# Towards the system-wide implementation of the international classification of functioning, disability and health in routine practice: lessons from a pilot study in China

**J. Li, Birgit Prodinger, J. Reinhardt, G. Stucki**

**SPECIAL REPORT**

# TOWARDS THE SYSTEM-WIDE IMPLEMENTATION OF THE INTERNATIONAL CLASSIFICATION OF FUNCTIONING, DISABILITY AND HEALTH IN ROUTINE PRACTICE: LESSONS FROM A PILOT STUDY IN CHINA

Jianan Li, MD[1,2], Birgit Prodinger, PhD[3,4,5], Jan Reinhardt, PhD[3,4,6] and
Gerold Stucki, MD, MS[3,4,5]

*From the [1]The First Affiliated Hospital of Nanjing Medical University, Nanjing, [2]Chinese Association of Rehabilitation Medicine, Beijing, China, [3]Swiss Paraplegic Research, Nottwil, [4]Department of Health Sciences and Health Policy, University of Lucerne, Lucerne, [5]ICF Research Branch, a cooperation partner within the World Health Organization (WHO) Collaborating Centre for the Family of International Classifications in Germany (at DIMDI), Nottwil, Switzerland and [6]Institute for Disaster Management and Reconstruction, Sichuan University and Hong Kong Polytechnic University, Chengdu, Sichuan, China*

**In 2011 the Chinese leadership in rehabilitation, in collaboration with the International Classification of Functioning, Disability and Health (ICF) Research Branch, embarked on an effort towards the system-wide implementation of the ICF in the healthcare system in China. We report here on the lessons learned from the pilot phase of testing the ICF Generic Set, a parsimonious set of 7 ICF categories, which have been shown to best describe functioning across the general population and people with various health conditions, for use in routine clinical practice in China. The paper discusses whether classification and measurement are compatible, what number of ICF categories should be included in data collection in routine practice, and the usefulness of a functioning profile and functioning score in clinical practice and health research planning. In addition, the paper reflects on the use of ICF qualifiers in a rating scale and the particularities of certain ICF categories contained in the ICF Generic Set when used as items in the context of Chinese rehabilitation and healthcare. Finally, the steps required to enhance the utility of system-wide implementation of the ICF in rehabilitation and healthcare services are set out.**

*Key words:* International Classification of Functioning, Disability and Health; rehabilitation; health information standards; classification; measurement.

## INTRODUCTION

In response to people's health problems and needs, health systems must be continuously improved. The World Health Organization (WHO) has identified 4 functions in the drive towards optimal service delivery and performance: policy-making and regulation; health services provision; financing; and managing of resources (1). For a health system aiming at continuous improvement, appropriate health information is fundamental. The collection of such information requires:

- standardization, ensuring comparability along the continuum of care and across health conditions, and
- practicability of collection, ensuring efficiency and compliance.

For standardization of health information countries rely on the classifications provided by the Classification, Terminology and Standards (CTS) team at the World Health Organization (WHO) (2). The first, and now in most countries widely used, classification is the International Classification of Diseases (ICD). The ICD serves a purpose in the context of morbidity and mortality statistics and reimbursement of predominantly acute medical care. Since 2001, the WHO has also provided the International Classification of Functioning, Disability and Health (ICF), which serves as the standard for the description of a person's health and functioning. The ICF therefore complements the ICD and provides essential information with regards to clinical health outcomes, rehabilitation management and financing, as well as work and, more generally, social integration (3).

