Japanese Journal of Comprehensive Rehabilitation Science (2020)

Original Article

Interrater agreement between clinician ratings and patient self-assessments for body function categories of ICF Rehabilitation Set

Yuki Senju, MD,[1] Masahiko Mukaino, MD, DMSc,[1] Megumi Ozeki, MD, DMSc,[2] Makoto Watanabe, OTR,[3] Yuki Okochi, OTR,[4] Koji Mizutani, RPT,[4] Eiichi Saito, MD, DMSc,[1] Shigeru Sonoda, MD, PhD[5]

[1]Department of Rehabilitation Medicine I, School of Medicine, Fujita Health University, Aichi, Japan
[2]Faculty of Rehabilitation, School of Health Sciences, Fujita Health University, Aichi, Japan
[3]Department of Rehabilitation Medicine, Fujita Health University Nanakuri Memorial Hospital, Mie, Japan
[4]Department of Rehabilitation Medicine, Fujita Health University Hospital, Aichi, Japan
[5]Department of Rehabilitation Medicine II, School of Medicine, Fujita Health University, Mie, Japan

ABSTRACT

Senju Y, Mukaino M, Ozeki M, Watanabe M, Okochi Y, Mizutani K, Saito E, Sonoda S. Interrater agreement between clinician ratings and patient self-assessments for body function categories of ICF Rehabilitation Set. Jpn J Compr Rehabil Sci 2020; 11: 9−16.

Objective: This study investigates interrater agreement between clinician ratings using the Rating reference guide developed in Japan for the International Classification of Functioning, Disability and Health (ICF) Rehabilitation Set and patient self-assessments for body function categories of the ICF Rehabilitation Set.

Methods: Eighty-eight inpatients and/or outpatients who received rehabilitation participated in this study. For the assessment of the nine body function categories of the ICF Rehabilitation Set, the patients were asked to complete the self-assessment questionnaires, and the clinicians rated the patients using the Rating reference guide. Interrater agreement between clinicians and patients was investigated using weighted kappa statistics and an intraclass correlation coefficient (ICC (2,1)) to determine interrater agreement of each category and the total score, respectively.

Results: The weighted kappa statistics ranged from 0.58 to 0.87. Eight out of nine categories presented with weighted kappa statistics greater than 0.61. The total score of all categories showed no significant difference between clinicians and patients and presented with an ICC (2,1) of 0.85.

Conclusion: No significant difference was observed between clinician ratings with the Rating reference guide for body function categories and patient self-assessments, showing feasibility of the Rating reference guide as a means of describing the status of patients' functioning.

Key words: ICF, ICF Rehabilitation Set, Rating reference guide, body function, interrater agreement

Introduction

The International Classification of Functioning, Disability and Health (ICF) is a framework for describing and organizing information on patient functioning and disability [1, 2]; it covers information ranging from the status of functioning related to the patient's health condition to the social system or social resources surrounding the patient. The ICF consists of more than 1,400 categories and provides both a standard language and a conceptual basis for defining and measuring health and disability that enable the ICF to comprehensively describe patients' status of functioning. Since the ICF was approved in May 2001, various actions have been taken to promote the implementation of the international framework.

The ICF Core Set project is one of the major projects promoted in the framework. This project refers to the select set of ICF categories that can serve as minimal standards for the assessment and documentation of the functioning and health of persons with various health conditions and/or diseases and that can be developed through the predefined objective process [3−8]. There are also ICF Core Sets that could be applied across various disease or health conditions. For this purpose, two Core Sets have been developed. The ICF Generic

Correspondence: Masahiko Mukaino, MD, DMSc
Department of Rehabilitation Medicine I, School of Medicine, Fujita Health University, 1−98 Dengakugakubo, Kutsukake, Toyoake, Aichi 470−1192, Japan.
E-mail: mmukaino@fujita-hu.ac.jp
Accepted: September 17, 2019.
The authors have no conflict of interest directly relevant to the content of this article.

