

Original Research

Evaluating Psychometric Properties of the International Knee Documentation Committee Subjective Knee Form in a Heterogeneous Sample of Post-Operative Patients

Ryan D Richardson¹, Madeline P Casanova¹ , Ashley J Reeves¹, Seungho Ryu¹, Adam C Cady², Russell T Baker¹ ^a¹ WWAMI Medical Education, University of Idaho, ² Woodland Hills Orthopedics, Kaiser Permanente

Keywords: structural validity, psychometric analysis, knee pathology, patient reported outcomes

<https://doi.org/10.26603/001c.83940>

International Journal of Sports Physical Therapy

Vol. 18, Issue 4, 2023

Background

The International Knee Document Committee Subjective Knee Form (IKDC-SKF) is a patient-reported outcome measure used in orthopedics and sports medicine. Further psychometric assessment is necessary to confirm measurement properties in a large, heterogeneous sample.

Purpose

The purpose of the study was to assess the psychometric properties of the IKDC-SKF in a large, heterogeneous sample.

Study Design

Cross-Sectional Study

Methods

An exploratory factor analysis (EFA) was conducted to identify a sound latent structure and to assess internal consistency in a large sample of patients who underwent knee arthroscopy. A confirmatory factor analysis (CFA) was conducted to confirm structural validity. Multi-group invariance was conducted to assess factorial stability across sex and age groups, while longitudinal invariance procedures were performed to assess stability over time.

Results

A 3-factor, 9-item IKDC-SKF short form was identified with EFA procedures. The model was confirmed with CFA (CFI = 0.983; TLI = 0.975; IFI = 0.983; RMSEA = 0.057), while a sound 2-factor, 6-item model was also identified (CFI = 1.0; TLI = 0.999; IFI = 1.0; RMSEA = 0.11). The 9-item IKDC-SKF short form was invariant across groups but not time; removal of a single item (i.e., 8-item IKDC-SKF short form) resulted in longitudinal invariance. The 6-item IKDC-SKF short form was invariant across groups and time.

Conclusion

The 6-item, 8-item, and 9-item short form versions of the IKDC-SKF exceed contemporary fit recommendations and present as plausible alternatives to the IKDC-SKF with improved measurement properties, reduced scale response burden, and evidence of multi-group and longitudinal invariance. Further, the 6- and 8-item IKDC-SKF short forms may be used to assess group differences or change across time.

^a Corresponding Author:
Russell T. Baker
875 Perimeter Drive,
Moscow, Idaho 83844
208-885-2065
russellb@uidaho.edu

Level of evidence here

Level 3

INTRODUCTION

The International Knee Documentation Committee (IKDC) Subjective Knee Form (SKF) is a 19-item joint-specific patient-reported outcome measure (PROM) used in orthopedics and sports medicine, with applications in clinical practice and research.¹ The IKDC-SKF is intended to be used across a variety of knee pathologies (e.g., ACL injury,² meniscal injury,³ cartilage damage,⁴ patellofemoral pain syndrome⁵) and has been translated into many languages (e.g., Chinese,⁶ Arabic,⁷ German⁸). Initial assessment of scale properties indicated the English version of the IKDC-SKF had adequate test-retest reliability (ICC range = 0.85–0.99)^{9–11} but large ceiling effects,¹² while translated versions have been reported to have good responsiveness (i.e., change can be detected) without floor or ceiling effects.^{8,10,11} However, other analysis procedures necessary to establish the measurement properties of the IKDC-SKF for use in clinical practice and research are either lacking (e.g., multi-group invariance testing) or have identified potential concerns with the scale.

For example, internal consistency, a measure of how similar the items are in a unidimensional scale or unique subscale, should be calculated for each construct included in a scale.^{13–16} High alpha levels (i.e., ≥ 0.90) have been interpreted as evidence of strong internal consistency; however, these high values may be more likely to indicate item redundancy, inclusion of too many items or parallel items, construct underrepresentation, or reduced construct precision.^{14,15,17} Low values (i.e., $\leq .70$), in contrast, indicate poor internal consistency within a scale or construct. While alpha values ≥ 0.70 and < 0.90 are often considered acceptable,^{13,14} a range of ≥ 0.80 and < 0.90 has been recommended for scale development.^{15,17} Reported Cronbach's alpha values for the IKDC-SKF have ranged from 0.77 to 0.97 across different versions of the scale.^{8–10,18} The reported Cronbach's alpha values outside the recommended range, particularly those well above 0.90, along with those calculated for the entire scale, raise concerns regarding the item design, internal consistency, and dimensionality of the IKDC-SKF. The results suggest further scale modification is needed to reduce redundancy, construct underrepresentation,^{15,17} and response burden,¹⁹ while also improving model fit.^{20,21}

Psychometrically sound reflective scales should also have consistent structural validity, which is often established through exploratory factor analysis (EFA)^{22–24} and confirmatory factor analysis (CFA) procedures^{20,25,26} or Rasch analysis.²⁶ When factor analysis is used, initial procedures should follow recommended extraction techniques,²⁷ and factor identification should follow recommended procedures, as under factoring or over factoring issues occur in scale development.²⁸ While principal component analysis (PCA) can be used initially to reduce the number of items and factors,^{23,29} a common factor approach is preferred,³⁰ and PCA should not be used as a sub-

stitute for EFA and CFA to determine the underlying latent factors.^{20,22,23,29} Further, once EFA procedures have been used to identify latent factors, CFA procedures in new samples to confirm the factor structure are recommended.^{20,22,23,29} An important step in this process is the identification of latent factors, which is often performed using multiple criteria, such as eigenvalues (e.g., Kaiser-Guttman criterion of values greater than 1.0), scree plots, percent variance explained, or minimum-average partial correlation.^{22–24} Parallel analysis is another approach that has been recommended because it performs well across PCA or EFA procedures for correctly identifying factors.^{24,30,31} As it relates to the IKDC-SKF, researchers have used PCA, EFA, and Bayesian CFA procedures to establish scale structure; however, multiple factor solutions have been found^{1,32,33} and best practice recommendations have not always been followed. For example, PCA has resulted in three-component solutions, with researchers supporting a one-factor solution,³² despite recommendations to not use PCA.^{20,22,23,29} Others have used EFA and Bayesian CFA methods which resulted in two-factor solutions,^{1,33} without implementing many recommendations (e.g., parallel analysis) in the identification of latent factors.^{20,22,23,29} Thus, the factor structures (e.g., number of factors) have not been consistent across studies,^{1,32,33} and the solutions have varied in the number of items to include in the final scale (e.g., 15 items across two-factors,³³ all 18 items across the PCA solution).³²

Short form versions have also been identified from EFA, CFA, and Rasch analysis procedures. The first short form (i.e., 15 items) solution, however, was not identified with EFA procedures using the most contemporary methods for item retention and factor identification (e.g., parallel analysis)^{24,30,31} and CFA results have indicated further modification of the scale is warranted to identify a sound short form version for use in clinical practice and research.⁸ The need for further item removal and the identification of a parsimonious short form was also supported with Rasch analysis; however, final model solutions differed with one retaining 5 items¹² and the other retaining 8 items.³⁴ Other concerns with these studies^{12,34} are the use of small samples (i.e., 77³⁴ and 160,¹² respectively) and respondent pools (e.g., healthy respondents) who are not representative of the patient population with which the scale is used in clinical practice and research. Thus, further research is needed to make clearer recommendations on a parsimonious IKDC-SKF short form that can be used in practice and research.

Finally, multi-group and longitudinal measurement invariance and hypothesis testing assessment results, which helps ensure scale suitability for use in research and clinical practice,^{20,25,26,35} have not been reported for the IKDC-SKF. Multi-group invariance testing should be conducted to ensure factorial stability exists across different populations, which establishes measurement properties are equivalent across various subgroups (e.g., sex, age, injury type).

