated by a scale (fineness of specification) [20, 21]. As many generic tools lack the precision required for effective health-care decision making, the use of the ICF as a reference tool for comparison can be of special importance to differentiate the precision among different instruments. For example, the level of differentiation of the NHP with regard to sleep, covering onset of sleep, maintenance of sleep, and quality of sleep can be of relevance in certain kinds of studies. When sleep is assessed by a condition-specific instrument in a determined study, it may not be of relevance whether this area is covered by the generic instrument or not. Consequently, other instruments without sleep could be selected as study outcomes. When selecting generic and condition specific instruments it would be worthwhile to examine the overlap of the HRQOL instruments with the condition and symptom-specific instruments. Such a comparison is possible through linkage of the items to the ICF. Ideally, there would be only a minimal overlap of condition and symptom-specific instruments to minimize the number of questions that need to be answered.

From the comparison, it also becomes clear that instruments not only differ with respect to the bandwidth and precision, but also on the level of differentiation regarding performance. An example is the SF-36, which covers both *walking short distances* and *walking long distances*. This is not a differentiation of precision, but one of performance of the instrument and this is an issue of measurement rather than of content. The obvious reason why such differentiation is often necessary in measurement is to avoid floor and ceiling effects or, in general, to cover the whole spectrum of performance encountered in a number of populations – young and old, healthy and ill people.

The comparison of instruments based on the ICF also illustrates that *contextual* and *environmental factors* are scarcely represented in the six selected instruments. This fact raises the question whether an addition should be made to these instruments. For example, attitudes of the family, significant others, support at work etc., are now recognized to be among the most important factors relating to life satisfaction, work performance and disability [22–27].

The linkage also revealed that different categories are addressed very frequently in the different instruments, which could mean that ICF does not differentiate enough in some categories, like the category b152, *emotional functions*. In a review of the items of the SF-36, and the NHP, the most common emotional functions that could be specified in a future version of the ICF are: sadness, happiness, anxiety, and anger.

The degree of agreement between health professionals is high according to the Kappa coefficient and the lower limits of confidence intervals which exceed 0.5. Nevertheless, estimation of the reliability of the linking process could be strengthened by augmenting the number of raters in future studies. A potential limitation in this study is the professional background of the two raters both being psychologists. It should be studied whether raters trained similarly in the ICF but coming from different professional backgrounds apply the linkage rules differently.

Most of the concepts contained in items which could not be linked to the ICF and, accordingly, were assigned *nc* (not covered) seem to be *personal factors*, which have not yet been classified in the ICF. Therefore, the linkage of HRQOL instruments to the ICF could also reveal a number of relevant *personal factors* that should be added to the ICF, if a classification of *personal factors* is intended in the future.

The use of the abbreviations *nd-gh* and *nd-qol* to summarize the content of determined concepts helps to overcome the difficulty encountered when attempting to link items which refer to health or to quality of life in general to the ICF. These abbreviations do not denote that something is missing in the ICF, but that these kinds of items/concepts refer to all categories contained in the ICF and to none concretely.

The abbreviation *nd* is used when the content of a determined item is not specific enough to decide to which ICF category it belongs. Therefore, the frequency with which the abbreviation *nd* is used could be interpreted as the frequency with which the linkage exercise was unsuccessful. The fact that the abbreviation *nd* was only applied to 19 of 226 concepts contained in 148 items can be interpreted as a positive result in this study.

In conclusion, the ICF provides an excellent framework when comparing the content of HRQOL instruments. The comparison of the instruments provides interesting insights into their differences with respect to the breadth and precision of their coverage of specific concepts. The link of HRQOL instruments to the ICF could also provide information on which ICF categories require more specification. This information could prove highly relevant for a future revision of the ICF.

