Turk J Phys Med Rehab 2023;69(3):350-365
DOI: 10.5606/tftrd.2023.12387
Available online at www.turkishjournalpmr.com

Original Article

Developing a common metric using current scales for assessing functioning in patients with knee osteoarthritis

Pervin Demir[1], Ayşe Adile Küçükdeveci[2], Şehim Kutlay[2], Atilla Halil Elhan[3]

[1]Department of Biostatistics and Medical Informatics, Ankara Yıldırım Beyazıt University Faculty of Medicine, Ankara, Türkiye
[2]Department of Physical Medicine and Rehabilitation, Ankara University Faculty of Medicine, Ankara, Türkiye
[3]Department of Biostatistics, Ankara University Faculty of Medicine, Ankara, Türkiye

Received: February 27, 2023 **Accepted:** April 19, 2023 **Published online:** July 07, 2023

ABSTRACT

Objectives: Various scales exist to assess different domains of functioning in knee osteoarthritis (OA). This study aimed to explore whether it is possible to develop a common metric (CM) from the frequently used scales to assess functioning in knee OA.

Patients and methods: The methodological study evaluated 411 patients (81 males, 330 females; mean age: 61.8±10.5 years; range, 41 to 88 years) with knee OA. Data from the Health Assessment Questionnaire, Oxford Knee Score, Medical Outcomes Study Short Form 36, Knee Injury and Osteoarthritis Outcome Score, Western Ontario and McMaster Universities Arthritis Index, and the Nottingham Health Profile were used, and the items focusing on self-care, mobility, and domestic activity domains based on the activities and participation component of the International Classification of Functioning, Disability, and Health were included. Concurrent calibration was performed to combine the items of the scales. The CM parameters were estimated using the Rasch measurement model. Reliability was assessed using the person separation index. The CM was utilized to generate a transformation table to convert the scale scores to each other based on the reference metric score.

Results: Each scale fitted the Rasch model. Item invariance was achieved for the CM (p=0.775). The CM had a person separation index of 0.827. Age, sex, and disease duration did not cause difference in item functions. The CM satisfied the assumptions of unidimensionality and local independence.

Conclusion: A reliable CM was created from the commonly used scales to measure functioning in individuals with knee OA. Thus, clinicians and researchers can refer to the transformation table to directly compare scores of those scales and use them interchangeably.

Keywords: Common metric, concurrent calibration, functioning, knee osteoarthritis, Rasch model.

Musculoskeletal disorders are common in the general population and constitute a heavy burden to society. They affect the quality of life negatively, particularly in terms of pain and physical function (PF).[1] Osteoarthritis (OA) is a musculoskeletal disorder that causes pain and restriction of joint movement, which leads to disability.[2] Therefore, it is important to assess the functioning of patients with OA. The International Classification of Functioning, Disability, and Health (ICF) published by the World Health Organization defines functioning as an umbrella term encompassing body functions and structures, as well as activities and participation.[3] Functional assessment generally addresses the measurement of an individual's abilities in performing tasks necessary to daily living, leisure activities, vocational pursuits, social interactions, and other required behaviors.[4]

Corresponding author: Pervin Demir, MD. Ankara Yıldırım Beyazıt Üniversitesi Tıp Fakültesi, Biyoistatistik ve Tıbbi Bilişim Anabilim Dalı, 06800 Çankaya, Bilkent, Ankara, Türkiye. E-mail: pervindemir2010@gmail.com

Cite this article as:
Demir P, Küçükdeveci AA, Kutlay Ş, Elhan AH. Developing a common metric using current scales for assessing functioning in patients with knee osteoarthritis. Turk J Phys Med Rehab 2023;69(3):350-365. doi: 10.5606/tftrd.2023.12387.

Generic or disease-specific scales are used to assess functioning in patients with OA. There are more than 90 generic or disease-specific instruments used to assess various domains of functioning (e.g., pain, PF, work, and general health status) in OA.[5,6] The heterogeneity of the scales used in different studies makes it hard to compare the results of those studies. In addition, most of those scales are ordinal-level scales, limiting their usability in monitoring change over time. Lack of comparable and interval-level scales might hinder clinicians and researchers in using clinical data for different goals, such as patient registries, hospital statistics, benchmarking, and research purposes, including meta-analyses.[7,8] Therefore, in such cases, a scale/metric to directly compare the results of two or more scales would be helpful.[9,10] Test equating methods are used to that end. Test equating can be conducted using different data collection designs. One can perform test equating either separately or concurrently.[11] Rasch measurement model allows conversion of ordinal scales to interval-level measures and thus performs concurrent calibration of various scales on the same common metric (CM).[12]

Prodinger et al.[7] used the Rasch measurement model to create a CM from four scales commonly used to assess functioning in OA and developed an easy-to-use transformation table. Oude Voshaar et al.[13] calibrated PF results from different scales to a single metric in inflammatory rheumatic diseases. They developed a standardized PF score metric based on item response theory for 10 scale items. Prodinger et al.[8] used data from four countries to compare different scales measuring functioning in rheumatoid arthritis. These researchers noted that using different patient-reported outcome measures made it difficult to compare scale outcomes. Therefore, they developed CMs based on item response theory. To our knowledge, no studies have ever developed a CM to assess functioning in patients with knee OA. Hence, this study aimed to develop a CM from frequently used scales to assess functioning in patients with knee OA and to use the CM to compare the results of different scales. The CM will allow us to determine what score of one scale corresponds to on the other scales and thus to directly compare functioning of patients who have been evaluated by different scales.

PATIENTS AND METHODS

The methodological study was conducted with 411 patients (81 males, 330 females; mean age: 61.8±10.5 years; range, 41 to 88 years) with knee OA. The data set was composed of recently and previously collected data. The previously collected data for secondary analysis was obtained from the responses of 284 patients who filled the scales Health Assessment Questionnaire (HAQ), Western Ontario and McMaster Universities Arthritis Index (WOMAC)-PF, and Nottingham Health Profile (NHP)-physical mobility (PM).[14] The newly collected data consisted of the responses of 127 patients with knee OA treated at the Department of Physical Medicine and Rehabilitation at Ankara University Faculty of Medicine; these patients filled the HAQ, the Medical Outcomes Study Short Form 36 (SF36)-PF, the Knee Injury and Osteoarthritis Outcome Score (KOOS)-PF, and the Oxford Knee Score (OKS). The inclusion criteria were *(i)* having been diagnosed with knee OA according to the American College of Rheumatology knee OA diagnostic criteria,[15] *(ii)* being over 40 years of age, *(iii)* speaking Turkish, and *(iv)* being literate. The exclusion criteria were *(i)* having comorbidities affecting functioning and *(ii)* having knee arthroplasty surgery in the past.