Establishing multi-group invariance of PROMs allows clinicians and researchers to answer substantive questions regarding group differences.^{20,26,36} Longitudinal invariance testing is valuable for PROMs because it helps establish if the underlying constructs are adequately measured across repeated testing to allow clinicians or researchers to interpret score changes as true change.^{20,25,36} Establishing multi-group and longitudinal invariance then allows for hypothesis testing by determining if the scale can be used to measure differences between relevant groups or across time.^{20,25,26}

Thus, further psychometric assessment of the IKDC-SKF is warranted given the reported inconsistencies, concerns with scale measurement properties, lack of invariance analysis results, and inconsistent findings on a short form version. Performing EFA and CFA procedures in large, diverse, and separate samples is valuable for determining and then confirming or refuting the structural validity of the IKDC-SKF or an identified short form version. These procedures will allow for identification of a parsimonious and psychometrically sound scale when following contemporary factor analysis procedure recommendations. Additionally, assessing the internal consistency of the identified factors (i.e., one, two, or three factors) is warranted to further confirm internal consistency and measurement precision without item redundancy. Finally, conducting multi-group and longitudinal invariance testing will provide insight into whether the scale can be used to measure group differences and change over time. Establishing these scale properties provides clinicians and researchers with a psychometrically sound scale to track patient progress or compare groups. Therefore, the purpose of the study was to assess the psychometric properties of the IKDC-SKF in a large, heterogeneous sample. This included four separate mechanisms: 1) to conduct EFA following best practice recommendations to identify a sound latent structure, which may include alternate forms (i.e., short forms), of the IKDC-SKF in a large, heterogeneous sample; 2) to assess the internal consistency of any identified constructs; 3) to use CFA procedures to confirm the structural validity of the identified scale structure in a separate sample; and 4) to perform relevant multi-group and longitudinal invariance procedures on the identified scale to inform practitioners and researchers on scale use for assessing group differences and change over time.

METHODS

A sample of patient data obtained from the Surgical Outcome System (SOS, Arthrex, Naples, Florida) was used for the study. Patients provided informed consent prior to using the SOS and were emailed PROMs at predetermined intervals. Institutional Review Board (IRB) approval for the project was granted by the Cedar-Sinai Office of Research Compliance and Quality Improvement as part of a larger research project using SOS data. University IRB was not required because the deidentified data set was not considered human subject research.

For this study, patients who were classified in an arthroscopic knee surgery group and who had completed the

IKDC-SKF at baseline (i.e., pre-arthroscopic knee surgery) were included in the study. For longitudinal invariance, only patients who completed the IKDC-SKF at four time points (i.e., baseline [pre-arthroscopic knee surgery], three months post-surgery, six months post-surgery, and 12-months post-surgery) were included in the analysis.

INSTRUMENTATION

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM

The IKDC-SKF is a 19-item knee joint specific PROM.³⁷ The IKDC-SKF includes one dichotomous item, four 11-point Likert scale items, and fourteen 5-point Likert scale items. Eighteen of the items are summed into one score which ranges from 0 to 100 (item #19 is not included in the score).³⁷ A higher score represents less dysfunction, less pain, and greater knee function.³⁷

DATA ANALYSIS

A total of 1,959 individuals completed the IKDC-SKF prior to knee arthroscopy and were exported from the SOS database into the Statistical Package for Social Sciences (SPSS v. 25.0, Chicago, IL) and Analysis of Moment Structures (AMOS v. 25.0, Chicago, IL) for analysis. Cases with a z-score equal to or greater than ± 3.3 were classified as univariate outliers and were subsequently removed. The dataset was also assessed for multivariate outliers using Mahalanobis distance; cases with a $p < 0.001$ according to the Chi-square test were removed from the data set. Respondent data were not excluded if demographic information was missing because the primary study purpose was to assess the IKDC-SKF. Finally, histograms and descriptive statistics (i.e., skewness and kurtosis values) were used to assess the normality of the data. Following data cleaning, the data set was randomly split into two equal samples (n_1 and n_2).

EXPLORATORY FACTOR ANALYSIS

An EFA with maximum likelihood extraction and direct oblimin rotation was conducted on sample n_1 to identify a parsimonious scale. Bartlett's test of sphericity (< 0.001) and Kaiser-Meyer-Olkin values (≥ 0.80) were assessed, with values outside of the specified ranges constituting a violation of the test.³⁸ Items were assessed individually and removed one at a time until a parsimonious solution was identified.^{22,23} Item removal was guided by theoretical (e.g., item content), design-related (e.g., item structure)¹⁹ and statistical (e.g., low factor loadings ≤ 0.40 , high cross-loadings ≥ 0.30 , high bivariate correlations with another item, poor contribution to internal consistency) criteria.^{15, 23,25,38} Factor retention was guided by eigenvalues ≥ 1.0 , scree plot examination, and factors that explained $\geq 5.0\%$ of the variance.^{25,28,38,39} Parallel analysis was used to confirm or refute factor retention; eigenvalues of the original data set were compared to a randomly ordered data set to inform final factor retention.⁴⁰ Cronbach's alpha was also

calculated for each factor retained. Items were considered for removal if the alpha value was ≥ 0.90 ; item removal was guided by statistical guidelines (i.e., which item was most redundant), theory (e.g., item content), and item design. The final EFA solution resulted in a parsimonious IKDC short form to be confirmed with CFA.

*CONFIRMATORY FACTOR ANALYSIS OF PROPOSED IKDC-SKF SHORT-FORM**

Sample n_2 was used to conduct a CFA of the proposed IKDC-SKF short form to confirm model structure using maximum likelihood estimation in AMOS. Model fit indices used for evaluation included the Comparative Fit Index (CFI) ≥ 0.95 , Tucker-Lewis Index (TLI) ≥ 0.95 , and root mean square error of approximation (RMSEA) ≤ 0.06 . Models with fit indices values outside of the specified ranges indicated poor model fit and were interpreted as not supporting the proposed factor structure of the IKDC-SKF.^{20,41} CFA procedures also included assessing localized areas of strain and the interpretability, size, and statistical significance of the model's parameter estimates (i.e., factor variances, covariances, and indicator errors).²⁵ If indicated, additional items were removed, and the CFA procedures were repeated with the new model.

MULTI-GROUP INVARIANCE TESTING

Multi-group invariance testing between participant sex and age groups was conducted on the full sample (i.e., samples n_1 and n_2 combined) in three stages: 1) structural invariance to assess equivalent factor structure between subgroups; 2) metric invariance to assess equal factor loadings between subgroups; and 3) scalar invariance to confirm equal loadings and intercepts between subgroups. Each model was more restricted than the previous model²⁰ and each step was used to assess whether the items were being interpreted equally across selected subgroups (i.e., sex, age group). These steps ensure the meanings of the common factors are consistent across groups and that mean scores are not contaminated by outside factors (e.g., group specific attributes), which then allows for substantive questions to be answered to support hypothesis testing (e.g., comparison of subgroup means).²⁰ If the metric model held, subgroups could be tested for equal variances on the latent constructs, and if the scalar model held, subgroups could be tested for equal latent means. For the purposes of multi-group analysis by age, participants were split into groups defined as youth (<18 years old), emerging adult (18-25 years old), early adulthood (26-40 years old), middle age (41-65 years old), and older adult (>65 years old)⁴²; however, the older adult group was not analyzed because of its small sample size ($n = 30$). The χ^2_{diff} and CFI_{diff} tests were both used to assess invariance, and the scale was considered invariant at each stage if the CFI_{diff} was ≤ 0.01 as compared to the configural model and the fit indices previously described were met. If the model was not found to be invariant at a given step, item loadings (i.e., metric model) or item intercepts (i.e., scalar model) were released one by one, and the model was retested. Once a problematic item

was identified (i.e., the one that improved CFI to be closest to the CFI of the configural model), it was removed, and the model was re-run. For the substantive questions, if the CFI_{diff} was > 0.01 compared to the configural model, it was deemed that the subgroups were not equal on the tested statistic (e.g., latent means). In these cases, another model was run in which one group served as the comparison group to determine relative latent variances or means for the other subgroups (i.e., greater than, less than, or equal to the comparison group). The χ^2_{diff} test was not weighted as heavily in the invariance process because of the effect sample size has on this statistic.^{20,21}

LONGITUDINAL INVARIANCE TESTING

Longitudinal invariance testing was evaluated using the same procedures outlined in the multi-group invariance section to confirm similar interpretation of items and common factors across time points. If all models held (i.e., all fit indices cut-off values were met), it indicated that substantive properties (e.g., change over time) could be evaluated, allowing for clinician assessment of patient scores over time (e.g., did scores change from baseline to 12-months post-arthroscopy). The same procedures were used as described in multi-group invariance testing to identify any problematic items and create a more parsimonious scale, when indicated.