doi.org/10.11336/jjcrs.11.9

Set, developed by psychometric study, consists of the seven ICF categories that are most relevant across the general population and across different clinical populations, contexts, settings, and purposes [9]. An extended version of this set for use in the rehabilitation setting, the ICF Rehabilitation Set with 30 ICF categories, was developed from the results of additional psychometric studies and expert reviews of the ICF Core Sets (Table 1) [9, 10]. Projects to create tools based on these ICF Core Sets has been advanced in the international framework by expert groups such as the International Society of Physical and Rehabilitation Medicine and the Physical and Rehabilitation Medicine Section and Board of the European Union of Medical Specialists [11, 12]. The focus of one ongoing project is to assist clinicians with categorization by developing simple, intuitive descriptions for ICF categories that have titles and definitions that are not intuitive [13, 14]. In Japan, a Japanese version of the simple, intuitive descriptions of ICF Rehabilitation Set was created in 2016 (Mukaino et al, in submission).

Another ongoing project aims to develop clinical rating tools [12]. For the coding of patients with functioning problems, the ICF has its own rating system that uses qualifiers. The qualifiers are defined in the ICF book as follows: 0: no problem, 1: mild problem (5−24%), 2: moderate problem (25−49%), 3: severe problem (50−95%), 4: complete problem (96−100%), 8: not specified, 9: not applicable [1]. Unlike most of the existing clinical scales, the qualifiers have no further explanation. This may cause the difficulty and low reliability in rating the patients' status in the categories the clinicians are not accustomed to evaluating. For example, Uhlig et al reported on the interrater reliability of clinician ratings of the ICF Core Set for rheumatoid arthritis. Although they could show relatively high interrater reliability in several categories with which the clinicians were familiar, such as toileting or dressing, most of the categories presented with low weighted kappa statistics and thus low interrater reliability [15].

Table 1. ICF Rehabilitation Set.

	Categories
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

The difficulties in rating could be seen particularly in the body function categories. As is done in many of the existing clinical scales, the problems in activities and participation may be simply estimated by clinicians based on the patients' status in performing the activity described in the given category. This could be considered in a single dimension. However, in the evaluation of body function, estimating the degree of patients' problems is not straightforward. For example, "b280 Sensation of pain" could be evaluated by clinicians in light of several aspects, including the extent of pain, frequency of pain, and site of pain. However, these aspects may differ between the raters. For example, mild and infrequent pain may be perceived as a problem by some patients but not others. In addition, what constitutes either a severe problem or a mild problem may easily vary according to what aspects the raters focus on.

More recently, the development of ICF-based data collection tools in an international framework was proposed to solve these problems [12]. Along with these international efforts, a Rating reference guide was developed in Japan. The Rating reference guide is a clinical tool designed to support ICF-based patient evaluation and developed from the analysis of clinician's thinking processes in using ICF by a cognitive interviewing of clinicians (Senju et al, in submission). For example, the Rating reference guide for body function categories indicates 1) what aspects the raters focus on when rating, and 2) what each response option (0−4) indicates. The study showed that the weighted kappa statistics of all categories exceeded 0.60, indicating that using such a Rating reference guide results in generally good interrater reliability [16]. However, toward the clinical implementation of the ICF, it is important to ensure the rating is not only reliable among clinicians but also in agreement with patient self-assessments. The ICF is developed to describe the "lived experience" of the patients [17]. In

other words, rather than just looking at the severity of symptoms according to clinicians' perspectives, patients' perspectives should be considered. Naturally, there will be differences in the evaluation based on the objective viewpoint of clinicians and the subjective problem recognition of patients. However, based on the ICF concept defined in this paper, the evaluation by clinicians should not diverge significantly, even if it does not completely match patients' own evaluations. Based on such considerations, we examined interrater agreement between clinician ratings using the Rating reference guide and patients' own evaluations using the self-assessment questionnaires and distribution of evaluation for each body function category in the ICF Rehabilitation Set.