In 2011, the Chinese leadership in rehabilitation, which comprises representatives of the president's cabinet from the 2 main rehabilitation societies, namely the Chinese Association of Rehabilitation Medicine (CARM) and the Chinese Society of Physical Medicine and Rehabilitation, in collaboration with colleagues from the ICF Research Branch, a cooperation partner within the WHO Collaborating Centre for the Family of International Classifications in Germany, has embarked in an effort towards the system-wide implementation of the ICF in the healthcare system in China, as described previously (4). A first step in this effort was a pilot testing of the usefulness of the ICF Generic Set for collecting functioning information in routine practice within the Chinese healthcare system. The ICF Generic Set is a parsimonious set of 7 ICF categories, which have been shown to best describe functioning across the general population and people with various

health conditions (Table I). Therefore, it can be used generically in the health system, as it derives from the ICF, which is universally accepted and is useful in practice, given that it consists of only 7 categories. Since the initial publication in 2011, the final version of this ICF Generic Set has been made available (5). Furthermore, there is also an extension of the ICF Generic Set available, the ICF Rehabilitation Set, which contains the 7 ICF categories of the ICF Generic Set and an additional 23 ICF categories reflecting aspects of functioning most relevant to describe functioning across health conditions and along the continuum of care. This set of 30 ICF categories can now serve as the foundation for what to report in routine practice (6). It is important to note that, for practical reasons, the ICF Rehabilitation Set is typically used as a pre-selection of categories that may be used by clinicians, depending on the specific situation. The Rehabilitation Set may therefore provide an additional standardization beyond the compulsory documentation of the ICF Generic Set.

The objectives of this paper are to report on the lessons from the pilot phase of testing the ICF Generic Set for use in routine clinical practice in China and to set out the way forward. The specific aims are:

- to clarify some questions relevant to the system-wide implementation of the ICF
- to summarize the lessons learned from this pilot study
- to outline steps needed to further enhance the utility of the

Table I. *International Classification of Functioning, Disability and Health (ICF) Categories contained in the ICF Generic Set (in bold) and ICF Rehabilitation Set*

| b130 | Energy and drive functions |
|------|----------------------------|
| b134 | Sleep functions |
| **b152** | **Emotional functions** |
| **b280** | **Sensation of pain** |
| b455 | Exercise tolerance functions |
| b620 | Urination functions |
| b640 | Sexual functions |
| b710 | Mobility of joint functions |
| b730 | Muscle power functions |
| **d230** | **Carrying out daily routine** |
| d240 | Handling stress and other psychological demands |
| d410 | Changing basic body position |
| d415 | Maintaining a body position |
| d420 | Transferring oneself |
| **d450** | **Walking** |
| **d455** | **Moving around** |
| d465 | Moving around using equipment |
| d470 | Using transportation |
| d510 | Washing oneself |
| d520 | Caring for body parts |
| d530 | Toileting |
| d540 | Dressing |
| d550 | Eating |
| d570 | Looking after one's health |
| d640 | Doing housework |
| d660 | Assisting others |
| d710 | Basic interpersonal interactions |
| d770 | Intimate relationships |
| **d850** | **Remunerative employment** |
| d920 | Recreation and leisure |

ICF as a standard for reporting functioning in rehabilitation, and ultimately the healthcare system at large, in China.

## CLARIFYING QUESTIONS REGARDING SYSTEM-WIDE IMPLEMENTATION OF THE ICF

Based on our discussions with various experts within and beyond the Chinese healthcare system, the topic related to the compatibility of classification and measurement, the "correct" number of ICF categories to be included in routine reporting and data collection, and the development and usefulness of a score of functioning remained salient.

### Are classification and measurement compatible?

The ICF is a health classification that builds on an integrative model and aims to describe the situation of a person within an array of health and health-related domains given the person's context. ICF categories are the unit of the classification and are in hierarchical order. Each category is nested within the next hierarchical level up the classification tree, in which a category at the third level (e.g. *d5300 Regulating urination*) is contained in its associated second-level category (e.g. *d530 Toileting*).

A category is more meaningful once assigned a qualifier, as it specifies the extent of an impairment, limitation or restriction. The qualifiers are built on a percentile scheme for reporting without any claim to meet the requirements of scientific measurement. As Bond and Fox (7) stated:

Classification and seriation are necessary precursors to the development of measurement systems, they are not sufficient for measurement. The distinctive attribute of a measurement system is the requirement for an arbitrary unit of differences that can be iterated between successive lengths (7, p. 28).