CORRELATION ANALYSES

Bivariate correlation analysis was conducted using scores from the 18-item IKDC-SKF and scores from any generated IKDC-SKF short forms. The preferred percentage of variance explained was set at $r \geq 0.90$ ($R^2 = 0.81$).^{43,44}

RESULTS

A total of 55 cases were removed during the data cleaning process (i.e., identified outliers) leaving 1,904 cases for analysis; the 1,904 total cases were then randomly split into two even data sets (i.e., 952 cases in n_1 and n_2). For the full sample, participants were 32.06 ± 14.16 years of age (range: 11-80 years) and included 874 males and 802 females. For sample n_1 , participants were 32.42 ± 14.38 years old (range: 11-74 years) and included 441 males and 388 females. For sample n_2 , participants were an average of 31.69 ± 13.93 years old (range: 12-80 years) and included 433 males and 414 females.

EXPLORATORY FACTOR ANALYSIS

The initial exploratory factor analysis (EFA) using all 18 items resulted in a four-factor solution with items that had low loadings and high cross-loadings. Parallel analysis indicated a three-factor solution was sufficient when all 18 items were used. Items were removed during the EFA procedures one at a time and the solution was respecified until an acceptable solution was identified; a total of nine items were removed, resulting in a 9-item, 3-factor solution, with three items in each factor (Athletic Activities, Ac-

Table 1. Exploratory factor analysis of the IKDC-SKF

Item	Factor 1 Athletic Activities	Factor 2 Activity Level	Factor 3 ADLs
16	0.925		
17	0.850		
15	0.713		
5		0.857	
1		0.797	
8		0.731	
9			0.786
12			0.657
13			0.647
Eigenvalue (% variance)	4.67 (51.87)	1.21 (13.44)	0.94 (10.43)
Cronbach's Alpha	0.88	0.84	0.76

ADLs = Activities of Daily Living.

tivity Level, Activities of Daily Living [ADLs]; [Table 1](#)). The solution accounted for 75.74% of the variance and Cronbach's alpha values fell within the suggested range for each subscale (ADLs: 0.76; Activity Level: 0.84; Athletic Activities: 0.88; [Table 1](#)) with item loadings ranging from 0.65 to 0.93. While certain criteria (e.g., scree plot, percent variance explained) supported the 3-factor structure solution, parallel analysis with the nine items supported a two-factor structure. The 3-factor, 9-item scale was retained for CFA as further modification could be conducted during those analysis procedures to support or refute factor structure.

CONFIRMATORY FACTOR ANALYSIS 9-ITEM IKDC-SKF SHORT FORM

The CFA of the 3-factor, 9-item IKDC-SKF short form met all model fit criteria (CFI = 0.983; TLI = 0.975; IFI = 0.983; RMSEA = 0.057; chi-square = 97.667; $p < .001$; Figure 1) and had factor loadings ranging from 0.55 to 0.87. Construct correlations ranged from 0.64 to 0.75, with the highest correlation between Athletic Activities and ADLs (56.25% shared variance). Modification indices indicated significant cross-loadings and potential model misspecification were present.

Although model fit indices were exceeded, inspection of the model (e.g., item design, latent variable correlations, modification indices) and consideration of the parallel analysis findings led to further refinement and the identification of a 2-factor, 6-item modified IKDC-SKF short form. The 2-factor (Activity Level and ADLs), 6-item model supported by parallel analysis, demonstrated excellent model fit (CFI = 1.0; TLI = 0.999; IFI = 1.0; RMSEA = 0.11; chi-square = 8.943; $p = 0.347$; Figure 2), and addressed concerns (e.g., high latent variable correlations, cross-loadings) identified in the 9-item IKDC-SKF short form. Invariance testing (multigroup and longitudinal) was conducted on both the 9-item and 6-item IKDC-SKF short forms to provide further insight on both proposed factor structures.

MULTIGROUP INVARIANCE TESTING

Multigroup invariance testing across sex and age groups was conducted using participant responses to the IKDC-SKF at baseline (i.e., pre-arthroscopy).

SEX

IKDC-SKF 9-ITEM SHORT FORM

A total of 1,676 individuals (males = 874; females = 802) reported sex and were used for analysis. Both individual models (i.e., males, females) met all fit indices criteria ([Table 2](#)). The configural model fit indices also met all recommended values (CFI = 0.990; RMSEA = 0.031; [Table 2](#)). The metric and scalar models passed the CFI_{diff} test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_{diff} test, indicating variances were equal across groups. The equal means model also passed the CFI_{diff} test, indicating the means were equal for all latent variables across males and females.

IKDC-SKF 6-ITEM SHORT FORM

A total of 1,676 individuals (males = 874; females = 802) reported sex and were used for analysis. Both individual models (i.e., males, females) met all fit indices criteria ([Table 3](#)). The configural model also met all recommended model fit values (CFI = 0.990; RMSEA = 0.031; [Table 3](#)). The metric and scalar models passed the CFI_{diff} test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_{diff} test, indicating variances were equal across groups. The equal means model also passed the CFI_{diff} test, indicating the means were equal for all latent variables across males and females.

Table 2. Invariance across sex for the 9-item IKDC-SKF Short Form

	Chi-Square	df	Chi-Square Diff	CFI	CFI Diff	TLI	RMSEA
Measure 1 (Male)	56.039	24	N/A	0.992	N/A	0.989	0.039
Measure 2 (Female)	68.907	24	N/A	0.987	N/A	0.980	0.048
Model A (Configural)	124.947	48	N/A	0.990	N/A	0.985	0.031
Model B (Metric)	127.682	54	2.735 (6)	0.990	NC	0.987	0.029
Model C (Equal Latent Variances)	136.188	57	11.241 (9)	0.990	NC	0.987	0.029
Model D (Scalar)	137.173	60	12.226 (12)	0.990	NC	0.988	0.028
Model E (Equal Latent Means)	151.039	63	26.092 (15)	0.989	0.001	0.987	0.029

Bold italic font: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

Table 3. Invariance across sex for the 6-item IKDC-SKF Short Form

	Chi-Square	df	Chi-Square Diff	CFI	CFI Diff	TLI	RMSEA
Measure 1 (Male)	14.813	8	N/A	0.997	N/A	0.994	0.031
Measure 2 (Female)	10.948	8	N/A	0.998	N/A	0.997	0.021
Model A (Configural)	25.761	16	N/A	0.997	N/A	0.995	0.019
Model B (Metric)	28.506	20	2.745 (4)	0.998	+0.001	0.997	0.016
Model C (Equal Latent Variances)	37.002	22	11.241 (6)	0.996	0.001	0.995	0.020
Model D (Scalar)	32.529	24	6.768 (8)	0.998	+0.001	0.987	0.015
Model E (Equal Latent Means)	44.991	26	19.23 (10)	0.995	0.002	0.994	0.021

Bold italic font: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

AGE GROUP

IKDC-SKF 9-ITEM SHORT FORM

A total of 1,762 individuals (youth = 321; emerging adults = 416; early adulthood = 558; middle age = 467) who reported an age (range: 11–65 years) were used for analysis. Baseline models (i.e., youth, emerging adults, early adults, middle age) met all fit indices (Table 4). The configural model fit indices met all recommended values (CFI = 0.993; RMSEA = 0.019; Table 4). The metric and scalar models passed the CFI_{diff} test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_{diff} test, indicating variances were equal across groups. The equal means model did not pass the CFI_{diff} test, indicating the means were not equal for all latent variables between age groups. When means were not constrained, the middle age group had significantly lower means than all groups (i.e., more dysfunction, more pain, and less knee ability) across all three latent variables (i.e., ADLs, Activity Level, and Athletic Activities). Additionally, the early adulthood group had a significantly lower mean (i.e., more dysfunction, more pain, and less knee ability) than the youth and emerging adult groups for the ADL latent variable. Statistically significant mean differences were not found

for any latent constructs between the youth and emerging adult groups.