Methods

1. Subjects

Eighty-eight patients who received rehabilitation from April 2017 to November 2018 in Fujita Health University Hospital and Fujita Health University Nanakuri Memorial Hospital participated in this study. The mean age of the patients was 65.4 ± 15.3 years. Of the 56 male and 32 female patients, 63 patients had neurological diseases such as stroke, 10 patients had musculoskeletal diseases, 4 patients had cardiovascular diseases, 4 patients had pulmonary diseases, and 7 had others; there were 77 inpatients and 11 outpatients (Table 2). We excluded the patients whose Mini-Mental State Examination (MMSE) score was less than 23 or who could not answer the self-assessments questionnaire.

2. Methods

All patients were asked to complete the self-assessments questionnaire for the nine body function categories of the ICF Rehabilitation Set (Table 3), and clinicians provided ratings using the Rating reference guide (Table 4). The assessment of body function categories of the ICF Rehabilitation Set by the clinicians was conducted by a single rater out of five rehabilitation experts (two physiatrists and three therapists, with a combined 6−18 years of experience (median: 16 years)). The Rating reference guide shows the following points for each category: 1) what aspects (e.g., frequency, extent) the raters focus on, and 2) what the "4: complete problem" indicates. In addition, common brief guidelines to frame each response option were provided: 0: no problem, 1: mild problem (may include problem that does not affect daily activities), 2: moderate problem (may include problem that exceeds "1" but still remains relatively minor problem in given category), 3: severe problem (may

Table 2. Demographic variables.

Sample size (*n*)	88
Age (years)	65.4 ± 15.3
Age, mean ± SD (years)	
Gender	*n* (%)
Male	56 (63.6)
Female	32 (36.4)
Diagnosis	*n* (%)
Cerebrovascular disease	63 (71.6)
-Supratentorial lesions	45 (51.1)
Orthopedic disease	10 (11.4)
Respiratory disease	4 (4.6)
Cardiovascular disease	4 (4.6)
Other	7 (8.0)
Inpatients	77 (87.5)
Outpatients	11 (12.5)

Table 3. Self-assessment questionnaire.

How much do you have problem in the following functions ?

	Categories	No problem	Mild problem	Moderate problem	Severe problem	Complete problem	I don't know
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						

Table 4. Rating reference guide.

	Categories	What would be complete problem	What aspect to score
b130	Energy and drive functions	The extreme should be that they have no motivation or appetite at any time	・The extent of the problem ・The frequency of the problem
b134	Sleep functions	The extreme should be that they cannot sleep at all or the sleeping cycle has completely collapsed every day	・The extent of the problem ・The frequency of the problem
b152	Emotional functions	The extreme should be that they completely lose control of emotions every day or are completely incapable of expressing emotions	・The extent of the problem ・The frequency of the problem
b280	Sensation of pain	The extreme should be that they have continuous intolerable pain at any time	・The extent of the problem ・The frequency of the problem ・The number of site with pain
b455	Exercise tolerance functions	The extreme should be that they cannot bear any single activity of daily living at any time	・The extent of the problem ・The frequency of the problem
b620	Urination functions	The extreme should be that they cannot urinate or control urination at all every day	・The extent of the problem ・The frequency of the problem
b640	Sexual functions	The extreme should be that they have no mental (may include having no sexual desire or cannot control sexual desire at all) or physical ability to perform sexual activity at any time	・The extent of the problem ・The frequency of the problem
b710	Mobility of joint functions	The extreme should be that all of their major joints are completely fixed	・The extent of the problem ・The ratio of the joint which have the problem
b730	Muscle power functions	The extreme should be that all of their major muscles are completely incapable of movement	・The extent of the problem ・The ratio of the joint which have the problem

Qualifiers

0 No problem
1 Mild problem: may include a problem which doesn't affect the patient's daily activities
2 Moderate problem: may include a problem which exceeds "1" but still remains a relatively minor problem in the given category
3 Severe problem: may include a problem which is a major problem in the given category
4 Complete problem
8 Not specified
9 Not applicable

include problem that is major in given category), 4: complete problem [16].