A single score of functioning must be metrically sound in order to be most relevant for further analyses and rigorous inferences (8). Once such a score is available it is most useful for the purpose of predicting, e.g. mortality, readmission rates, length of stay and for monitoring functioning across time. Hence, developing a score that is suitable for use in various statistical analyses based on the domains contained in the ICF is reasonable. In this case, the ICF determines the domains to be assessed. How to assess the extent of a problem for a given domain in a reliable and valid way must be addressed from a measurement perspective.

### What is the "correct" number of ICF categories to include in routine reporting and data collection?

Functioning is a complex construct that must be seen in interaction with a person's health condition and the context in which the person is situated. For routine reporting and data collection the challenge becomes to define a set of essential categories that is sufficiently succinct to be administered in routine practice, and yet comprehensive enough to represent the complexity inherent in WHO's conceptualization of health. To foster the utility of the ICF in practice and research, and enhance the comparability of functioning information locally

(including across patients and time), nationally (including across settings and regions), and internationally (including across countries), ICF Core Sets for specific health conditions or settings have been developed based on a multi-stage international consensus process (9), as well as the ICF Generic Set and Rehabilitation Set. It must be kept in mind that the ICF Generic Set constitutes a parsimonious set of essential categories that can be included at all levels of the healthcare system (5). For more specific settings, such as rehabilitation, where functioning is the primary outcome (10), it is recommended to use the ICF Rehabilitation Set as it allows more comprehensive description of the lived experience of people living with a certain health condition and discriminates different levels of functioning within and between relevant groups (6). The multi-stage process of developing the ICF Rehabilitation Set, described elsewhere (11), points to its content validity, and yet the proposed sets represent a first step and reference framework for collecting and reporting functioning across general and clinical populations.

Any set of ICF categories can be complemented with additional ICF categories from specific ICF Core Sets or the ICF itself to meet local needs. The approach of having an essential set of ICF categories (ICF Generic and Rehabilitation Set), more extensive sets and yet consensus-based parsimonious sets (health condition and context specific Core Sets), as well as the spectrum of ICF categories from the ICF, allows a minimum standard for reporting to be maintained, and yet adds flexibility for professionals in routine practice and enhances the practicability of the ICF. The overlap and step-wise process toward selecting ICF Core Sets is shown in Fig. 1.

*Usefulness of a functioning profile and functioning score for clinical practice and health service planning*

ICF Core Sets can be integrated into a documentation template to gain a functioning profile of an individual as a starting point for interdisciplinary team discussions to monitor the rehabilita-



*Fig. 1.* Stepwise selection of International Classification of Functioning, Disability and Health (ICF) sets. Overlapping circles indicate that the ICF Generic Set is recommended to be included in any other set. The ICF Rehabilitation Set may not completely overlap with the health condition or context-specific Core Sets.

tion process, including assessment, goal-setting, intervention, and evaluation. A functioning profile is an illustration of the levels of functioning across domains included in the selected ICF sets (12). Functioning profiles are valuable tools to support decision-making in clinical practice and monitor change across aspects of functioning over time. While the visualization of a person's functioning profile is meaningful and useful, decision-makers in health service planning often rely on sum scores across these different aspects of functioning. To aggregate information obtained from a functioning profile into a score, it requires that all items relate to a single specific construct (unidimensionality; 7) and have the same meaning across various groups of patients (invariance; 13) from a psychometric point of view. Both unidimensionality and invariance are requirements in aggregating information across categories and must be confirmed using modern psychometric methods prior to aggregating information (14, 15). If unidimensionality and invariance are confirmed it implies that all ICF categories relate to a common construct, i.e. functioning, and creating a sum score is valid.