IKDC-SKF 6-ITEM SHORT FORM

A total of 1,762 individuals (youth = 321; emerging adults = 416; early adulthood = 558; middle age = 467) who reported an age (range = 11–65 years) were used for analysis. Baseline models (i.e., youth, emerging adults, early adults, middle age) met all fit indices (Table 5). The configural model fit indices met all recommended values (CFI = 0.993; RMSEA = 0.019; Table 5). The metric and scalar models passed the CFI_{diff} test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_{diff} test, indicating variances were equal across groups. The equal means model did not pass the CFI_{diff} test, indicating the means were not equal for all latent variables between age groups. When means were not constrained, the middle age group had significantly lower means (i.e., more dysfunction, more pain, and less knee ability) than all groups across both latent variables (i.e., ADLs and Activity Level). Additionally, the early adulthood group had a significantly lower mean (i.e., more dysfunction, more pain, and less knee ability) than the youth and emerging adult groups for the ADL latent variable. There

Table 4. Invariance across age group for the 9-item IKDC-SKF Short Form

	Chi-Square	df	Chi-Square Diff	CFI	CFI Diff	TLI	RMSEA
Measure 1 (Youth)	21.997	24	N/A	1.000	N/A	1.000	0.000
Measure 2 (Emerging Adult)	38.464	24	N/A	0.992	N/A	0.988	0.038
Measure 3 (Early Adult)	56.629	24	N/A	0.988	N/A	0.982	0.049
Measure 4 (Middle Age)	36.968	24	N/A	0.993	N/A	0.989	0.034
Model A (Configural)	154.051	96	N/A	0.993	N/A	0.989	0.019
Model B (Metric)	188.910	114	34.859 (18)	0.990	0.003	0.988	0.019
Model C (Equal Latent Variances)	225.757	123	71.706 (27)	0.987	0.006	0.987	0.022
Model D (Scalar)	236.479	132	82.428 (36)	0.987	0.006	0.985	0.021
Model E (Equal Latent Means)	332.783	141	178.732 (45)	0.975	0.022	0.975	0.028

Bold italic font: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

Table 5. Invariance across age group for the 6-item IKDC-SKF Short Form

	Chi-Square	df	Chi-Square Diff	CFI	CFI Diff	TLI	RMSEA
Measure 1 (Youth)	2.211	8	N/A	1.000	N/A	1.000	0.000
Measure 2 (Emerging Adult)	5.923	8	N/A	1.000	N/A	1.000	0.000
Measure 3 (Early Adult)	7.380	8	N/A	1.000	N/A	1.000	0.000
Measure 4 (Middle Age)	10.380	8	N/A	0.998	N/A	0.995	0.025
Model A (Configural)	25.892	32	N/A	1.000	N/A	1.000	0.000
Model B (Metric)	47.957	44	22.065 (12)	0.999	0.001	0.999	0.007
Model C (Equal Latent Variances)	69.335	50	43.443 (18)	0.995	0.005	0.994	0.015
Model D (Scalar)	63.078	56	37.186 (24)	0.998	0.002	0.998	0.008
Model E (Equal Latent Means)	153.676	62	127.784 (30)	0.976	0.022	0.977	0.029

Bold italic font: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

were no significant mean differences for any latent constructs between the youth and emerging adult groups.

LONGITUDINAL INVARIANCE TESTING

A total of 792 individuals completed the IKDC-SKF at all four time points and were retained for longitudinal invariance. The average age of participants in this subsample was 33.02 ± 15.00 years (range: 11-78 years; 354 females; 353 males).

IKDC-SKF 9-ITEM SHORT FORM

All baseline models (i.e., baseline, 3-months post-surgery, 6-months post-surgery, 12-months post-surgery) met all fit indices (Table 6). The configural model fit indices met all recommended values (CFI = 0.993; RMSEA = 0.019; Table 6). The metric model passed the CFI_{diff} test, warranting examination of an equal variances model. The equal variance model did not pass the CFI_{diff} test, indicating variances were not equal across time points for latent variables. The

scalar model, however, did not pass the CFI_{diff} test, indicating potential item-level bias which did not support testing of the equal latent means model. Follow-up analysis indicated slight item bias for item #15 (i.e., “How does your knee affect your ability to run straight ahead?”).

Due to the item bias findings, invariance testing was conducted on an 8-item IKDC-SKF short form (i.e., the remaining items from the 9-item scale after item #15 was removed). All baseline models (i.e., baseline, 3-months post-surgery, 6-months post-surgery, 12-months post-surgery) met model fit indices (Table 7). The configural model fit indices met all recommended values (CFI = 0.997; RMSEA = 0.014; Table 7). The metric and scalar models passed the CFI_{diff} test, warranting examination of an equal variances and equal means model. The equal variance model did not pass the CFI_{diff} test, indicating variances were not equal across time points for latent variables. The equal means model also did not pass the CFI_{diff} test, indicating means were not equal across time. When not constrained to be equal, Activity Level, ADLs, and Athletic Activities latent means at 3-, 6-, and 12-months post-surgery were signif-

Table 6. Longitudinal Invariance for the 9-item IKDC-SKF Short Form

	Chi-Square	df	Chi-Square Diff	CFI	CFI Diff	TLI	RMSEA
Measure 1 (Pre)	53.313	24	N/A	0.990	N/A	0.985	0.043
Measure 2 (3 months)	52.882	24	N/A	0.992	N/A	0.987	0.039
Measure 3 (6 months)	37.748	24	N/A	0.997	N/A	0.995	0.027
Measure 4 (12 months)	65.321	24	N/A	0.992	N/A	0.989	0.047
Model A (Configural)	614.743	474	N/A	0.993	N/A	0.990	0.019
Model B (Metric)	740.369	492	125.626 (18)	0.987	0.006	0.984	0.025
Model C (Equal Latent Variances)	1108.942	501	494.199 (27)	0.969	0.024	0.961	0.039
Model D (Scalar)	908.477	510	293.734 (36)	0.980	0.013	0.975	0.031

Bold italic font: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

Table 7. Longitudinal Invariance for the 8-item IKDC-SKF Short Form

	Chi-Square	df	Chi-Square Diff	CFI	CFI Diff	TLI	RMSEA
Measure 1 (Pre)	24.382	17	N/A	0.997	N/A	0.995	0.023
Measure 2 (3 months)	27.127	17	N/A	0.996	N/A	0.994	0.027
Measure 3 (6 months)	25.566	17	N/A	0.997	N/A	0.996	0.025
Measure 4 (12 months)	28.512	17	N/A	0.997	N/A	0.996	0.029
Model A (Configural)	402.353	350	N/A	0.997	N/A	0.995	0.014
Model B (Metric)	473.321	365	70.968 (15)	0.993	0.004	0.991	0.019
Model C (Equal Latent Variances)	811.086	374	408.733 (24)	0.973	0.024	0.964	0.038
Model D (Scalar)	564.617	380	162.264 (30)	0.989	0.008	0.975	0.031
Model E (Equal Latent Means)	1505.912	386	1103.559 (36)	0.930	0.067	0.910	0.061

Bold italic font: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

icantly higher than baseline (i.e., pre-arthroscopy) scores (i.e., less dysfunction, less pain, and higher knee ability), except for Activity Level latent means at three months. Scores increased/improved across time, except for Activity Level at three months, indicating patients reported scores with improved function, pain, and knee ability after surgery.