The weighted kappa statistics were used to determine interrater agreement between clinicians and patients for each category. To examine the difference between patients with brain lesions and others, we conducted a comparison between the two groups of patients. We also evaluated the distribution and average total scores of clinicians' and patients' evaluations, where intraclass correlation coefficient (ICC (2,1)) was used to determine the agreement in total scores between clinicians and patients.

3. Data analysis

The weighted kappa statistics and ICC (2,1) were used to determine interrater agreement between clinicians and patients in each category and the total score. The weighted kappa statistics with linear weight were calculated for each category. The qualifier response options of 8 for "not specified" and 9 for "not applicable" were not included in the calculation of kappa statistics and were considered missing data. The following standards for interpreting kappa statistics presented by Landis and Koch were employed in our interpretation of the results: <0.20, poor; 0.21−0.40, fair; 0.41−0.60, moderate; 0.61−0.80, substantial; and >0.81, excellent [18]. For interpretation of the ICC (2,1), the evaluation criteria of Cicchetti et al. were used (<0.40, poor; 0.40−0.59, fair; 0.60−0.74, good; >0.75, excellent [19]).

Table 5. Weighted kappa statistics and ICC (2,1).

	Categories	Missing value		Weighted kappa
		Clinician ratings *n* (%)	Patient self-reports *n* (%)	
b130	Energy and drive functions	0 (0.0)	0 (0.0)	0.72
b134	Sleep functions	0 (0.0)	0 (0.0)	0.74
b152	Emotional functions	0 (0.0)	0 (0.0)	0.58
b280	Sensation of pain	2 (2.3)	3 (3.5)	0.69
b455	Exercise tolerance functions	0 (0.0)	2 (2.3)	0.63
b620	Urination functions	0 (0.0)	0 (0.0)	0.87
b640	Sexual functions	31 (35.2)	22 (25.0)	0.67
b710	Mobility of joint functions	0 (0.0)	2 (2.3)	0.70
b730	Muscle power functions	0 (0.0)	2 (2.3)	0.62
				ICC (2,1)
	Total	31 (35.2)	22 (25.0)	0.85

4. REDCap

These study data were collected and managed using REDCap (Research Electronic Data Capture) electronic data capture tools hosted at Fujita Health University [20]. REDCap is a secure, web-based application designed to support data capture for research studies. The application provides 1) an intuitive interface for validated data entry, 2) audit trails for tracking data manipulation and export procedures, 3) automated export procedures for seamless data downloads to common statistical packages, and 4) procedures for importing data from external sources.

Results

The results of interrater agreement between clinician ratings and patient self-assessments for body function categories are shown in Table 5. The weighted kappa statistics ranged from 0.58 to 0.87. Eight out of nine categories presented high weighted kappa statistics, and these results were classified as substantial to excellent. Only the weighted kappa statistics of "b152 Emotional functions" were classified as moderate reliability. There was a category ("b620 Urination functions") with more than 0.81 of weighted kappa statistics classified as excellent in agreement. A marked number of missing values was observed in "b640 Sexual functions", which presented with 35.2% missing values in clinicians' evaluations and 25.0% in patients' evaluations; nevertheless, other items showed less than 5% of missing values.

The ICC (2,1) of the total score of all categories excluding missing value was 0.85, which was classified as excellent. The distributions of the total score of

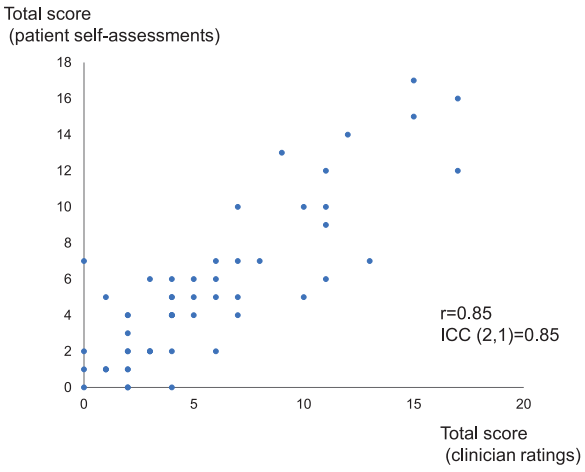

Figure 1. The total scores of the clinician ratings and the patient self-assessments.