## LESSONS FROM USE OF THE ICF GENERIC SET IN ROUTINE PRACTICE

A more detailed review of the design and results of the pilot study, including feasibility, metric properties, and sensitivity to change, has been published elsewhere (16).

*Using ICF qualifiers as rating scale*

In the pilot study, ICF categories were used as items and complemented with ICF qualifiers as rating scale. ICF qualifiers constitute a range from 0 to 4, whereby the margins of a code are indicated with a percentage range (17) (Table II).

The percentiles are relevant for standardized reporting of the extent of a problem with respect to a specific ICF category; however, when used as a rating scale in clinical practice there is evidence that the margins of the qualifiers are ambiguous as response categories and are not used consistently (18–20). Thus, there is no consistent rating of, for example *2 moderate problems* and *3 severe problems*, which is a thread to the validity of the data. Rasch measurement methods facilitate the examination of the ordering of response categories by evaluating thresholds. The threshold represents the point between 2 response options, e.g. *2 moderate problems* and *3 severe prob-*
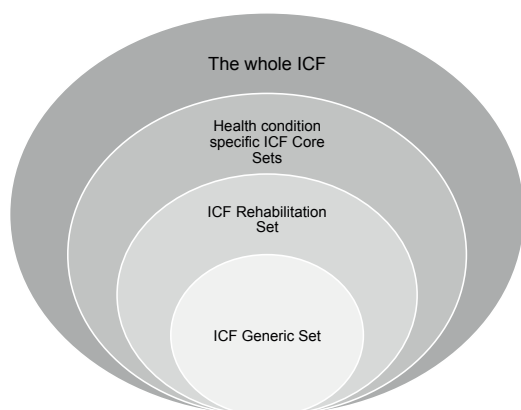
Table II. *Descriptions and percentile ranges of International Classification of Functioning, Disability and Health (ICF) qualifiers as specified in the ICF (16)*

| ICF qualifier | | Percentile range (%) |
|---|---|---|
| 0 = No problem | none, absent, negligible,… | 0–4 |
| 1 = Mild problem | slight, low, … | 5–24 |
| 2 = Moderate problem | medium, fair, … | 25–49 |
| 3 = Severe problem | high, extreme, … | 50–95 |
| 4 = Complete problem | total, … | 96–100 |
| 8 = Not specified | | |
| 9 = Not applicable | | |

*lems*, where the person has an equal probability of obtaining 1 of these adjacent response options (21). Disordered thresholds are a sign that response options are not working as intended, and thus provide empirical evidence on the ambiguity inherent in the use of ICF qualifiers as rating scale. Fig. 2 illustrates ordered thresholds for the ICF category *b130 Energy and drive functions* and disordered thresholds for the ICF category *d450 Walking*. In psychometric analyses disordered thresholds can be dealt with by collapsing response options. As this remedy implies some loss of information, it is recommended that response options are developed to be as unambiguous as possible. The challenges identified in this pilot study underscore the necessity raised by WHO (17) of developing clinical data collection tools that facilitate the proposed quantification of the extent of a problem in a consistent manner.

*Particularities of some ICF categories contained in the ICF Generic Set when used as items in the context of Chinese rehabilitation and healthcare*

For instruments to be used in routine clinical practice, their ability to detect change in functioning over time is important. *d850 Remunerative employment i*s considered relevant within a minimum generic set of domains to describe functioning across the general population and people with varying health conditions. Previous psychometric research indicated that this ICF category does not discriminate well between individuals with different levels of functioning, which might be an indicator that remunerative employment is influenced more by the interaction of the person with the environment rather than the person's level of functioning (22, 23). Furthermore, in certain countries remunerative employment would not change

throughout a hospital stay; the capacity or actual performance of a person to conduct work-related tasks may change, but the employment status of a person would not. Hence, for the future, one might consider reporting *d850 Remunerative employment* in a functioning profile if reasonable within a given context, but not including it in a functioning score.