IKDC-SKF 6-ITEM SHORT FORM

All baseline models (i.e., baseline, three months post-surgery, six months post-surgery, 12-months post-surgery) met all fit indices (Table 8). The configural model fit indices met all recommended values (CFI = 0.998; RMSEA = 0.012). The metric model and scalar model passed the CFI_{diff} test, warranting examination of an equal variances and equal means model. The equal variance model did not pass the CFI_{diff} test, indicating variances were not equal across time points. The equal means model also did not pass the CFI_{diff} test, indicating means were significantly different across time points. When not constrained, Activity Level and ADL latent means at three-, six-, and 12-months post-surgery

were significantly higher than baseline (i.e., pre-arthroscopy) scores (i.e., less dysfunction, less pain, and higher knee ability), except for Activity Level at six months post-surgery. Scores increased/improved across time, except for Activity Level at three months, indicating patients reported scores with improved function, pain, and knee ability after surgery.

CORRELATION ANALYSIS

Individual scores for the IKDC-SKF 9-item short form were highly correlated ($r = 0.924$, $R^2 = 0.854$) with the scores for the original 18-item IKDC-SKF. Individual scores for the IKDC-SKF 6-item short form were highly correlated ($r = 0.889$, $R^2 = 0.790$) with the scores for the original 18-item IKDC-SKF. Scores for the IKDC-SKF 9-item short form were also highly correlated ($r = 0.940$, $R^2 = 0.884$) with the scores for the 6-item IKDC-SKF short form. Finally, scores on the modified 8-item (3-dimension) IKDC-SKF short form were highly correlated with scores on the original 18-item IKDC-SKF ($r = 0.919$, $R^2 = 0.845$), the 9-item IKDC-SKF short form ($r = .992$, $R^2 = 0.984$), and the 6-item IKDC-SKF short form ($r = .962$, $R^2 = 0.925$).

Table 8. Longitudinal Invariance for the 6-item IKDC-SKF Short Form

	Chi-Square	df	Chi-Square Diff	CFI	CFI Diff	TLI	RMSEA
Measure 1 (Pre)	15.254	8	N/A	0.995	N/A	0.992	0.034
Measure 2 (3 months)	15.961	8	N/A	0.996	N/A	0.992	0.035
Measure 3 (6 months)	6.284	8	N/A	1.000	N/A	1.000	0.000
Measure 4 (12 months)	12.202	8	N/A	0.999	N/A	0.997	0.026
Model A (Configural)	210.782	188	N/A	0.998	N/A	0.997	0.012
Model B (Metric)	264.528	200	53.746 (12)	0.994	0.004	0.992	0.020
Model C (Equal Variances)	554.915	206	344.133 (18)	0.969	0.029	0.958	0.046
Model D (Scalar)	344.718	212	133.936 (24)	0.988	0.01	0.984	0.028
Model E (Equal Latent Means)	1214.108	218	1003.326 (30)	0.911	0.087	0.887	0.076

Bold italic font: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

DISCUSSION

Best practice recommendations for assessing the structural validity have not always been followed or reported in measurement studies of the IKDC-SKF,^{20,22-31} which may explain the inconsistent structural findings reported.^{1,32,33,37} Further, multiple short form versions of the IKDC-SKF have been suggested in the literature, but initial efforts have primarily used small samples that do not well represent the patient population who completes the IKDC-SKF.^{12,34} Therefore, assessment of the IKDC-SKF using recommended classical test theory procedures was warranted and the purpose of our study was to conduct EFA, CFA, and invariance testing procedures on the IKDC-SKF in a large, heterogeneous sample of patients to assess the measurement properties of the scale or an alternate, psychometrically sound short form version of the scale. EFA resulted in a 9-item, 3-factor IKDC-SKF short form (IKDC-SKF-9; Appendix 1) supported by CFA and multi-group invariance testing; however, the proposed model did not meet all recommended measurement criteria and did not pass longitudinal invariance requirements. Due to potential concerns with the identified 9-item version, subsequent 8-item (3-factor; IKDC-SKF-8; Appendix 2) and 6-item (2-factor; IKDC-SKF-6; Appendix 3) IKDC-SKF short forms (Appendices 1 and 2) were also tested with CFA and invariance procedures.

FACTOR STRUCTURE

The IKDC-SKF has been reported to have a unidimensional^{32,37} and a multidimensional^{1,8,33} factor structure with different items in the final models. Study methodology differences may contribute to the inconsistent findings as differences in samples (e.g., size, respondent population [e.g., healthy,¹² ACL injury,^{1,34}]) and analysis methodologies (e.g., EFA/CFA,³³ PCA,³² Bayesian SEM,¹ Rasch,^{12,34} factor and item retention criteria) exist between studies. For example, our study included a large, heterogeneous sample of patients who had undergone arthroscopic knee surg-

eries, while others have included smaller samples, healthy respondents,¹² or focused on different patient populations (e.g., ACL reconstruction,^{1,34} meniscal lesions,⁴⁵ various patient pathologies^{32,33,37}). Research^{1,8,33} using more contemporary and recommended factor analysis measurement techniques has generally supported a multidimensional structure; however, evidence exists to suggest structural validity and model fit could be improved with further item reduction.^{8,12,34}

Three-dimensional and two-dimensional factor structures that exceeded most recommended contemporary fit criteria were identified.^{20,41} The retained solutions included fewer items than those found in prior research on the IKDC-SKF.^{1,8,33} Removing items with poor fit (e.g., cross-loadings, item redundancy) or poor design (e.g., item structure, item reading level, etc.) can improve internal consistency and scale structure, while also reducing response burden with a more concise instrument.^{20,38,43} The proposed short form versions improved model precision and scale structure without losing much of the information captured with all 18 items. Despite using nine or fewer items, scores on the short form versions accounted for 79% ($r = .889$), 84% ($r = .919$), and 85% ($r = .924$) of the variance in scores on the 18-item original IKDC-SKF with the 6-, 8-, and 9-items, respectively. Our correlational findings are in line with prior research using similar procedures to produce short-form versions of previously established PROMs^{43,44} and indicate the newly proposed models capture similar enough information to warrant use in comparison to the original scale. One concern, however, was the identification of an internal consistency value (0.76) for the ADLs construct that is outside of the preferred 0.80 to 0.90 range for scale development.^{15,17} Thus, the ADLs construct may not have the preferred precision for assessing the construct; however, scale design recommendations caution against using constructs with fewer than three items.^{20,21,46} Future work may be recommended to alter or add items to measure the ADLs construct more precisely. Until that time, researchers and clinicians should be aware that this construct

does not meet the strictest contemporary recommendations for internal consistency.