The correlation coefficient between the total score of clinician ratings and the patient self-assessments was 0.85, and ICC (2,1) was 0.85. *r*: Pearson's correlation coefficient, ICC: intraclass correlation coefficient.

Table 6. Weighted kappa statistics with and without brain lesion.

		Weighted kappa	
		With brain lesion	Without brain lesion
b130	Energy and drive functions	0.63	0.90
b134	Sleep functions	0.71	0.76
b152	Emotional functions	0.50	0.78
b280	Sensation of pain	0.65	0.76
b455	Exercise tolerance functions	0.52	0.77
b620	Urination functions	0.86	0.89
b640	Sexual functions	0.67	1.00
b710	Mobility of joint functions	0.68	0.71
b730	Muscle power functions	0.58	0.73

clinician ratings and patient self-assessments are shown in Figure 1. There was no significant difference in total score between clinician ratings and patient self-assessments (clinician ratings: 5.6 ± 4.5, patient self-assessments: 5.2 ± 4.1). The weighted kappa statistic of patients with and without brain lesion are shown in Table 6. In the patients with brain lesions, interrater agreement between the rating of clinicians and patients tended to be generally low compared to the patients without brain lesions.

Discussion

In this study, we examined the agreement in evaluation between clinician ratings with the Rating reference guide and patient self-assessments for the nine body function categories of the ICF Rehabilitation Set. Eight out of nine categories presented high weighted kappa statistics, and these results were classified as substantial to excellent. The ICC (2,1) was more than 0.75, which was classified as excellent. Thus, there was relatively good clinical rating agreement between clinicians and patients in these categories. This would indicate that the clinician ratings using the Rating reference guide properly reflect the problems the patients experience and are thus clinically useful.

The weighted kappa statistic of "b152 Emotional functions" was 0.58, which was relatively lower than the other categories. This may be due to clinicians' insufficient understanding of patients' emotional functions, which may come from the patients' lack of expression or clinicians' insufficient observation of the problems in emotional functions. Another possible reason could be the existence of anosognosia on the part of the patients, which may result in the clinicians' poor evaluation of patients' problems [21–23]. Patients with obvious cognitive problems were excluded from the study. However, approximately 50% of our target included patients with neurological diseases such as stroke, and it is possible that they had mild cognitive problems (e.g., anosognosia) that may not have been reflected in the MMSE. The high interrater agreement of "b152 Emotional functions" in patients without brain lesions and the low interrater agreement of "b455 Exercise tolerance functions" and "b730 Muscle power functions", where clinician ratings relied on objective information, might indicate the influence of cognitive problems on the difference in ratings between clinicians and patients.

The previous studies suggest that clinicians tend to underestimate the problems of body function categories, particularly in evaluating those categories strongly influenced by subjectivity. For example, it is reported that clinicians often underestimate pain compared to patients' evaluations [24]. It has also been reported that clinicians tend to underestimate psychological problems such as depression [25]. These reports indicate the difficulties faced by clinicians when estimating the extent of subjective problems in the functioning of patients. However, in this study, there was good agreement between clinician ratings and patient self-assessments for each category. In addition, the ICC (2,1) of total points was in good agreement with 0.85, and the average of total score of clinician ratings and patient self-assessments showed no significant difference. The Rating reference guide used in this study was developed while taking several aspects into account, including the frequency and extent of functioning problems, in rating each category. The Rating reference guide also refers to the influence of the functioning problem on daily activities that are likely to strongly affect the subjective evaluation of the problem by patients (0: no problem, 1: mild problem (may include problem that does not affect daily activities), 2: moderate problem (may include problem that exceeds "1" but still remains a relatively minor problem in the given category), 3: severe problem (may include problem that is a major problem in given category), 4: complete problem). Therefore, it may be easier to agree with the patients' subjective evaluations than when simply focusing on single aspects such as the severity of the symptoms and extent of functioning problems.

