

**Reykjavik, Iceland
24-30 October, 2004**

Title: **Operationalizing ICF for Measurement: Calibration, Qualifier, Instruments**

Authors: **Nenad Kostanjsek, Bedirhan Üstün**
Classification, Assessment and Terminology Team (CAT)
World Health Organization

Purpose: for discussion & decision

Recommendations:

The WHO FIC Network shall adopt the proposed framework and actions for operationalizing the ICF for measurement.

Abstract:

ICF already is a much more operational framework. It has categories with definitions which are comprehensive, specific, reliable and cross-cultural applicable. ICF is using a qualifier scale with refined anchor points which will be linked with ICF based population norms from more than 70 countries. However further operationalization of the ICF for measurement purposes is vital in order to increase and sustain the utility of the classification. Two pathways are proposed for further operationalization of ICF. First, the mapping of existing health status and disability assessment instruments to ICF and secondly, the use of ICF as a homogenous assessment base with application and further development of the WHO DAS II. To synergize ongoing and upcoming activities within the WHO FIC Network the activities are proposed: (i) protocol development for mapping assessment instruments to ICF and WHO DAS II further development; (ii) establish a portal on the WHO FIC website; (iii) development and field testing of maps and WHO questions on impairments and environmental factors.

I. Introduction

The ICF allows to qualify the magnitude of the functioning problem (impairment, decrement in capacity and/or performance) and the extent to which an environmental factor is a facilitator or barrier. The severity scale of the qualifiers distinguishes between no, mild, moderate, severe, complete, not specified and not applicable. Generic indicators/synonym and broad percentages are given for all qualifiers in their gradation. For example:

- the term “mild” is generally used to indicate a slight, low degree of problem, which means a problem that is present less than 25% of the time, with a slight alternation in functioning, which happens rarely over the last 30 days.
- the term “moderate” is generally used to indicate a medium, fair degree of problem, which means that a problem that is present less than 50% of the time, with a medium alternation in functioning, which happens occasionally over the last 30 days.
- the term “severe” is generally used to indicate high, extreme degree of problem, which means a problem that is present more than 50% of the time, high alternation in functioning, which happens frequently over the last 30 days.
- the term “complete” is generally used to indicate a total degree of problem, which means complete means that a problem that is present more than 95% of the time, total alternation in functioning, which happens every day over the last 30 days.

The rationale for the percentage ranges which are assigned to the qualifiers is as follows. Category 1 (No Problem) and Category 4 (Complete Problem) include a percentage margin of 5% to prevent measurement errors in the next higher or lower categories. For example a person who reports that during the last month he/she had one day where he/she experienced twice a minor and momentary difficulty in walking could be considered as having no problem, especially in case of self-reported difficulty where detailed cross examination is necessary to uncover the momentary problem.

A person within Category 3 with over 50% either in terms frequency, intensity or duration of the functioning problem is recognized as meeting the clinical threshold and requires some sort of health or health related intervention.

Because of the uncertainty about what is clinical significant below 50% and in view that the majority of the population is expected to be below 50%, the percentile range from 5%-49% was divided into Category 1 (Mild Problem 5%-24%) and Category 2 (Moderate problem 25%-49%). In Category 2 we expect to have a mixture of people, who are either above or below the clinical threshold. In Category 1 we expect only people who are below the clinical threshold.

It is important to note that the percentages ranges assigned to the qualifiers are to be calibrated in different domains with reference to population standards as percentiles. For this quantification to be used in a uniform manner, ICF based population norms and cross-

mapping with assessment instruments need to be developed. Until then ICF users have to use clinical judgement while using the qualifier scale.

II. Pathways for further operationalization of ICF for measurement

Compared to its predecessor ICIDH, the ICF already is a much more operational framework. It has categories with definitions which are comprehensive, specific, reliable and cross-cultural applicable. ICF is using a qualifier scale with refined anchor points which will be linked with ICF based population norms from more than 70 countries. However further operationalization of the ICF for measurement purposes is vital in order to increase and sustain the utility of the classification. Two pathways are proposed for further operationalization of ICF.

1. *ICF as the standard coding system for the different measures commonly used in the health and disability field*

The concept of measuring functioning, disability or health is not new. There are hundreds of well-established assessment tools. Mostly clinicians in different specialties have developed condition specific assessment tools (e.g. Arthritis Impact Measurement Scale, AIMS 2; Hamilton Rating Scale of Depression, HAMD; McGill Pain Assessment Questionnaire, MPQ; Outcome Measures in Rheumatology Clinical Trials, OMERACT). There are also some generic measures (SF 36, Nottingham Health Profile, EuroQol-5D. These measures have proven useful to track outcomes, but they are neither comprehensive nor do they fully map to the ICF. Furthermore, no single measure appears to be commonly used internationally for assessment of functioning and disability and those few which have been developed internationally or used in different cultures have no proven cross-population comparability.