MULTI-GROUP INVARIANCE TESTING

This study is the first to perform multi-group invariance testing with the IKDC-SKF. Multi-group invariance testing helps to ensure the association between the items and dimensions are stable between groups, which supports scale validity and allows for an instrument to be used to assess group differences (e.g., group mean differences in older individuals compared to younger individuals would be outside scale measurement error).^{20,21,25} Both the 9-item and 6-item IKDC-SKF short form versions in this study were found to be invariant across configural, metric, and scalar models for sex and age groups, indicating the short form models have sound measurement properties across the tested groups. Thus, researchers and clinicians could use these versions of the scale to assess differences among these groups. The findings also allow for substantive testing of whether variances or means are equal between groups, which can also support scale validity.^{20,21}

The current multi-group findings could provide theoretical support for the validity of the two short form versions if the findings align with expectations based on the literature. For example, widespread pain, which often includes longstanding knee pain, is more common in patients over 50 years of age,⁴⁷ and self-reported knee pain has been found to be higher in people over the age of 40 compared to those under the age of 40.⁴⁸ Additionally, OA, a leading diagnosis and cause of disability in older populations,^{49,50} has a higher prevalence and more radiographic signs with increases in age and population longevity,⁵¹ with those over the age of 45 accounting for over 98% of total knee arthroplasties.⁵² Further, the presence of all types of knee abnormalities (e.g., osteophytes, cartilage damage, ligamentous damage, OA) has been found to increase with age,⁵³ and it has been reported that 85% or more of patients 50 years of age or older demonstrate articular cartilage changes to at least one knee compartment compared to 32% of patients between the ages of 20 to 29 years of age and 13% of patients 20 years or younger.⁵⁴ Researchers have also indicated knee functional difficulties increase with age.⁵⁵⁻⁵⁸ Gradual functional deterioration was found across the lifespan on the Knee Injury and Osteoarthritis Outcome Score (KOOS) and reported functional impairment was more apparent with functional tasks of greater difficulty (e.g., sport and recreational functional activities) across adults aged 18-84.⁵⁹ Similarly, Baldwin et al. reported group mean score differences across age groups (e.g., 18 to 29 years, 30 to 39 years, 40 to 49 years, 50 to 59 years), with consistent findings of less knee impairment on the pain, ADL, sport/recreation, and quality of life constructs of the KOOS for those under 40 years of age.⁴⁸ Thus, if the proposed IKDC-SKF short forms are measuring the intended constructs, it would be expected to find similar patterns in our multi-group invariance results (e.g., higher levels of impairment in the older age groups in our sample).

The proposed models were able to identify age group differences at initial examination: statistically significant

group differences were found across age groups for the 6-item and 9-item solutions. Specifically, the middle age (41-65 years of age) group had lower means (i.e., more impaired knee health) than the other younger age groups (i.e., early adult, emerging adult, youth) at baseline (i.e., pre-arthroscopic surgery). Additionally, the early adulthood (26-40 years of age) group had a lower mean score for the ADL construct than the youth and emerging adult groups. The middle age group reporting greater impairment across all three factors (i.e., ADLs, Activity Level, Athletic Activity) than the three younger age groups align with expectations based on KOOS findings^{48,59} and expectations for functional impairment across the life span.^{48,55-59} Similarly, the early adult age group reporting greater impairment in the Activity Level factor but similar mean scores for ADLs with younger age groups also aligns with expectations based on the literature.^{48,55-59} The lack of a statistically significant difference for the Athletic Activity construct might be explained by the final items included in that construct; however, prior KOOS Sport/Recreation construct findings indicate meaningful age group differences were not found until after 40 years of age for items assessing this type of construct.^{48,59} Finally, statistically significant differences were not found between the emerging adult and youth groups, which aligns with prior KOOS findings^{48,59} and expectations for the presence of knee pathological changes being less likely to have occurred this early in the lifespan.^{51,53,54} A weakness of our results, however, is the low number of responses ($n = 30$) in the older adult (66 years of age or older) group, which prevented us from including this group in the multi-group analyses of the 6- and 9-item short form versions. It would be valuable to confirm that similar group differences are found in older or elderly populations. Similarly, it would be valuable to conduct this analysis across different pathology groups (e.g., total knee arthroplasty patients vs. arthroscopy patients) to ensure the scale has the necessary measurement properties to assess groups differences based on pathology and if greater levels of diagnosed pathology results in greater reported knee health impairment on the proposed short forms.

LONGITUDINAL INVARIANCE TESTING

To the authors knowledge, this study is also the first to assess longitudinal invariance of the IKDC-SKF or proposed short forms. Longitudinal invariance testing is valuable because it allows for the determination of whether the items and dimensions are stable across time, which supports scale validity and allows for an instrument to be used to assess change over time.^{20,21,25} We found that the 6-item IKDC-SKF short form was invariant across time based on the configural, metric, and scalar model findings. The 9-item IKDC-SKF short form was not invariant across time, and further analysis revealed item #15 exhibited bias. Follow-up analysis indicated that the remaining items (i.e., 9-item IKDC-SKF short form except for item #15; 8-item IKDC-SKF short form) and factor structure were invariant across time. The findings allow for the assessment of score change over time to determine when and where patient reported improvement occurred following arthroscopic

surgery on the 6- and 8-item short form versions. Finding expected improvement over time would support scale validity, while also indicating whether patients perceived improvements in their condition across time.^{20,21,25}

In this study, individuals reported the lowest scores (i.e., greatest impairment in knee health) at baseline and the highest scores (i.e., lowest knee impairment) at 12-months post-surgery. The score improvements were statistically significant for all latent means (e.g., Activity Level, ADLs) at 3-, 6-, and 12-months post-surgery except Activity Level at three months post-surgery across the 6- and 8-item versions of the scale. The current findings indicate patients reported statistically significant improvements across all dimensions six months post-surgery and the improvements were maintained at 12-months post-surgery. The longitudinal findings were consistent across both the 6- and 8-item short form versions of the IKDC-SKF. Thus, the two versions of the scale identified patient-reported improvement across the measured latent constructs across time similarly.

The current findings support scale validity as the results are consistent with what we would expect for individuals recovering from surgery. Specifically, patients in the rehabilitation process would be expected to report improvements across items intended to measure how the prior injury impaired the previously measured constructs (i.e., Activity Levels, ADLs, and Athletic Activities) because the patient should experience health status improvement (e.g., decreased pain, increased ROM, increased strength) after surgery through a combination of treatment effectiveness, natural healing, and placebo. We would also expect to find that improvements in certain constructs (e.g., Activity Level) might not occur as quickly as other constructs (e.g., ADLs) because patients may have activity/rehabilitation restrictions or more substantial pathology that may slow improvements in specific constructs; however, the authors would also expect to then see significant improvements in those dimensions at later time points that are in line with the improvements found across the other constructs over time.

Thus, the current findings support the use of the 6- and 8-item IKDC-SKF short-form versions: sound measurement properties were demonstrated and theoretical support (e.g., patient-perceived improvements match expectations for the recovery process of the included patients in our study) was found. These results also provide support for clinicians who want to use the short form versions of the IKDC-SKF (i.e., 6- and 8-item versions) to measure change across time. The 9-item short form could be used with caution to assess change across time because it did not meet the strictest criterion for longitudinal invariance due to one problematic item. However, it is also important to note limitations with the 3-dimensional solutions: 1) parallel analysis better supported a 2-dimensional factor structure once problematic items had been removed; and 2) the Athletic Activities factor in the 8-item IKDC-SKF short form only contained two items and three to five items per factor has been recommended.^{20,46,60} Clinicians and researchers should consider summary of findings when deciding which version of the IKDC-SKF to use within their clinical practice

or research; however, the 6-item IKDC-SKF likely has the greatest measurement support for its use across various research and clinical practice scenarios. Further scale development work is needed to develop items to accurately capture the desired information of the Athletic Activities factor and truly support a 3-dimensional IKDC-SKF factor structure.