Limitation

In this study, the raters in the field study consisted of experienced rehabilitation experts. As previous studies have shown, lack of clinical experience may lower interrater agreement in the rating of clinical scales [26–28]. Thus, the results may not be replicated with less experienced clinicians evaluating patients' functioning and should be further investigated.

In "b640 Sexual functions", missing values were high for both clinicians and patients. One possible reason could be that this category includes more private information than the other categories; thus, both clinicians and patients faced strong resistance when rating this category. Although the weighted kappa statistic was 0.67 and classified as moderate in agreement, the percentage of patients rated 0 by both clinicians and patients reached 84%, which was much higher than the other categories. Because sexual function is not a function required in daily life, the frequency of opportunities for sexual activity may affect the detection of the sexual function problem. It is also possible that clinicians and patients scored 0 because subjects were relatively old, 65 years on average, and many were inpatients; thus, both clinicians and patients had few opportunities to obtain objective clinical findings and subjective symptoms that could be used as a reference for evaluation. In evaluating interrater agreement between clinician ratings and patient self-assessments for sexual function, a study of younger patients with higher

sexual function needs may be necessary.

Conclusion

In this study, we investigated interrater agreement between clinician ratings using the Rating reference guide and patient self-assessments for the nine body function categories of the ICF Rehabilitation Set and evidenced good agreement in rating. Our findings support the feasibility of clinician ratings using the Rating reference guide for body function categories as a means of describing the functioning problems from which patients suffer.

Acknowledgments

We would like to express our gratitude to the members of the Department of Rehabilitation of Fujita Health University Hospital and the Department of Rehabilitation of Fujita Health University Nanakuri Memorial Hospital for their cooperation in this study. This research was supported by the Grants for Research on Health and Welfare for the research project titled “The establishment of a mechanism for evaluation and information sharing using the International Classification of Functioning, Disability and Health to integrate medical and long-term care services” (Principal Investigator: Masahiko Mukaino).