In the pilot study, *b280 Sensation of pain* did not fit into a unidimensional scale with other categories of the ICF Generic Set (16). Clinically, pain is a difficult to measure construct as it may vary depending on its aetiology, e.g. being nociceptive or neuropathic, its quality, intensity, and its course, e.g. being episodic or chronic. Instances whereby sources of variation in assessing the level of pain introduced by other factors than the trait to be assessed are referred to as differential item functioning (DIF). For example, when persons with a particular health condition, with the same level of functioning, report differently to a specific item. From a measurement point of view this can be dealt with using specific methods dealing with DIF. Items are split by groups displaying DIF and then treated as group-specific items. Underlying these approaches is the assumption that, in theory, pain can be assessed in a generic manner across health conditions. In future studies, the ICF category *b280 Sensation of pain* requires further clarification to reliably integrate the item derived from this category into a functioning score.

Based on the insights gained from the pilot study in China and for the reasons outlined above, *d850 Remunerative employment* and *b280 Sensation of pain* are not necessarily recommended to be included in a functioning score across ICF categories of the ICF Generic Set at this point. However, it remains important that these categories are reported inde-
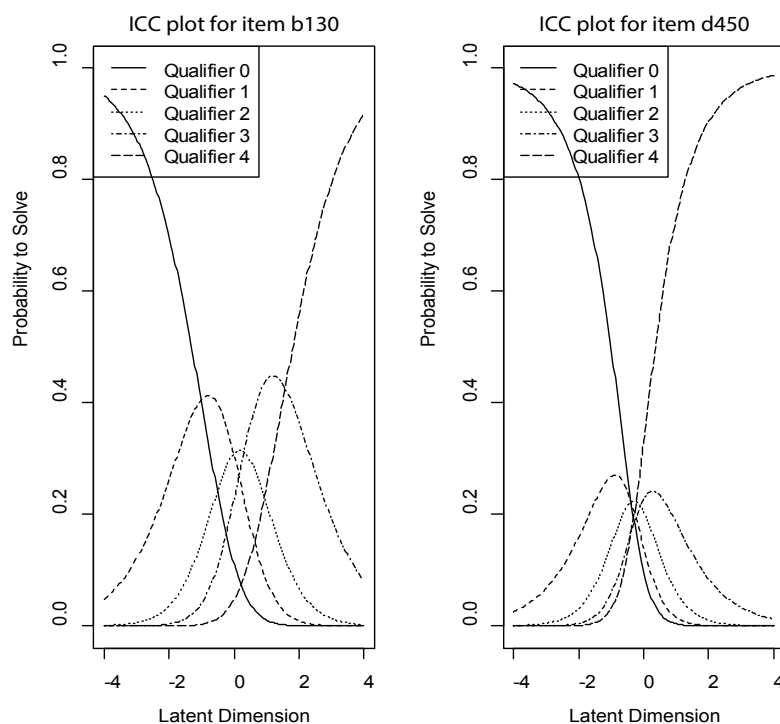


*Fig. 2.* Ordered and disordered thresholds of 2 International Classification of Functioning, Disability and Health (ICF) categories that have been administered with the ICF qualifiers as rating scale. The left-hand box shows the item characteristic curves (ICC) for the ICF category *b130 Energy and drive functions*, which has been administered with the ICF generic problem qualifier. Each curve represents 1 response category that has been shown to be nicely ordered. In contrast, the right-hand box presents the ICC for the ICF category *d450 Walking*, where the thresholds of the response categories are not ordered consecutively, as in *b130 Energy and drive functions*.

pendently of a functioning score, to ensure that all relevant categories of the ICF Generic Set are represented.

## ENHANCING THE UTILITY OF THE ICF AS A STANDARD FOR REPORTING FUNCTIONING

Based on the lessons from the pilot study, we can identify further steps to be addressed to enhance the utility of the ICF as a standard for assessing and reporting functioning in the system-wide implementation in rehabilitation and healthcare services in general.