Scales should be conceptually equivalent to develop a CM.[16] Therefore, this study considered the conceptual definition of the contents of the scales depending on the ICF categories. We reviewed common scales for assessing patients with knee OA and decided to use three disease-specific and three generic patient-reported outcome measures.[5,6] These scales were the HAQ,[17,18] SF36,[19,20] NHP,[21,22] KOOS,[23,24] OKS,[25,26] and WOMAC[27,28] (Table 1).

Subscales focusing on PF of the SF36, KOOS, WOMAC, and NHP, as well as all items of the HAQ and the OKS, were included in this study. The items were selected such that they would measure "self-care," "mobility," and "domestic activity" categories based on the activities and participation component of the ICF. Higher HAQ, KOOS-PF, OKS, WOMAC-PF, and NHP-PM scores indicate worse PF, while higher SF36-PF scores indicate better PF. Therefore, SF36-PF items were recorded such that high scores indicated worse functioning, similar to the other five scales. The total score for each scale was created by summing the responses to respective items.

Statistical analysis

A power analysis was performed to determine the sample size. The results showed that a sample of 384 patients would be large enough to detect significant differences (an effect size of ±0.20 logit in item difficulty estimation with 95% reliability).[29]

TABLE 1
Properties of the outcome measures used in this study

Scale name	Type-Aim	The number of items and the subscales*		Scoring items	The score range of the PF/PM domain†	Abbreviation
The Health Assessment Questionnaire (HAQ)	G - Assessing the difficulties in performing activities of daily living	**Items** **Dressing** **Arising** **Eating** **Walking** **Hygiene** **Reach** **Grip** **Activities**	**20** 2 2 3 2 3 2 3 3	0: Without any difficulty- 3: Unable to do	0-24	HAQ
The Medical Outcomes Study Short Form 36 (SF36)	G - Allows for the evaluation of one's health status	**Items** **Physical functioning** Role limitations due to physical problems Role limitations due to emotional problems Pain Social functioning Mental health Energy/fatigue General health perceptions Change in health	**36** 10 4 3 2 2 5 4 5 1	In different types as 3-5-6 point Likert, yes-no. re-coded as 0: No, not limited at all-2: Yes, limited a lot.	0-20	SF36-PF
The Knee Injury and Osteoarthritis Outcome Score (KOOS)	DS - Evaluates functional status due to knee OA	**Items** Symptoms Pain **Function-activities of daily living** Sports and recreational activities Knee related quality of life	**42** 7 9 17 5 4	0: No problem- 4: Severe problem	0-68	KOOS-PF
Oxford Knee Score (OKS)	DS - The patient with knee OA evaluates herself or himself	**Items** **Physical function & pain during activities**	**12**	0: Easily/no pain/ difficulty- 4: Impossible/severe pain/ difficulty	0-48	OKS
The Western Ontario and McMaster Universities Arthritis Index (WOMAC)	DS - Specific scale for hip and/or knee OA	**Items** Pain Stiffness **Physical function**	**24** 5 2 17	0: No difficulty/ pain-4: Extreme difficulty/pain	0-68	WOMAC-PF
The Nottingham Health Profile (NHP)	G - Aims to measure the self-perceived health status of the person in terms of physical, emotional, and social aspects	**Part 1: items** Sleep Energy Emotional reaction Social isolation **Physical mobility** Pain	**38** 5 3 9 5 8 8	No: 0-yes: Percentage no: 0-yes: 1	0-8	NHP-PM

PF: Physical function; PM: Physical mobility; DS: Disease specific; G: Generic; OA: Osteoarthritis; * The items/domains used in the study are shown in bold; † A higher score indicates poor functioning.

TABLE 2

Results of Rasch analyses for HAQ, SF36-PF, KOOS-PF, OKS, WOMAC-PF, NHP-PM scales, and CM*

	HAQ	SF36-PF	KOOS-PF	OKS	WOMAC-PF	NHP-PM	Common metric
Item location	0.000±0.865	0.000±1.183	0.000±0.530	0.000±1.018	0.000±0.603	0.000±3.766	**0.000±0.201**
Person location	-1.134±1.945	1.063±1.384	0.147±1.530	0.195±1.475	0.412±1.467	-0.073±2.318	**0.006±0.618**
Item fit residuals	0.176±0.746	-0.011±0.929	0.156±1.213	0.215±1.344	-0.140±1.468	-0.082±1.543	**-0.103±2.539**
Person fit residuals	-0.316±0.958	-0.202±0.956	-0.386±1.649	-0.277±1.347	-0.291±1.267	-0.275±0.227	**0.681±1.134**
Item trait interaction (χ^2 (df); *p*‡)	35.638 (35); 0.438	33.451 (14); 0.003	39.756 (28); 0.070	31.986 (24); 0.127	100.605 (75); 0.026	96.875 (32); <0.001	**23.935 (30); 0.775**
Person separation index	0.849	0.822	0.945	0.899	0.912	0.715	**0.827**
Tests of unidimensionality							
Number of sample	399	124	126	127	279	276	**395**
n (%) of significant t-test	24 (6.0)	6 (4.8)	12 (9.5)	11 (8.7)	18 (6.5)	5 (1.8)	**9 (2.3)**
95% CI†	3.7-8.3	1.1-8.6	4.4-14.6	3.8-13.6	3.6-9.3	0.2-3.4	**0.8-3.8**
Local dependency	✓	✓	X	✓	✓	✓	✓
DIF	✓	✓	✓	✓	✓	✓	✓

HAQ: Health Assessment Questionnaire; SF36: Short Form-36; KOOS: The Knee Injury and Osteoarthritis Outcome Score; PF: Physical function; OKS: The Oxford Knee Score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; NHP: Nottingham Health Profile; PM: Physical mobility; CM: Common metric; df: Degrees of freedom; CI: Confidence interval; DIF: Differential item functioning (was examined for age, sex, and disease duration); * The results of the final analysis, in which the Rasch model assumptions were met; ✓ Assumption was satisfied; χ: Assumption was not satisfied; ‡ If it is higher than 0.05/number of items, it indicates model fit; † If this value is lower than or equal to 5% then the unidimensionality is acceptable; The mean ± standard error was used as descriptive statistics unless indicated otherwise.