LIMITATIONS AND FUTURE RESEARCH

While the current study included the use of contemporary analysis procedures on a large, diverse sample of patients, it does have limitations. First, the data set did not include information on the type of knee pathology or procedure performed. One of the preconditions for a viable IKDC-SKF instrument is that the model is stable over a variety of knee pathologies. Without the relevant demographic information, we were unable to conduct multi-group invariance tests by pathology or intervention type. Additionally, our sample had a small sample of patients classified in the older age (66 years or older) group, which prevented their inclusion in the multi-group age analysis. Further, responses to all 18-items were used to produce short-form versions; while the analysis processes used are common for instrument refinement, it is possible that respondents were influenced by items not included in the final models. Additional psychometric analyses could also be conducted; for example, the new models could be tested against a criterion standard scale to support validity, Rasch analysis (e.g., person differentiation) could be performed, and responsiveness (e.g., minimal clinically important difference [MCID] values) and test-retest reliability of the new models could be assessed. Finally, as this PROM was delivered via email, potential response biases could have affected results and it was not possible to examine if completion mode (i.e., paper or electronic) influenced results.

Future analysis should include multi-group invariance testing in older populations, while also examining the multi-group invariance properties across pathology or intervention groups. Further, researchers should examine the structural validity of the scale in different respondent groups who only answer the short form versions of the scale, while also incorporating additional items to measure the Athletic Activities factor more effectively. In addition to confirming the measurement properties of the short form versions, these analyses could provide insight into whether the 6-item short form may have other psychometric concerns (e.g., ceiling effects) when used in certain populations (e.g., competitive athletes) that could be resolved by developing an effective 3-dimensional scale. Finally, future research should also work to establish the test-retest reliability, responsiveness (e.g., MCIDs), and criterion validity of the short form versions.

CONCLUSION

The EFA and CFA resulted in short form versions of the IKDC-SKF that exceed contemporary fit recommendations. The identified models present as plausible alternatives to

the IKDC-SKF as the original item pool was reduced by more than 50%, but the short forms still accounted for most of the variance in participant responses on the IKDC-SKF. Further, the 6- and 8-item IKDC-SKF short forms met all criteria for applied multi-group and longitudinal invariance tests, which indicates the scales may be used to assess group differences or change across time. The overall analysis indicated the short form versions of the IKDC-SKF were structurally valid alternatives to the IKDC-SKF with improved measurement properties, reduced scale response burden, and evidence to support the assessment of patient improvement across time.

.....

FINANCIAL DISCLOSURE

This publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under Grant #P20GM103408 and an Idaho WWAMI Research Training Support Award.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Submitted: March 10, 2023 CDT, Accepted: June 16, 2023 CDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-nc/4.0> and legal code at <https://creativecommons.org/licenses/by-nc/4.0/legalcode> for more information.

REFERENCES

1. Williams T, Burley D, Evans L, et al. The structural validity of the IKDC and its relationship with quality of life following ACL reconstruction. *Scand J Med Sci Sports*. 2020;30(9):1748-1757. [doi:10.1111/sms.13738](https://doi.org/10.1111/sms.13738)
2. Hambly K, Griva K. IKDC or KOOS: which one captures symptoms and disabilities most important to patients who have undergone initial anterior cruciate ligament reconstruction? *Am J Sports Med*. 2010;38(7):1395-1404. [doi:10.1177/0363546509359678](https://doi.org/10.1177/0363546509359678)
3. Crawford K, Briggs KK, Rodkey WG, Steadman JR. Reliability, validity, and responsiveness of the IKDC score for meniscus injuries of the knee. *Arthroscopy*. 2007;23(8):839-844. [doi:10.1016/j.arthro.2007.02.005](https://doi.org/10.1016/j.arthro.2007.02.005)
4. Hambly K, Griva K. IKDC or KOOS? Which measures symptoms and disabilities most important to postoperative articular cartilage repair patients? *Am J Sports Med*. 2008;36(9):1695-1704. [doi:10.1177/0363546508317718](https://doi.org/10.1177/0363546508317718)
5. Latt L, Fornalski S, Raiszadeh K, Inacio M, Fithian D, Csintalan R. Medial Patellofemoral Ligament (MPFL) reconstruction for the treatment of patellofemoral instability. *J Knee Surg*. 2014;27(2):139-146. [doi:10.1055/s-0033-1360652](https://doi.org/10.1055/s-0033-1360652)
6. Fu SN, Chan YH. Translation and validation of Chinese version of International Knee Documentation Committee Subjective Knee Form. *Disabil Rehabil*. 2011;33(13-14):1186-1189. [doi:10.3109/09638288.2010.524274](https://doi.org/10.3109/09638288.2010.524274)
7. Ahmed KM, Said HG, Ramadan EKA, Abd El-Radi M, El-Assal MA. Arabic translation and validation of three knee scores, Lysholm Knee Score (LKS), Oxford Knee Score (OKS), and International Knee Documentation Committee Subjective Knee Form (IKDC). *SICOT-J*. 2019;5(6):1-13. [doi:10.1051/sicotj/2018054](https://doi.org/10.1051/sicotj/2018054)
8. Kümmel D, Preiss S, Harder LP, Leunig M, Impellizzeri FM. Measurement properties of the German version of the IKDC subjective knee form (IKDC-SKF). *J Patient Rep Outcomes*. 2018;2(1):2-9. [doi:10.1186/s41687-018-0058-1](https://doi.org/10.1186/s41687-018-0058-1)
9. Kanakamedala AC, Anderson AF, Irrgang JJ. IKDC Subjective Knee Form and Marx Activity Rating Scale are suitable to evaluate all orthopaedic sports medicine knee conditions: a systematic review. *J ISAKOS*. 2016;1(1):25-31. [doi:10.1136/jisakos-2015-00014](https://doi.org/10.1136/jisakos-2015-00014)
10. Grevnerts HT, Terwee CB, Kvist J. The measurement properties of the IKDC-subjective knee form. *Knee Surg Sports Traumatol Arthrosc*. 2015;23(12):3698-3706. [doi:10.1007/s00167-014-3283-z](https://doi.org/10.1007/s00167-014-3283-z)
11. Esculier JF, Roy JS, Bouyer LJ. Psychometric evidence of self-reported questionnaires for patellofemoral pain syndrome: a systematic review. *Disabil Rehabil*. 2013;35(26):2181-2190. [doi:10.3109/09638288.2013.774061](https://doi.org/10.3109/09638288.2013.774061)
12. Farnsworth JL II, Evans T, Binkley H, Kang M. Evaluation of knee-specific patient-reported outcome measures using rasch analysis. *J Sport Rehabil*. 2021;30(2):278-285. [doi:10.1123/jsr.2019-0263](https://doi.org/10.1123/jsr.2019-0263)
13. Taber KS. The use of cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ*. 2018;48(6):1273-1296. [doi:10.1007/s11165-016-9602-2](https://doi.org/10.1007/s11165-016-9602-2)
14. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. *Optom Vis Sci*. 2007;84(8):663-674. [doi:10.1097/OPX.0b013e318141fe75](https://doi.org/10.1097/OPX.0b013e318141fe75)
15. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*. 2003;80(1):99-103. [doi:10.1207/s15327752jpa8001_18](https://doi.org/10.1207/s15327752jpa8001_18)
16. Panayides P. Coefficient alpha: interpret with caution. *Eur J Psychol*. 2013;9(4):687-696. [doi:10.5964/ejop.v9i4.653](https://doi.org/10.5964/ejop.v9i4.653)
17. Lance CE, Butts MM, Michels LC. The sources of four commonly reported cutoff criteria: what did they really say? *Organ Res Methods*. 2006;9(2):202-220. [doi:10.1177/1094428105284919](https://doi.org/10.1177/1094428105284919)
18. Metsavaht L, Leporace G, Riberto M, de Mello Sposito MM, Batista LA. Translation and cross-cultural adaptation of the Brazilian version of the International Knee Documentation Committee Subjective Knee Form: validity and reproducibility. *Am J Sports Med*. 2010;38(9):1894-1899. [doi:10.1177/0363546510365314](https://doi.org/10.1177/0363546510365314)
19. Dillman DA, Smyth JD, Christian LM. *Internet, Phone, Mail, and Mixed Mode Surveys: The Tailored Design Method*. 4th ed. John Wiley & Sons; 2014.
20. Kline RB. *Principles and Practices of Structural Equation Modeling*. 4th ed. The Guilford Press; 2015.