References

1. Spilker B. Introduction. In: Spilker B. (ed), Quality of Life and Pharmacoeconomics in Clinical trials, 2nd ed, Lippincott-Raven Publishers, 1996; 1–10.
2. Ware JE and Sherbourne CD. The MOS 36-item short-form health survey (SF-36). A. Conceptual framework and item selection. Medical Care 1992; 30: 473–483.
3. Hunt SM, McEwen J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. J R Coll Gen Pract 1985; 35(273): 185–188.
4. World Health Organization. World Health Organization Disability Assessment Schedule (WHODASII). Training Manual: a guide to administration. Geneva: WHO (2000) Retrieved April 9, 2003, from < http://www.who.int/icidh/whodas/ > .
5. The Euroqol Group. Euroqol-a facility for the measurement of health-related quality of life. Health Policy 1990; 16: 199–208.
6. World Health Organization. International Classification of Functioning, Disability and Health: ICF. Geneva: WHO, 2001.
7. Stucki G, Cieza A, Ewert T, Kostanjsek N, Chatterji S, Uestuen TB. Application of the International Classification of Functioning, Disability and Health (ICF) in clinical practice. Disability and Rehabilitation 2002; 24(5): 281–282.
8. Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, Bedirhan Üstün B, Stucki G. Linking Health-Status measurements to the International Classification of Functioning, Disability and Health. J Rehabil Med 2002; 34: 1–6.
9. World Health Organization. International Classification of Impairments, Disabilities, and Handicaps: ICIDH. Geneva: WHO, 1980.
10. Stucki G, Ewert T, Cieza A. Value and application of the ICF in rehabilitation medicine. Disability and Rehabili tation 2002; 24(17): 932–938.
11. Doyle PJ. Measuring health outcomes in stroke survivors. Arch Phys Med Rehabil 2002 Dec; 83(12 Suppl 2): S39–S43.
12. Harwood RH, Gompertz P, Ebrahim S. Handicap one year after a stroke: validity of a new scale. J Neurol Neurosurg Psychiatry 1994 Jul; 57(7): 825–829.
13. Spitzer WO, Dobson AJ, Hall J, et al. Measuring the quality of life of cancer patients: A concise QL-index for use by physicians. J Chronic Dis 1981; 34(12): 585–597.
14. The WHOQOL Group The development of the World Health Organization quality of life assessment instrument (the WHOQOL). In: Orley J and Kuyken W (eds), Quality of Life Assessment: International Perspectives. Heidelberg: Springer Verlag, 1994.
15. Williams A. The measurement and valuation of health: a chronicle. Discussion Paper 136. Centre for Health Economics, York Health Economics Consortium, NHS Centre for Reviews and Dissemination. York: University of York, 1995.
16. Brook R with the EuroQol group. EuroQol: the current state of play. Health Policy 1996; 37: 53–72; Dolan P.

Modeling valuations for EuroQol health states. Med Care 1997; 35: 1095–108.
17. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960; 20: 37–46.
18. Vierkant RA (2000). A SAS macro for calculating bootstrapped confidence intervals about a kappa coefficient. SAS Users Group International Online Proceedings. Retrieved July 23, 2004, from <http://www2.sas.com/proceedings/sugi22/STATS/PAPER295.PDF>.
19. Weigl M, Cieza A, Harder M, Geyh S, Amann E, Kostanjsek N, Stucki G. Linking osteoarthritis-specific health-status measures to the International Classification of Functioning, Disability, and Health (ICF). Osteoarthritis Cartilage 2003; 11(7): 519–523.
20. Kessler RC, Mroczek DK. Measuring the effects of medical interventions. Med Care 1995; 33(4): 109–119.
21. McHorney CA. Generic health measurement: Past accomplishments and a measurement paradigm for the 21st century. Ann Intern Med 1997; 15(27): 743–750.
22. Waltz M, Kriegel W, van't Pad Bosch P. The social environment and health in rheumatoid arthritis: Marital quality predicts individual variability in pain severity. Arthritis Care Res 1998; 11: 356–374.
23. Kirwan J, Heiberg T, Hewlett S, et al. Outcomes from the Patient Perspective Workshop at OMERACT 6. J Rheumatol 2003; 30(4): 868–872.
24. Johansson U, Bernspang B. Life satisfaction related to work re-entry after brain injury: A longitudinal study. Brain Inj 2003; 17(11): 991–1002.
25. Vestling M, Tufvesson B, Iwarsson S. Indicators for return to work after stroke and the importance of work for subjective well-being and life satisfaction. J Rehabil Med 2003; 35(3): 127–131.
26. Avlund K, Lund R, Holstein BE, Due P. Social relations as determinant of onset of disability in aging. Arch Gerontol Geriatr 2004; 38(1): 85–99.
27. Travis LA, Lyness JM, Shields CG, King DA, Cox C. Social support, depression, and functional disability in older adult primary-care patients. Am J Geriatr Psychiatry 2004; 12(3): 265–271.

Address for correspondence: Gerold Stucki, Department of Physical Medicine and Rehabilitation, University of Munich, Marchioninistr. 15, 81377 Munich, Germany
Phone: +49 89 7095 4051; Fax: +49 89 7095 8836
E-mail: gstucki@med.uni-muenchen.de