The data were analyzed using IBM SPSS version 21.0 (IBM Corp., Armonk, NY, USA) at a significance level of 0.05. Psychometric properties were examined using RUMM2020 (RUMM Laboratory, Perth, Australia).[30] Reference metrics were developed using Microsoft Office Excel 2016 (Microsoft Corp., Redmond, WA, USA). Descriptive statistics [mean ± standard deviation, median, minimum, maximum, and frequency (percentage)] were used for demographic statistics. Initially, the assumptions of the Rasch measurement model were tested in each scale or subscale. A binary Rasch model was used to examine the psychometric properties of the NHP-PM with two response categories,[31] whereas for scales with more than two response categories, the partial credit model was used according to the likelihood ratio test result.[32] The concurrent calibration method based on the internal-anchor design and item response theory was used to develop a CM.[33] The difficulty level of items (β) and the level of the examined characteristic of the individual (θ) were estimated by concurrent calibration. The partial credit model was used to obtain parameter estimates since the distance between the number of response categories and threshold values of the items in the CM was not equal. Reliability was assessed using the person separation index. The protocol followed to examine the Rasch model assumptions for CM was described in the following paragraph.[34]

First, the threshold ordering of polytomous items was performed. The category probability curves were examined to determine whether the thresholds of consecutive response categories were sequential. Second, local independence was assessed. An item residual correlation of 0.2 and above was considered local dependence. Third, data were tested for fit to the model. The Bonferroni-corrected chi-square test (item trait interaction statistics) showed that the p-value was above 0.05, indicating that the dataset fit the model and that invariance across the trait was satisfied. The model fit was also evaluated to determine whether the standardized fit residuals were within ±2.5. Fourth, residual principal components analysis was performed on the residuals to assess unidimensionality, resulting in two subscales with at least 12 thresholds. There was no difference between the means of the θ estimates obtained from the two subscales (items with positive and negative factor loadings), indicating that unidimensionality was achieved. Unidimensionality is ensured if the 95% confidence interval obtained for the percentages of subjects with differences between subscales includes 5% or its lower bound is less than 5%.

Finally, differential item functioning was evaluated. Two-way analysis of variance was used to test whether the probabilities of the response categories given to an item by individuals in different groups with similar θ levels were different. Bonferroni correction was applied for multiple testing.[35] Differential item functioning was analyzed for groups defined by the median for age (<62 years *vs.* ≥62 years), disease duration (<5 years *vs.* ≥5 years), and sex (female *vs.* male).

The compliance of the metric with the assumptions showed that the internal validity of the scale was ensured. Thus, ordinal raw scores (RS) could be transformed into logit scores (LS).

Creating common metric scores and a transformation table

Each scale combined with concurrent calibration over the response matrix was taken as a single item or subtest to develop the CM. The RS and LS were calculated for each scale or subscale. The RS is the original ordinal score of the scale. The logit is the mathematical interval unit estimated by the Rasch measurement model. First, the LS, estimated by the Rasch model, corresponding to all possible RS for CM obtained from the total number of scales of subtests (items), was recorded. Afterward, reference metric score (RMS) in the range of 0 to 100 was defined with the help of CM LS obtained by concurrent analysis. The RMS equivalents of other scale RS were calculated. A transformation table with the CM and the corresponding RS, LS, and RMS for each scale was created.

RESULTS

The median age of the participants was 62 years, and the mean of disease duration was 7.2±7.7 (median: 5; range, 0.5 to 40) years. The median of missing responses per item value for the six scales ranged from 0.0 to 3.4%.

The results of the Rasch analyses of each scale are presented in Table 2. All scales except KOOS-PF satisfied the assumption of local independence. Subtests were created for the KOOS-PF analysis to overcome this problem. All scales met the assumption of unidimensionality and fit the model. Differential item functioning was not observed in the scale items according to age, sex, and disease duration.

After the scales were shown to fit the model, item invariance was ensured for the CM obtained by concurrent equating (p=0.775, Table 2). Therefore, the CM included the items of all scales with 84 items. The CM had a person separation index of 0.827, indicating high reliability. Differential item functioning was not determined according to age, sex, and disease duration in the CM. The residual principal component analysis revealed no statistically significant differences between the subsets [number of significant test results was nine (2.3%)], indicating that unidimensionality was achieved. The local independence assumption was satisfied. The β for CM items ranged from -3.544 to 3.600 logits. The θ ranged from -2.583 to 2.850 logits. Thus, the CM contained a sufficient number of items for all levels of functioning (Figure 1).

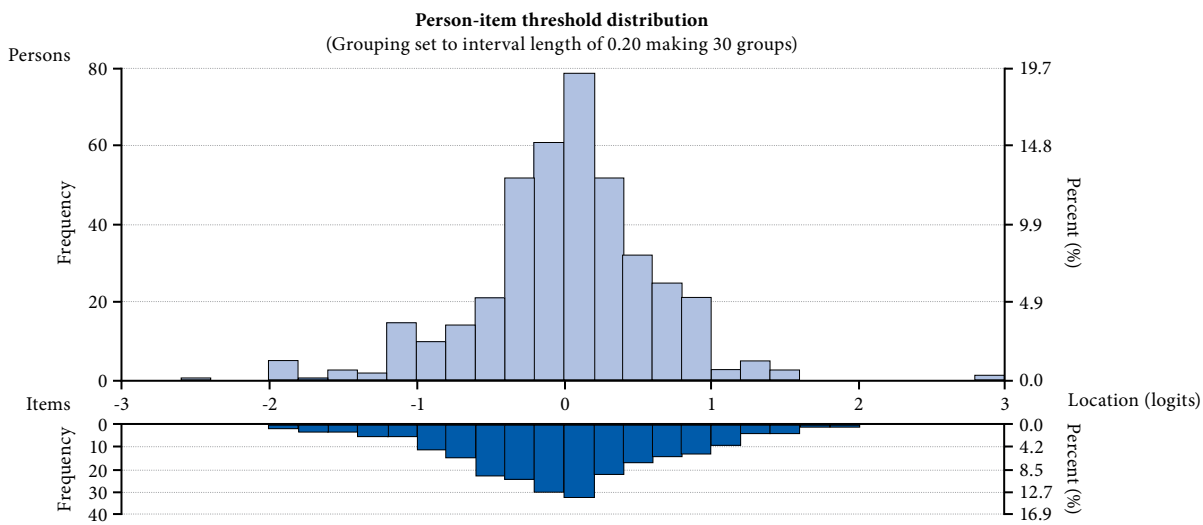

Figure 1. Person and item threshold distribution for common metric.