The result, well-known and much criticized, is 'data silos' in which assessment data acquired in one episode of care – emergency, medical, rehabilitative, out-patient and community clinical care -- cannot be carried over to another episode of care involving a different clinical focus. Furthermore, from an international perspective these 'data silos' are problematic in terms of data comparability. To compare outcome data across diseases, interventions, countries and cultures we do need a common framework that will serve as a "Rosetta Stone". The ICF makes it possible to link together these data across conditions or interventions, eliminating the frustrating data silo effect, and making for more efficient, transparent, and cost-effective health care.

2. *ICF as a homogenous assessment base: WHO DAS II, ICF Checklist, ICF Core Sets and others*

A classification needs to be exhaustive by its very nature and becomes very complex for daily use unless it is transformed into practice-friendly tools. For example, a clinician cannot easily take the main volume of ICF and consistently apply it to his or her patients. In daily practice, clinicians will only need a fraction of the categories found in the ICF. As the general rule

goes, 20 % of the codes will explain 80% of the variance observed in practice. With this need in mind, WHO has already created a series of instruments based on the ICF, like the WHO Disability Assessment Schedule II (WHO DAS II) and the ICF Checklist.

The WHO DAS II explores what people do in different areas of life and does not include signs or symptoms of diseases, or feelings of subjective well-being. WHO DAS II was conceived as a general health state assessment measure that can be used for multiple purposes such as epidemiological surveys, clinical use or as a potential description system to contribute to summary measure of population health. It gives a general score as well as different profiles on the following six domains, which were selected after a careful review of existing research and survey instruments and the cross-cultural applicability study:

- Understanding and communicating with the world (cognition)
- Moving and getting around (mobility)
- Self care (attending to one's hygiene, dressing, eating and staying alone)
- Getting along with people (interpersonal interactions)
- Life activities (domestic responsibilities, leisure, and work)
- Participation in society (joining in community activities)

Across these six domains there are thirty-six questions, which directly map to the ICF. All domains and their questions have proven metric qualities in terms of sensitivity, specificity, reliability, validity and comparability shown in general population surveys and in clinical sensitivity to change studies and cost-effectiveness studies. The results of these studies are expected to be published end of 2004 or beginning of 2005.

WHO- DAS II has now been translated into 16 languages and comes in several versions: fully structured self-administrated, 12 and 36 question versions; fully structure interviewer administrated, in 12, 36 and 12 plus 24 item versions; and fully structured proxy versions. The administration of questions takes an average time of 8 minutes and requires only a minimum training of lay interviewers. Training material and tools for data analysis are also available.

The ICF Checklist is a practical translation of the ICF for clinical practice. Items from the classification were chosen by experts to list the most commonly used domains and later field tested to verify the selection and make additions of missing items. The ICF Checklist gives a thumbnail sketch of the main functioning of any individual in terms of body functions and structures; activities and participation and environmental factors.

However, the generic character of the WHO DAS II and ICF Checklist may be a drawback in specialty settings. For example, a clinician dealing with patients with arthritis will need a wider range of categories to identify functions in the neuromusculoskeletal and movement-related area. A speech and language therapist, on the other hand, will require detailed description of voice and speech functions and related structures. Here comes the dilemma: on the one hand we need a "common base" to compare with other health conditions and interventions; on the other hand we need "variability" to capture the detail to describe the profile of a unique group. For such specialized clinical settings, "one (generic) size does not fit all" and the "devil is in the detail".

In response to this dilemma ICF Core Sets are developed. The ICF Core Sets have "common" categories, that will help to address the comparability issue. These common categories are comparable to the generic ICF Checklist., The ICF Core Sets have "additional items" that give a more detailed picture for certain clinical conditions.

II. Proposed methodology for operationalization

1. Cross-walking health status and disability assessment instruments to ICF

Mapping existing health status and disability assessment instruments to ICF comprises two major methodology challenges. First, the cross-walking of assessment instrument domains to ICF domains and secondly the linking of the response scales used in those instruments with the ICF qualifiers. In both cases systematic and standardized approach is needed.

1.1 Domain Mapping

A mapping typology and mapping rules are the essential elements of a linking methodology. Table 1. summarizes the main features of a proposed typology. Type 1 is defined as a one to one full match where one an instrument item with one specific concept is linked with a single ICF domain. Type 2 refers to a one to many full match where an instrument item with many specific concept is linked with a multiple ICF domains. Type 3 is defined as a one to one approximated match where one an instrument item with one non-specific concept is linked with an appropriate single ICF domain. Type 4 is defined as a one to many approximated match where one an instrument item with many non-specific concepts is linked with appropriate multiple ICF domains.

The typology needs to be complemented by rules, which guide the mapping process. Such rule set would include among others the following rules:

- Mapping should be as precise as possible;
- Include concepts mentioned in response options of an item;
- Exclude items which use functioning terms as a synonym for symptom severity;
- Flag instrument items which make causal attributions.