### *Develop simple, intuitive descriptions of ICF categories to enhance clarity*

The descriptions of the ICF categories constitute statements that define the concept designated by a category and its essential attributes, but they are not necessarily simple and intuitive descriptions that are commonly understood and consistently used by professionals within an interdisciplinary team. The psychometric analysis of data collected in the pilot study has highlighted some ambiguities with reference to a number of ICF categories, as discussed above. These ambiguities may also refer to different interpretations and understandings of what is meant by an ICF category across different professionals and raters. Hence, further specification of the ICF categories by finding consensus on simple and intuitive descriptions of each of the ICF categories contained in the ICF Rehabilitation Set is important to facilitate its utility in routine practice.

### *Generating an ICF reference metric for standardized reporting*

These simple and intuitive descriptions can subsequently serve as the foundation for expert ratings on the extent of a problem related to a specific ICF category in routine practice. A numeric rating scale, e.g. ranging from 0 to 10, can be used to rate the extent of a problem. The raw scores can subsequently be transferred into a user-friendly metric from 0 to 100 using the Rasch measurement model. This metric can be then used for reporting the extent of a problem a person experiences with respect to a given aspect of functioning.

### *Conclusion*

To ensure comparability of functioning information at all levels of the health system, across health conditions and time, a minimal universally applicable set of relevant domains is required. Such requirements are also called upon in the WHO Global Disability Action Plan 2014–2021, which defines "the collection of relevant and internationally comparable data on disability" as 1 of the 3 objectives (24). The International Society of Physical and Rehabilitation Medicine (ISPRM) has developed a work plan in response to the Global Disability Action Plan, which points to the development and implementation of the ICF in routine clinical practice (25). The lessons from the pilot study in China will inform future initiatives toward the development of clinical data collection tools. Most often the challenge in clinical practice is not a lack of information, but rather varying

sources of information derived from clinical tests, technical investigations, standardized patient-reported outcomes, and patient narratives, which clinicians have to integrate into an overall assessment of the patient's level of functioning. Once a standard set of ICF categories for reporting functioning has been implemented, the information gathered provides a comprehensive picture understood by various stakeholders within the health system as it builds on the unified language of the ICF for evidence-based decision-making at all levels. On the micro-level of clinical practice, a functioning profile may assist decision-making on intervention goals and targets. Once the dimensionality of the domains contained in the profile has been examined, the information can be aggregated into a score and serve as a foundation for decision-making on meso- and macro-levels. Based on our previous research, it can be concluded that the ICF Generic Set serves as a good starting point for facilitating the standardization of information on functioning and health across clinical (sub-) populations, time and settings. To strengthen the system-wide implementation of the ICF in China, further issues must be addressed, as outlined in this paper.

## REFERENCES

1. World Health Organization (WHO). Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies. Geneva: WHO; 2010.
2. World Health Organization (WHO). The WHO Family of International Classifications. Geneva: World Health Organization; 2014 [cited 2014 Aug 28]. Available from: http://www.who.int/classifications/en/.
3. Kostanjsek N, Rubinelli S, Escorpizo R, Cieza A, Kennedy C, Selb M, et al. Assessing the impact of health conditions using the ICF. Disabil Rehabil 2011; 33: 1475–1482.
4. Stucki G, Qiu Z-y, Li J-a, Wu X-g. Towards the system-wide implementation of the ICF in rehabilitation in China. Chin J Rehabil Theory Pract 2011; 17: 5–10.
5. Cieza A, Oberhauser C, Bickenbach J, Chatterji S, Stucki G. Towards a minimal generic set of domains of functioning and health. BMC Public Health 2014; 14: 218.
6. Prodinger B, Bickenbach J, Stucki G, Cieza A, editors. Toward the system-wide implementation of the ICF in clinical and rehabilitation practice to describe functioning along the continuum of care. WHO – Family of International Classifications Network Annual Meeting; Barcelona: 2014.
7. Bond TG, Fox CM, editors. Applying the Rasch model: fundamental measurement in the human sciences. 2nd ed., Mahwah, NJ: Lawrence Erlbaum Associates, Inc.; 2007.
8. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? J Rehabil Med 2012; 44: 97–98.
9. Selb M, Escorpizo R, Kostanjsek N, Stucki G, Ustün B, Cieza A. A guide on how to develop an international classification of functioning, disability and health core set. Eur J Phys Rehabil Med 2014; 51: 105–117.
10. Stucki G, Cieza A, Melvin J. The International Classification of Functioning, Disability and Health: a unifying model for the conceptual description of the rehabilitation strategy. J Rehabil Med 2007; 39: 279–285.
11. Prodinger B, Cieza A, Oberhauser C, Bickenbach J, Üstün TB, Chatterji S, et al. Toward the International Classification of Functioning, Disability and Health (ICF) Rehabilitation Set: A Minimal Generic Set of Domains for Rehabilitation as a Health Strategy.