TABLE 3
The transformation table between outcome scales commonly used in knee OA*

Common metric			HAQ			SF36-PF			KOOS-PF			OKS			WOMAC-PF			NHP-PM		
Raw score (RS)	Logit score (LS)	Reference metric score (RMS)	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS
0	**-3.544**	**0.00**	0	-1.143	33.61	0	-1.272	31.80	0	-1.350	30.71	0	-1.865	23.50	0	-2.315	17.20	0	-2.380	16.29
1	**-2.889**	**9.17**	1	-0.834	37.93	1	-1.013	35.43	1	-1.132	33.76	1	-1.538	28.08	1	-2.004	21.56	1	-1.817	24.17
2	**-2.489**	**14.77**	2	-0.622	40.90	2	-0.844	37.79	2	-0.989	35.76	2	-1.316	31.19	2	-1.788	24.58	2	-1.313	31.23
3	**-2.246**	**18.17**	3	-0.477	42.93	3	-0.735	39.32	3	-0.895	37.08	3	-1.164	33.31	3	-1.639	26.67	3	-0.823	38.09
4	**-2.075**	**20.56**	4	-0.363	44.53	4	-0.648	40.54	4	-0.825	38.06	4	-1.044	34.99	4	-1.522	28.30	4	-0.228	46.42
5	**-1.944**	**22.40**	5	-0.264	45.91	5	-0.575	41.56	5	-0.765	38.90	5	-0.942	36.42	5	-1.418	29.76	5	0.441	55.78
6	**-1.839**	**23.87**	6	-0.173	47.19	6	-0.51	42.47	6	-0.715	39.60	6	-0.851	37.70	6	-1.325	31.06	6	1.070	64.59
7	**-1.750**	**25.11**	7	-0.087	48.39	7	-0.451	43.30	7	-0.669	40.24	7	-0.769	38.84	7	-1.239	32.26	7	1.713	73.59
8	**-1.673**	**26.19**	8	-0.003	49.57	8	-0.391	44.13	8	-0.628	40.82	8	-0.693	39.91	8	-1.157	33.41	8	2.421	83.50
9	**-1.606**	**27.13**	9	0.078	50.70	9	-0.332	44.96	9	-0.589	41.36	9	-0.623	40.89	9	-1.078	34.52			
10	**-1.545**	**27.98**	10	0.160	51.85	10	-0.268	45.86	10	-0.553	41.87	10	-0.559	41.78	10	-1.002	35.58			
11	**-1.489**	**28.77**	11	0.243	53.01	11	-0.201	46.79	11	-0.518	42.36	11	-0.500	42.61	11	-0.930	36.59			
12	**-1.438**	**29.48**	12	0.326	54.17	12	-0.128	47.82	12	-0.485	42.82	12	-0.446	43.37	12	-0.862	37.54			
13	**-1.391**	**30.14**	13	0.411	55.36	13	-0.047	48.95	13	-0.454	43.25	13	-0.396	44.06	13	-0.797	38.45			
14	**-1.347**	**30.75**	14	0.498	56.58	14	0.047	50.27	14	-0.423	43.69	14	-0.351	44.69	14	-0.736	39.31			
15	**-1.305**	**31.34**	15	0.588	57.84	15	0.157	51.81	15	-0.394	44.09	15	-0.31	45.27	15	-0.678	40.12			
16	**-1.266**	**31.89**	16	0.681	59.14	16	0.285	53.60	16	-0.365	44.50	16	-0.271	45.81	16	-0.623	40.89			
17	**-1.229**	**32.40**	17	0.777	60.48	17	0.440	55.77	17	-0.338	44.88	17	-0.234	46.33	17	-0.571	41.62			
18	**-1.194**	**32.89**	18	0.878	61.90	18	0.637	58.52	18	-0.312	45.24	18	-0.200	46.81	18	-0.523	42.29			
19	**-1.161**	**33.36**	19	0.986	63.41	19	0.923	62.53	19	-0.286	45.60	19	-0.168	47.26	19	-0.477	42.93			
20	**-1.129**	**33.80**	20	1.104	65.06	20	1.340	68.37	20	-0.261	45.95	20	-0.137	47.69	20	-0.434	43.53			
21	**-1.099**	**34.22**	21	1.239	66.95				21	-0.237	46.29	21	-0.107	48.11	21	-0.394	44.09			
22	**-1.070**	**34.63**	22	1.409	69.33				22	-0.214	46.61	22	-0.078	48.52	22	-0.356	44.62			
23	**-1.043**	**35.01**	23	1.651	72.72				23	-0.192	46.92	23	-0.049	48.92	23	-0.321	45.11			
24	**-1.016**	**35.39**	24	1.997	77.56				24	-0.170	47.23	24	-0.019	49.34	24	-0.288	45.58			
25	**-0.991**	**35.74**							25	-0.148	47.54	25	0.009	49.73	25	-0.257	46.01			
26	**-0.967**	**36.07**							26	-0.127	47.83	26	0.039	50.15	26	-0.228	46.42			
27	**-0.944**	**36.39**							27	-0.106	48.12	27	0.070	50.59	27	-0.201	46.79			
28	**-0.922**	**36.70**							28	-0.086	48.40	28	0.101	51.02	28	-0.175	47.16			
29	**-0.900**	**37.01**							29	-0.065	48.70	29	0.135	51.50	29	-0.150	47.51			
30	**-0.880**	**37.29**							30	-0.045	48.98	30	0.170	51.99	30	-0.127	47.83			
31	**-0.860**	**37.57**							31	-0.026	49.24	31	0.208	52.52	31	-0.105	48.14			
32	**-0.840**	**37.85**							32	-0.006	49.52	32	0.248	53.08	32	-0.084	48.43			
33	**-0.822**	**38.10**							33	0.014	49.80	33	0.292	53.70	33	-0.063	48.73			
34	**-0.804**	**38.35**							34	0.034	50.08	34	0.339	54.35	34	-0.042	49.02			