Table 1: Typology for mapping assessment instruments into ICF framework

Type of Match	Assessment Instrument	Example
1. One to one, full match	Item with one specific concept	"Do you have difficulties in walking." (<i>WHO DAS II</i>)
2. One to many, full match	Item with many specific concepts incl. concepts mention in response options	"Does your health now limit you in dressing/bathing yourself?" (<i>SF36</i>)
3. One to one, approx.	Item with one non-specific concept a) not explicitly mentioned in ICF b) more general	
4. One to many, approx.	Item with many non-specific concept (e.g. compound question with category examples)	" In the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives etc.)?" (<i>SF 12</i>)
5. Not definable (nd)	Item not defined in ICF	
6. Not contained (nc)	Item not represented in ICF	

1.2 Proposed approaches for qualifier mapping

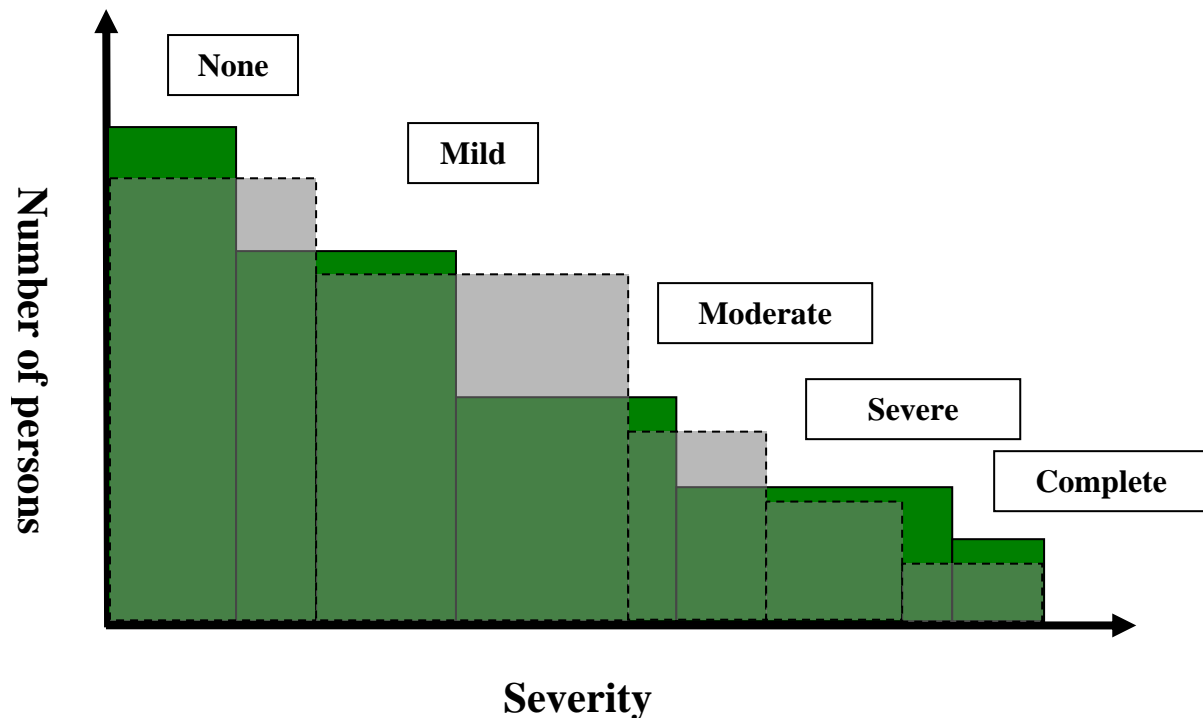
For mapping the response scale of existing assessment instruments with the ICF qualifier scale two approaches are proposed. First, to develop a topographic map which shows the categorical match between the a response category and the corresponding ICF qualifier with its anchor points.

Table 2. Establish response scale equivalence through categorical match

ICF Severity Scale	Duration	Intensity	Frequency	Vision scale
0 = none	≤ 5%	Not noticeable	Never	18-20/20
1 = mild	< 25%	tolerable	Rarely	16-17/20
2 = moderate	≥ 25%	interference with daily life	Occasionally	12-15/20
3 = severe	> 50%	Partial disruption with daily life	Frequently	04-11/20
4 = complete	> 95%	full disruption with daily life	Constantly	0-3/20

The second approach is to establish the equivalence between different response scales through matching frequency and difficulty.

Table 3: Establish equivalence between different response scales through matching frequency and difficulty



2. WHO DAS II further development

In response to growing WHO DAS user demands for recording information on impairments and environmental factors in addition to information on activities and participation WHO is planning to develop a WHO DAS impairment module and a probe on Environmental Factors.

The following WHO DAS domains for impairments in Body Functions and Body Structures are proposed:

- Pain and discomfort
- Cognition (memory)
- Vision
- Sleep and energy
- Affect (mood)
- Additional domains

For capturing Environmental Factor information the development of questions or probes are proposed. Furthermore, it is suggested to conduct comparability studies.

III. Next steps

Within the WHO FIC Network numerous activities are ongoing or upcoming in terms of mapping assessment instruments to ICF as well as using and further development of ICF as a homogenous assessment base. To synergize this activities and experience and avoid duplication of efforts the following steps are proposed:

- Protocol development for mapping assessment instruments to ICF and WHO DAS II further development
- Establish a portal on the WHO FIC website
- Development and field testing of maps and WHO questions on impairments and environmental factors.