Arch Phys Med Rehabil 2016 Jan 28. [Epub ahead of print].

12. Bickenbach J, Cieza A, Rauch A, Stucki G, editors. ICF Core Sets. Manual for Clinical Practice. Göttingen: Hogrefe Publishing; 2012.

13. Engelhard Jr G. Invariant measurement: using Rasch models in the social, behavioral, and health sciences. New York: Routledge; 2013.

14. Pallant J, Tennant A. Unidimensionality matters. Rasch Measurement Transactions 2006; 20: 1048–1051.

15. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007; 57: 1358–1362.

16. Reinhardt J, Zhang X, Prodinger B, Bostan C, Selb M, Stucki G, et al. Toward the system-wide implementation of the International Classification of Functioning, Disability and Health in routine clinical practice: Empirical findings of a pilot study from Mainland China. J Rehabil Med 2016; 48: 515–521.

17. World Health Organization (WHO). International Classification of Functioning, Disability and Health. Geneva: WHO; 2001.

18. Uhlig T, Lillemo S, Moe RH, Stamm T, Cieza A, Boonen A, et al. Reliability of the ICF Core Set for rheumatoid arthritis. Ann Rheum Dis 2007; 66: 1078–1084.

19. Røe C, Bautz-Holter E, Cieza A. Low back pain in 17 countries, a Rasch analysis of the ICF core set for low back pain. Int J Rehabil Res 2013; 36: 38–47.

20. Alviar MJ, Olver J, Pallant JF, Brand C, de Steiger R, Pirpiris M, et al. Can the ICF osteoarthritis Core Set represent a future clinical tool in measuring functioning in persons with osteoarthritis undergoing hip and knee joint replacement? J Rehabil Med 2012; 44: 955–961.

21. Tennant A. Disordered Thresholds: an example from the Functional Independence Measure. Rasch Meas Trans 2004; 17: 945–948.

22. Cieza A, Hilfiker R, Chatterji S, Kostanjsek N, Üstün B, Stucki G. The International Classification of Functioning, Disability, and Health could be used to measure functioning. J Clin Epidemiol 2009; 62: 899–911.

23. Cieza A, Hilfiker R, Boonen A, Van der Heijde D, Braun J, Stucki G. Towards an ICF-based clinical measure of functioning in people with ankylosing spondylitis: a methodological exploration. Disabil Rehabil 2009; 31: 528–537.

24. WHO. WHO Global Disability Action Plan 2014–2021. Geneva: WHO; 2015 [cited 2015 Aug 3]. Available from: http://www.who.int/disabilities/actionplan/en/.

25. Gutenbrunner C, Negrini S, Kiekens C, Zampolini M, Nugraha B. The Global Disability Action Plan 2014–2021 of the World Health Organization (WHO): a major step towards better health for all people with disabilities. Change and challenge for Physical and Rehabilitation Medicine (PRM). Eur J Phys Rehabil Med 2015; 51: 1–4.