TABLE 3
Continued

Common metric			HAQ			SF36-PF			KOOS-PF			OKS			WOMAC-PF			NHP-PM		
Raw score (RS)	Logit score (LS)	Reference metric score (RMS)	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS
35	**-0.787**	**38.59**							35	0.054	50.36	35	0.392	55.10	35	-0.022	49.30			
36	**-0.770**	**38.83**							36	0.075	50.66	36	0.449	55.89	36	-0.004	49.55			
37	**-0.753**	**39.07**							37	0.096	50.95	37	0.511	56.76	37	0.015	49.82			
38	**-0.737**	**39.29**							38	0.117	51.25	38	0.580	57.73	38	0.034	50.08			
39	**-0.722**	**39.50**							39	0.138	51.54	39	0.654	58.76	39	0.052	50.34			
40	**-0.707**	**39.71**							40	0.160	51.85	40	0.734	59.88	40	0.072	50.62			
41	**-0.692**	**39.92**							41	0.183	52.17	41	0.821	61.10	41	0.091	50.88			
42	**-0.677**	**40.13**							42	0.206	52.49	42	0.914	62.40	42	0.110	51.15			
43	**-0.663**	**40.33**							43	0.230	52.83	43	1.016	63.83	43	0.130	51.43			
44	**-0.650**	**40.51**							44	0.255	53.18	44	1.129	65.41	44	0.150	51.71			
45	**-0.636**	**40.71**							45	0.281	53.54	45	1.261	67.26	45	0.171	52.00			
46	**-0.623**	**40.89**							46	0.307	53.91	46	1.425	69.55	46	0.192	52.30			
47	**-0.610**	**41.07**							47	0.335	54.30	47	1.663	72.89	47	0.214	52.60			
48	**-0.597**	**41.25**							48	0.364	54.70	48	2.009	77.73	48	0.237	52.93			
49	**-0.585**	**41.42**							49	0.394	55.12				49	0.262	53.28			
50	**-0.573**	**41.59**							50	0.425	55.56				50	0.288	53.64			
51	**-0.561**	**41.76**							51	0.458	56.02				51	0.315	54.02			
52	**-0.549**	**41.92**							52	0.492	56.49				52	0.344	54.42			
53	**-0.537**	**42.09**							53	0.527	56.98				53	0.375	54.86			
54	**-0.526**	**42.25**							54	0.564	57.50				54	0.407	55.31			
55	**-0.515**	**42.40**							55	0.602	58.03				55	0.442	55.80			
56	**-0.504**	**42.55**							56	0.642	58.59				56	0.479	56.31			
57	**-0.493**	**42.71**							57	0.683	59.17				57	0.518	56.86			
58	**-0.482**	**42.86**							58	0.726	59.77				58	0.560	57.45			
59	**-0.472**	**43.00**							59	0.771	60.40				59	0.604	58.06			
60	**-0.461**	**43.16**							60	0.818	61.06				60	0.652	58.73			
61	**-0.451**	**43.30**							61	0.868	61.76				61	0.702	59.43			
62	**-0.441**	**43.44**							62	0.922	62.51				62	0.758	60.22			
63	**-0.431**	**43.58**							63	0.981	63.34				63	0.820	61.09			
64	**-0.421**	**43.72**							64	1.049	64.29				64	0.890	62.07			
65	**-0.411**	**43.85**							65	1.131	65.44				65	0.975	63.26			
66	**-0.401**	**43.99**							66	1.237	66.92				66	1.086	64.81			
67	**-0.392**	**44.12**							67	1.397	69.16				67	1.253	67.15			
68	**-0.382**	**44.26**							68	1.637	72.52				68	1.504	70.66			
69	**-0.373**	**44.39**																		

TABLE 3
Continued

Common metric			HAQ			SF36-PF			KOOS-PF			OKS			WOMAC-PF			NHP-PM		
Raw score (RS)	Logit score (LS)	Reference metric score (RMS)	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS
70	**-0.364**	**44.51**																		
71	**-0.355**	**44.64**																		
72	**-0.346**	**44.76**																		
73	**-0.337**	**44.89**																		
74	**-0.328**	**45.02**																		
75	**-0.319**	**45.14**																		
76	**-0.311**	**45.25**																		
77	**-0.302**	**45.38**																		
78	**-0.294**	**45.49**																		
79	**-0.285**	**45.62**																		
80	**-0.277**	**45.73**																		
81	**-0.269**	**45.84**																		
82	**-0.261**	**45.95**																		
83	**-0.253**	**46.07**																		
84	**-0.245**	**46.18**																		
85	**-0.237**	**46.29**																		
86	**-0.229**	**46.40**																		
87	**-0.221**	**46.51**																		
88	**-0.214**	**46.61**																		
89	**-0.206**	**46.72**																		
90	**-0.198**	**46.84**																		
91	**-0.191**	**46.93**																		
92	**-0.183**	**47.05**																		
93	**-0.176**	**47.14**																		
94	**-0.169**	**47.24**																		
95	**-0.161**	**47.35**																		
96	**-0.154**	**47.45**																		
97	**-0.147**	**47.55**																		
98	**-0.140**	**47.65**																		
99	**-0.133**	**47.75**																		
100	**-0.126**	**47.84**																		
101	**-0.119**	**47.94**																		
102	**-0.112**	**48.04**																		
103	**-0.105**	**48.14**																		
104	**-0.098**	**48.24**																		

TABLE 3
Continued

Common metric			HAQ			SF36-PF			KOOS-PF			OKS			WOMAC-PF			NHP-PM		
Raw score (RS)	Logit score (LS)	Reference metric score (RMS)	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS
105	**-0.091**	**48.33**																		
106	**-0.084**	**48.43**																		
107	**-0.078**	**48.52**																		
108	**-0.071**	**48.61**																		
109	**-0.064**	**48.71**																		
110	**-0.057**	**48.81**																		
111	**-0.051**	**48.89**																		
112	**-0.044**	**48.99**																		
113	**-0.037**	**49.09**																		
114	**-0.031**	**49.17**																		
115	**-0.024**	**49.27**																		
116	**-0.017**	**49.37**																		
117	**-0.010**	**49.47**																		
118	**-0.004**	**49.55**																		
119	**0.003**	**49.65**																		
120	**0.009**	**49.73**																		
121	**0.016**	**49.83**																		
122	**0.023**	**49.93**																		
123	**0.029**	**50.01**																		
124	**0.036**	**50.11**																		
125	**0.043**	**50.21**																		
126	**0.049**	**50.29**																		
127	**0.056**	**50.39**																		
128	**0.063**	**50.49**																		
129	**0.070**	**50.59**																		
130	**0.076**	**50.67**																		
131	**0.083**	**50.77**																		
132	**0.090**	**50.87**																		
133	**0.097**	**50.97**																		
134	**0.104**	**51.06**																		
135	**0.111**	**51.16**																		
136	**0.118**	**51.26**																		
137	**0.125**	**51.36**																		
138	**0.132**	**51.46**																		
139	**0.139**	**51.55**																		

TABLE 3
Continued

Common metric			HAQ			SF36-PF			KOOS-PF			OKS			WOMAC-PF			NHP-PM		
Raw score (RS)	Logit score (LS)	Reference metric score (RMS)	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS
140	**0.146**	**51.65**																		
141	**0.153**	**51.75**																		
142	**0.160**	**51.85**																		
143	**0.168**	**51.96**																		
144	**0.175**	**52.06**																		
145	**0.182**	**52.16**																		
146	**0.190**	**52.27**																		
147	**0.198**	**52.38**																		
148	**0.205**	**52.48**																		
149	**0.213**	**52.59**																		
150	**0.221**	**52.70**																		
151	**0.229**	**52.81**																		
152	**0.237**	**52.93**																		
153	**0.245**	**53.04**																		
154	**0.253**	**53.15**																		
155	**0.261**	**53.26**																		
156	**0.269**	**53.37**																		
157	**0.278**	**53.50**																		
158	**0.286**	**53.61**																		
159	**0.295**	**53.74**																		
160	**0.304**	**53.86**																		
161	**0.312**	**53.98**																		
162	**0.321**	**54.10**																		
163	**0.330**	**54.23**																		
164	**0.340**	**54.37**																		
165	**0.349**	**54.49**																		
166	**0.358**	**54.62**																		
167	**0.368**	**54.76**																		
168	**0.377**	**54.89**																		
169	**0.387**	**55.03**																		
170	**0.397**	**55.17**																		
171	**0.407**	**55.31**																		
172	**0.417**	**55.45**																		
173	**0.428**	**55.60**																		
174	**0.438**	**55.74**																		