References

1. World Health Organization (WHO). International classification of Functioning, Disability and Health. 1st ed. Geneva; 2001.
2. Stucki G, Prodinger B, Bickenbach J. Four steps to follow when documenting functioning with the International Classification of Functioning, Disability and Health. Eur J Phys Rehabil Med 2017; 53: 144–9.
3. Selb M, Escorpizo R, Kostanjsek N, Stucki G, Ustun B, Cieza A. A guide on how to develop an International Classification of Functioning, Disability and Health Core Set. Eur J Phys Rehabil Med 2015; 51: 105–17.
4. Ewert T, Grill E, Bartholomeyczik S, Finger M, Mokrusch T, Kostanjsek N, et al. ICF Core Set for patients with neurological conditions in the acute hospital. Disabil Rehabil 2005; 27: 367–73.
5. Grill E, Ewert T, Chatterji S, Kostanjsek N, Stucki G. ICF Core Sets development for the acute hospital and early post-acute rehabilitation facilities. Disabil Rehabil 2005; 27: 361–6.
6. Grill E, Hermes R, Swoboda W, Uzarewicz C, Kostanjsek N, Stucki G. ICF Core Set for geriatric patients in early post-acute rehabilitation facilities. Disabil Rehabil 2005; 27: 411–7.
7. Wildner M, Quittan M, Portenier L, Wilke S, Boldt C, Stucki G, et al. ICF Core Set for patients with cardiopulmonary conditions in early post-acute rehabilitation facilities. Disabil Rehabil 2005; 27: 397–404.
8. Boldt C, Grill E, Wildner M, Portenier L, Wilke S, Stucki G, et al. ICF Core Set for patients with cardiopulmonary conditions in the acute hospital. Disabil Rehabil 2005; 27: 375–80.
9. 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.
10. Prodinger B, Cieza A, Oberhauser C, Bickenbach J, Ustun 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; 97: 875–84.
11. Gimigliano F, Selb M, Mukaino M, Baffone C, Bickenbach J, Patrick JE, et al. Strengthening rehabilitation in health systems worldwide by implementing information on functioning in rehabilitation practice, quality management, and policy: 2018 status report. J Int Soc Phys Rehabil Med 2018; 1: 37.
12. Frontera W, Gimigliano F, Melvin J, Li J, Li L, Lains J, et al. ClinFIT: ISPRM’s Universal Functioning Information Tool based on the WHO’s ICF. J Int Soc Phys Rehabil Med 2019; 2: 19.
13. Prodinger B, Reinhardt JD, Selb M, Stucki G, Yan T, Zhang X, et al. Towards system-wide implementation of the International Classification of Functioning, Disability and Health (ICF) in routine practice: developing simple, intuitive descriptions of ICF categories in the ICF Generic and Rehabilitation Set. J Rehabil Med 2016; 48: 508–14.
14. Selb M, Gimigliano F, Prodinger B, Stucki G, Pestelli G, Iocco M, et al. Toward an International Classification of Functioning, Disability and Health clinical data collection tool: the Italian experience of developing simple, intuitive descriptions of the Rehabilitation Set categories. Eur J Phys Rehabil Med 2017; 53: 290–8.
15. 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: 1078–84.
16. Mukaino M, Saitoh E, Otaga M. Development of ICF-based data collection tools and system for information sharing between medical services and long term care services. Health and Labor Sciences Research Grant Report 201802007A; 2019. p. 21–7. Japanese.
17. Bickenbach J. Reconciling the capability approach and the ICF. Alter 2014; 8: 10–23.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159–74.
19. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994; 6: 284.
20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009; 42: 377–81.
21. Verhulsdonk S, Quack R, Hoft B, Lange-Asschenfeldt C, Supprian T. Anosognosia and depression in patients with Alzheimer’s dementia. Arch Gerontol Geriatr 2013;

57: 282–7.
22. Conde-Sala JL, Turro-Garriga O, Pinan-Hernandez S, Portellano-Ortiz C, Vinas-Diez V, Gascon-Bayarri J, et al. Effects of anosognosia and neuropsychiatric symptoms on the quality of life of patients with Alzheimer's disease: a 24-month follow-up study. Int J Geriatr Psychiatr 2016; 31: 109–19.
23. Starkstein SE, Jorge RE, Robinson RG. The frequency, clinical correlates, and mechanism of anosognosia after stroke. Can J Psychiatry 2010; 55: 355–61.
24. Solomon P. Congruence between health professionals' and patients' pain ratings: a review of the literature. Scand J Caring Sci 2001; 15: 174–80.
25. Zastrow A, Faude V, Seyboth F, Niehoff D, Herzog W, Lowe B. Risk factors of symptom underestimation by physicians. J Psychosom Res 2008; 64: 543–51.
26. Carr EK, Kenney FD, Wilson-Barnett J, Newham DJ. Inter-rater reliability of postural observation after stroke. Clin Rehabil 1999; 13: 229–42.
27. Brunnekreef JJ, van Uden CJ, van Moorsel S, Kooloos JG. Reliability of videotaped observational gait analysis in patients with orthopedic impairments. BMC Musculoskelet Disord 2005; 6: 17.
28. Starrost K, Geyh S, Trautwein A, Grunow J, Ceballos-Baumann A, Prosiegel M, et al. Interrater reliability of the extended ICF core set for stroke applied by physical therapists. Phys Ther 2008; 88: 841–51.