TABLE 3
Continued

Common metric			HAQ			SF36-PF			KOOS-PF			OKS			WOMAC-PF			NHP-PM		
Raw score (RS)	Logit score (LS)	Reference metric score (RMS)	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS
175	**0.449**	**55.89**																		
176	**0.460**	**56.05**																		
177	**0.470**	**56.19**																		
178	**0.482**	**56.35**																		
179	**0.493**	**56.51**																		
180	**0.504**	**56.66**																		
181	**0.516**	**56.83**																		
182	**0.527**	**56.98**																		
183	**0.539**	**57.15**																		
184	**0.551**	**57.32**																		
185	**0.564**	**57.50**																		
186	**0.576**	**57.67**																		
187	**0.589**	**57.85**																		
188	**0.601**	**58.02**																		
189	**0.614**	**58.20**																		
190	**0.628**	**58.40**																		
191	**0.641**	**58.58**																		
192	**0.655**	**58.78**																		
193	**0.669**	**58.97**																		
194	**0.683**	**59.17**																		
195	**0.697**	**59.36**																		
196	**0.712**	**59.57**																		
197	**0.726**	**59.77**																		
198	**0.742**	**59.99**																		
199	**0.757**	**60.20**																		
200	**0.773**	**60.43**																		
201	**0.789**	**60.65**																		
202	**0.805**	**60.88**																		
203	**0.822**	**61.11**																		
204	**0.840**	**61.37**																		
205	**0.857**	**61.60**																		
206	**0.876**	**61.87**																		
207	**0.894**	**62.12**																		
208	**0.914**	**62.40**																		
209	**0.933**	**62.67**																		

TABLE 3
Continued

Common metric			HAQ			SF36-PF			KOOS-PF			OKS			WOMAC-PF			NHP-PM		
Raw score (RS)	Logit score (LS)	Reference metric score (RMS)	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS	RS	LS	RMS
210	**0.954**	**62.96**																		
211	**0.975**	**63.26**																		
212	**0.997**	**63.56**																		
213	**1.020**	**63.89**																		
214	**1.044**	**64.22**																		
215	**1.069**	**64.57**																		
216	**1.095**	**64.94**																		
217	**1.123**	**65.33**																		
218	**1.152**	**65.73**																		
219	**1.183**	**66.17**																		
220	**1.216**	**66.63**																		
221	**1.251**	**67.12**																		
222	**1.289**	**67.65**																		
223	**1.329**	**68.21**																		
224	**1.374**	**68.84**																		
225	**1.423**	**69.53**																		
226	**1.477**	**70.28**																		
227	**1.537**	**71.12**																		
228	**1.606**	**72.09**																		
229	**1.686**	**73.21**																		
230	**1.779**	**74.51**																		
231	**1.892**	**76.09**																		
232	**2.035**	**78.09**																		
233	**2.220**	**80.68**																		
234	**2.488**	**84.43**																		
235	**2.916**	**90.43**																		
236	**3.600**	**100.00**																		

OA: Osteoarthritis; HAQ: Health Assessment Questionnaire; SF36: Short Form-36; KOOS: The Knee Injury and Osteoarthritis Outcome Score; PF: Physical function; OKS: The Oxford Knee Score; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; NHP: Nottingham health profile; PM: Physical mobility; * Logit score (LS), Raw score (RS), Reference metric score (RMS)- ranging from 0 to 100 based on the common metric (0: no problems - 100: severe problems). A higher score indicates poor functioning.

The HAQ, SF36-PF, KOOS-PF, OKS, WOMAC-PF, and NHP-PM had a maximum RS of 24, 20, 68, 48, 68, and 8, respectively. The CM had a maximum RS of 236. The RMS values between 0 and 100 from the CM and the CM equivalents of other scale RS are presented in Table 3. The Figure 2 shows the operational ranges of the scales in logits along with the interval-level CM. The HAQ, SF36-PF, and KOOS-PF better identify those with RMS values above 30, while the WOMAC-PF better identifies those with RMS values below 70.

The equivalent of an RS from one scale in other scales can be obtained using the corresponding RMS on the reference metric scale (Figure 2a), which is defined in the CM in the range of 0 to 100. It can also be obtained with the help of the corresponding LS on the logit scale (Figure 2b), which is defined in the logit range of -3.544 to 3.600. An RMS or LS equivalent to the RS of one scale was determined for the other scales, and the closest value was taken in the absence of the same value. Some examples of the interpretation of the transformation table (Table 3) are given below for clarity and comprehensibility.

An individual with 16 RS (or 0.681 LS) on the HAQ scale has an RMS of 59.14 on the CM (Figure 2a). The RS values corresponding to an RMS of approximately 59.14 in the CM for the SF36-PF, KOOS-PF, OKS, WOMAC-PF, and NHP-PM are 18, 57, 39, 61, and 5, respectively. The RMS equivalent of 57 RS from the KOOS-PF scale is 59.17 in the CM. The RS values corresponding to an approximate RMS value of 59.17 in the CM for the HAQ, SF36-PF, OKS, WOMAC-PF, and NHP-PM are 16, 18, 39, 61, and 5, respectively. The RMS equivalent of 7 RS from the OKS scale is 38.84 in the CM. The RS values corresponding to approximately 38.84 RMS in the CM for the HAQ, SF36-PF, KOOS-PF, WOMAC-PF, and NHP-PM are 1, 3, 5, 13, and 3, respectively. The RMS equivalent of 3 RS from the SF36-PF scale is 39.32 in the CM. The RS values corresponding to approximately 39.32 RMS value in the CM for the HAQ, KOOS-PF, OKS, WOMAC-PF, and NHP-PM are 1, 5, 7, 14, and 3, respectively. The RMS equivalent of 38 RS from the WOMAC-PF scale is 50.08 in the CM. The RS values corresponding to an RMS of approximately 50.08 in the CM for the HAQ, SF36-PF, KOOS-PF, OKS, and NHP-PM are 8, 14, 34, 26, and 4, respectively. The RMS equivalent of 8 RS from the NHP-PM scale is 83.50 in the CM. The RS values corresponding to an RMS value of approximately 83.50 in the CM for the HAQ, SF36-PF, KOOS-PF, OKS, and WOMAC-PF are 24, 20, 68, 48, and 68, respectively.

DISCUSSION

In the present study, we developed a reliable CM based on the Rasch model containing sufficient items from the six common scales that assess PF in terms of self-care, mobility, and domestic activities in patients with knee OA. We also created an easy-to-use standardized transformation table to ensure direct comparability of the scores. Thus, clinicians and researchers can directly compare their data and levels of functioning by the scales used in this study through the transformation table. In addition, this CM will enable the transformation of patient scores evaluated at different settings with different scales and therefore help with follow-up of the individual patient in terms of functioning.

The CM had an item difficulty estimates of -3.5 to 3.6 logits, while it had a person ability estimate of

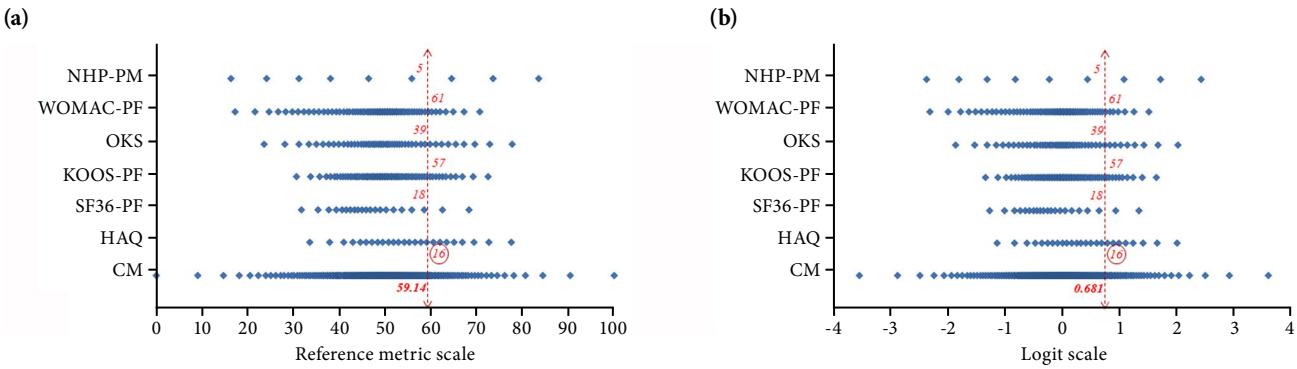

Figure 2. (a, b) The widths of each scale along the reference metric ranging from 0 to 100 (from -3.5 to 3.6 logits) based on the CM. (Dashed line shows the expected raw scores of an individual with a raw score of 16 on the HAQ scale in other scales).

NHP: Nottingham health profile; PM: Physical mobility; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; PF: Physical function; OKS: The Oxford Knee Score; KOOS: The Knee Injury and Osteoarthritis Outcome Score; SF36: Short Form-36; HAQ: Health Assessment Questionnaire; CM: Common metric.

-2.6 to 2.9 logits. These results showed that the CM developed in this study covered a greater location of functioning compared to each scale. The range of variation over both the logit scale and reference metric scale for other scales showed that the disease-specific scales (WOMAC-PF, KOOS-PF, and OKS) provided a wider range of prediction and more precise measurements than the generic scales (HAQ, SF36-PF, and NHP-PM).

Lundgren-Nilsson et al.[6] conducted a systematic review of the use and psychometric properties of 78 patient-reported scales in OA. It is important to make a transformation of the scale scores to provide standardization in patient follow-up and to compare results without changing the measurement tools used in daily practice and research. The CM developed in this study ensures that the original RS of the ordinal scales are sufficient to estimate patients' level of functioning[36] as we can make comparisons among converted scales. The commonly used scales in OA were reported to be WOMAC, SF36, KOOS, and OKS.[6] As all these scales are included in our study, the transformation table we created can have a potential for widespread use in clinical practice and research.

Meta-analyses integrate the results of several independent studies and are considered to provide the top level of evidence in scientific research. However, researchers conducting a meta-analysis may encounter different scale scores used for the same purpose in different studies. de Rooij et al.[37] conducted a systematic review and meta-analysis on functioning and pain in patients with knee OA. They reported that despite the large number of studies identified, few studies were included in the meta-analyses because different scales and metrics were used to assess the outcome. Similarly, Raposo et al.[38] showed that different scales were used by studies in their review. These studies address that the effect size of some interventions was unavailable due to the variety of scales used to assess the same construct in the same group of patients. Thus, if we want to perform a systematic review of functioning in patients with knee OA, a direct comparison of results may not be possible since different researchers use different scales. In that case, the CM and the transformation table created in this study for the relevant scales can be used to convert the scale scores.

One limitation of the study is that the scales except the HAQ were not answered by all patients. However, the use of Rasch model eliminates this problem. Another limitation is that this study was conducted in a single center. Therefore, future multicenter studies are needed to confirm our results. Finally, the differential item functioning assessment for sex was not adequate due to the unbalanced distribution (80.3% of the participants were female). The strength of the study is that the content of the CM is based on the ICF and includes PF in terms of self-care, mobility, and domestic activities. Additionally, the distribution of the patients' functioning levels was well calibrated on the CM, which means that the CM is able to measure a wider range of the patients' ability levels compared to each scale.

In conclusion, we created a CM including six common scales/subscales for assessing functioning in terms of self-care, mobility, and domestic activities in patients with knee OA. This CM covers a greater location of functioning compared to each scale assessed in the study, and together with the transformation table, it allows interchangeability of those scale scores. Thus, clinicians and researchers can refer to the transformation table to directly compare scores of those scales and use them interchangeably.

Ethics Committee Approval: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the clinical research ethics committee of Ankara University Faculty of Medicine (date: 25.09.2017, no: 15-952-17). The approval no for the previously collected data is 138-3999 (Project no: BAP 0809241).

Patient Consent for Publication: A written informed consent was obtained from each patient.

Data Sharing Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions: Had access to all data, literature review: P.D.; Study design, data collection, writing the article, and interpretation of results: P.D., A.A.K., Ş.K., A.H.E.; All authors critically evaluated the manuscript and approved the final version to be submitted for publication.

Conflict of Interest: The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding: The authors received no financial support for the research and/or authorship of this article.

REFERENCES

1. Reginster JY. The prevalence and burden of arthritis. Rheumatology (Oxford) 2002;41 Supp 1:3-6.
2. Michael JW, Schlüter-Brust KU, Eysel P. The epidemiology, etiology, diagnosis, and treatment of osteoarthritis of the knee. Dtsch Arztebl Int 2010;107:152-62. doi: 10.3238/arztebl.2010.0152.

3. World Health Organization. International classification of functioning, disablity and health (ICF). Geneva: World Health Organization; 2001.
4. Küçükdeveci AA. Osteoartritte işlevsel değerlendirme ölçütleri. Turk J Geriatr 2011;14 Supplement 4:37-44.
5. Aaos Project Team. Osteoarthritis: Function&Pain Assessment Measure Methodology Report. AAOS 2017. Available at: https://www.aaos.org/globalassets/quality-and-practice-resources/osteoarthritis-of-the-knee/osteoarthritis-function-and-pain-assessment---final-report-approved-by-bod-sept-2017.pdf [Acccessed: 22.05.2019].
6. Lundgren-Nilsson Å, Dencker A, Palstam A, Person G, Horton MC, Escorpizo R, et al. Patient-reported outcome measures in osteoarthritis: A systematic search and review of their use and psychometric properties. RMD Open 2018;4:e000715. doi: 10.1136/rmdopen-2018-000715.
7. Prodinger B, Stamm T, Peterson D, Stucki G, Tennant A; International Classification of Functioning, Disability, and Health Info Network. Toward a standardized reporting of outcomes in hand osteoarthritis: Developing a common metric of outcome measures commonly used to assess functioning. Arthritis Care Res (Hoboken) 2016;68:1115-27. doi: 10.1002/acr.22816.
8. Prodinger B, Coenen M, Hammond A, Küçükdeveci AA, Tennant A. Scale banking for patient-reported outcome measures that measure functioning in rheumatoid arthritis: A daily activities metric. Arthritis Care Res (Hoboken) 2022;74:579-87. doi: 10.1002/acr.24503.
9. Angoff WH. Scales, norms and equivalent scores. princeton, New Jersey: Educational Testing Service; 1984.
10. Kolen M, Brennan R. Test equating, scaling and linking: Methods and practices. 2nd ed. New York: Springer-Vergal; 2004.
11. Kim SH, Cohen AS. A comparison of linking and concurrent calibration under item response theory. Appl Psychol Meas 1998;22:131-43. doi: 10.1177/01466216980222003.
12. Luo G, Seow A, Chin CL. Linking and anchoring techniques in test equating using the Rasch model. In: Proceedings of the 5th CAA International Computer Assisted Assessment Conference, Loughborough, UK; 2001.
13. Oude Voshaar MAH, Vonkeman HE, Courvoisier D, Finckh A, Gossec L, Leung YY, et al. Towards standardized patient reported physical function outcome reporting: Linking ten commonly used questionnaires to a common metric. Qual Life Res 2019;28:187-97. doi: 10.1007/s11136-018-2007-0.
14. Öztuna D, Elhan AH, Küçükdeveci AA, Kutlay S, Tennant A. An application of computerised adaptive testing for measuring health status in patients with knee osteoarthritis. Disabil Rehabil 2010;32:1928-38. doi: 10.3109/09638281003777572.
15. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum 1986;29:1039-49. doi: 10.1002/art.1780290816.
16. Prodinger B, Tennant A, Stucki G. Standardized reporting of functioning information on ICF-based common metrics. Eur J Phys Rehabil Med 2018;54:110-7. doi: 10.23736/S1973-9087.17.04784-0.
17. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45. doi: 10.1002/art.1780230202.
18. Küçükdeveci AA, Sahin H, Ataman S, Griffiths B, Tennant A. Issues in cross-cultural validity: Example from the adaptation, reliability, and validity testing of a Turkish version of the Stanford Health Assessment Questionnaire. Arthritis Rheum 2004;51:14-9. doi: 10.1002/art.20091.
19. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
20. Kocyigit H, Aydemir O, Fisek G, Olmez N, Memis A. Validity and reliability of Turkish version of Short form 36: A study of a patients with romatoid disorder. Drug Therapy 1999;12:102-6.
21. The Europen Group for Quality of Life Assessment and Health Measurement: European Guide to the Nottingham Health Profile. Surrey: Brookwood Medical Publications; 1983. p. 1-16.
22. Kücükdeveci AA, McKenna SP, Kutlay S, Gürsel Y, Whalley D, Arasil T. The development and psychometric assessment of the Turkish version of the Nottingham Health Profile. Int J Rehabil Res 2000;23:31-8. doi: 10.1097/00004356-200023010-00004.
23. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88-96. doi: 10.2519/jospt.1998.28.2.88.
24. Paker N, Buğdayci D, Sabirli F, Özel S, Ersoy S. Knee injury and osteoarthritis outcome score: Reliability and validation of the Turkish version. Turkiye Klinikleri J Med Sci 2007:27:350-6.
25. Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford hip and knee scores. J Bone Joint Surg [Br] 2007;89:1010-4. doi: 10.1302/0301-620X.89B8.19424.
26. Tuğay BU, Tuğay N, Güney H, Kınıklı Gİ, Yüksel İ, Atilla B. Oxford Knee Score: Cross-cultural adaptation and validation of the Turkish version in patients with osteoarthritis of the knee. Acta Orthop Traumatol Turc 2016;50:198-206. doi: 10.3944/AOTT.2015.15.0127.
27. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833-40.
28. Tüzün EH, Eker L, Aytar A, Daşkapan A, Bayramoğlu M. Acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC osteoarthritis index. Osteoarthritis Cartilage 2005;13:28-33. doi: 10.1016/j.joca.2004.10.010.
29. Linacre JM. Sample size and item calibration stability. Rasch Measurement Transactions 1994;7:328.

30. Andrich D, Lyne A, Sheridan B, Luo G. RUMM2020. Rasch Unidimensional Measurement Models Software. Perth: RUMM Laboratory; 2003.
31. Rasch G. Probabilistic models for Some Intelligence and Attainment Tests. Chicago: University of Chicago Press; 1960.
32. Masters GN. A rasch model for partial credit scoring. Psychometrika 1982;47:149-74. doi: 10.1007/BF02296272.
33. Demir P. Evaluation the factors affecting common item selection to develop parallel test by using simulation study. [Master Thesis]. Ankara University, Department of Biostatistics; 2014.
34. 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;57:1358-62. doi: 10.1002/art.23108.
35. Bland JM, Altman DG. Multiple significance tests: The Bonferroni method. BMJ 1995;310:170. doi: 10.1136/bmj.310.6973.170.
36. Andrich D. Rating scales and Rasch measurement. Expert Rev Pharmacoecon Outcomes Res 2011;11:571-85. doi: 10.1586/erp.11.59.
37. de Rooij M, van der Leeden M, Heymans MW, Holla JF, Häkkinen A, Lems WF, et al. Prognosis of pain and physical functioning in patients with knee osteoarthritis: A systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2016;68:481-92. doi: 10.1002/acr.22693.
38. Raposo F, Ramos M, Lúcia Cruz A. Effects of exercise on knee osteoarthritis: A systematic review. Musculoskeletal Care 2021;19:399-435. doi: 10.1002/msc.1538.