21. Byrne BM. *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. 3rd ed. Routledge; 2016. [doi:10.4324/9781315757421](https://doi.org/10.4324/9781315757421)
22. Thompson B. *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. American Psychological Association; 2004. [doi:10.1037/10694-000](https://doi.org/10.1037/10694-000)
23. Matsunaga M. How to factor-analyze your data right: do's, don'ts, and how-to's. *Int J Psychol Res*. 2010;3(1):97-110. [doi:10.21500/20112084.854](https://doi.org/10.21500/20112084.854)
24. Courtney M, Gordon R. Determining the number of factors to retain in EFA: using the SPSS R-Menu v2.0 to make more judicious estimations. *Pract Assessment Res Eval*. 2013;18:1-14. [doi:10.7275/9CF5-2M72](https://doi.org/10.7275/9CF5-2M72)
25. Brown T. *Confirmatory Factor Analysis for Applied Research*. 2nd ed. The Guilford Press; 2015.
26. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN Study Design checklist for patient-reported outcome measurement instruments. COSMIN Web site. Published 2019. Accessed February 12, 2022. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf
27. Wood ND, Akloubou Gnonhosou DC, Bowling JW. Combining parallel and exploratory factor analysis in identifying relationship scales in secondary data. *Marriage Fam Rev*. 2015;51(5):385-395. [doi:10.1080/1494929.2015.1059785](https://doi.org/10.1080/1494929.2015.1059785)
28. Hayton JC, Allen DG, Scarpello V. Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. *Organ Res Methods*. 2004;7(2):191-205. [doi:10.1177/1094428104263675](https://doi.org/10.1177/1094428104263675)
29. Park HS, Dailey R, Lemus D. The use of exploratory factor analysis and principal components analysis in communication research. *Hum Commun Res*. 2002;28(4):562-577. [doi:10.1111/j.1468-2958.2002.tb00824.x](https://doi.org/10.1111/j.1468-2958.2002.tb00824.x)
30. Pearson R, Mundfrom D, Piccone A. A comparison of ten methods for determining the number of factors in exploratory factor analysis. *Gen Linear Model J*. 2013;39:1-15.
31. Ruscio J, Roche B. Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychol Assess*. 2012;24(2):282-292. [doi:10.1037/a0025697](https://doi.org/10.1037/a0025697)
32. Schmitt LC, Paterno MV, Huang S. Validity and internal consistency of the international knee documentation committee subjective knee evaluation form in children and adolescents. *Am J Sports Med*. 2010;38(12):2443-2447. [doi:10.1177/0363546510374873](https://doi.org/10.1177/0363546510374873)
33. Higgins LD, Taylor MK, Park D, et al. Reliability and validity of the International Knee Documentation Committee (IKDC) Subjective Knee Form. *Joint Bone Spine*. 2007;74(6):594-599. [doi:10.1016/j.jbspin.2007.01.036](https://doi.org/10.1016/j.jbspin.2007.01.036)
34. Duckett T, Fox CM, Hart JM, Norte GE. Rationale for a parsimonious measure of subjective knee function after ACL Reconstruction: a rasch analysis. *J Athl Train*. 2021;56(12):1340-1348. [doi:10.4085/490-20](https://doi.org/10.4085/490-20)
35. Noorduyn JCA, van de Graaf VA, Mokkink LB, et al. Responsiveness and minimal important change of the IKDC of middle-aged and older patients with a meniscal tear. *Am J Sports Med*. 2019;47(2):364-371. [doi:10.1177/0363546518812880](https://doi.org/10.1177/0363546518812880)
36. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Model*. 2007;14(3):464-504. [doi:10.1080/10705510701301834](https://doi.org/10.1080/10705510701301834)
37. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the International Knee Documentation Committee Subjective Knee Form. *Am J Sports Med*. 2001;29(5):600-613. [doi:10.1177/03635465010290051301](https://doi.org/10.1177/03635465010290051301)
38. Leech NL, Barrett KC, Morgan GA. *IBM SPSS for Intermediate Statistics: Use and Interpretation*. 5th ed. Routledge; 2015. [doi:10.4324/9780203122778](https://doi.org/10.4324/9780203122778)
39. Schönrock-Adema J, Heijne-Penninga M, van Hell EA, Cohen-Schotanus J. Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Med Teach*. 2009;31(6):e226-e232. [doi:10.1080/01421590802516756](https://doi.org/10.1080/01421590802516756)
40. O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behav Res Methods Instruments Comput*. 2000;32(3):396-402. [doi:10.3758/bf03200807](https://doi.org/10.3758/bf03200807)
41. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model*. 1999;6(1):1-55. [doi:10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)
42. Sigelman CK, Rider EA. *Life-Span Human Development*. 6th ed. Cengage Learning; 2009.

43. Ware JE Jr, Kosinski M, Keller SD. A 12-item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-233. doi:10.1097/00005650-199603000-00003
44. Baker RT, Burton D, Pickering MA, Start A. Confirmatory factor analysis of the Disablement in the Physically Active Scale and preliminary testing of short-form versions: a calibration and validation study. *J Athl Train*. 2019;54(3):302-318. doi:10.4085/1062-6050-355-17
45. van de Graaf VA, Wolterbeek N, Scholtes VAB, Mutsaerts ELAR, Poolman RW. Reliability and validity of the IKDC, KOOS, and WOMAC for patients with meniscal injuries. *Am J Sports Med*. 2014;42(6):1408-1416. doi:10.1177/0363546514524698
46. Carpenter S. Ten steps in scale development and reporting: a guide for researchers. *Commun Methods Meas*. 2018;12(1):25-44. doi:10.1080/19312458.2017.1396583
47. Bergman S, Herrstrom P, Hogstrom K, Petersson IF, Svensson B, Jacobsson LT. Chronic musculoskeletal pain, prevalence rates, and sociodemographic associations in a Swedish population study. *J Rheumatol*. 2001;28:1369-1377.
48. Baldwin JN, McKay MJ, Simic M, et al. Self-reported knee pain and disability among healthy individuals: reference data and factors associated with the Knee injury and Osteoarthritis Outcome Score (KOOS) and KOOS-Child. *Osteoarthr Cartil*. 2017;25(8):1282-1290. doi:10.1016/j.joca.2017.03.007
49. Heidari B. Knee osteoarthritis prevalence, risk factors, pathogenesis and features: part I. *Caspian J Intern Med*. 2011;2:205-212.
50. Anderson AS, Loeser RF. Why is osteoarthritis an age-related disease? *Best Pract Res Clin Rheumatol*. 2010;24(1):15-26. doi:10.1016/j.berh.2009.08.006
51. Abramson SB, Attur M. Developments in the scientific understanding of osteoarthritis. *Arthritis Res Ther*. 2009;11(3):227. doi:10.1186/ar2655
52. Williams SN, Wolford ML, Bercovitz A. Hospitalization for total knee replacement among inpatients aged 45 and over: United States, 2000–2010. Published 2015. Accessed February 18, 2022. <https://www.cdc.gov/nchs/data/databriefs/db210.pdf>
53. Guermazi A, Niu J, Hayashi D, et al. Prevalence of abnormalities in knees detected by MRI in adults without knee osteoarthritis: population based observational study (Framingham Osteoarthritis Study). *BMJ*. 2012;345:e5339. doi:10.1136/bmj.e5339
54. Ciccotti MC, Kraeutler MJ, Austin LS, et al. The prevalence of articular cartilage changes in the knee joint in patients undergoing arthroscopy for meniscal pathology. *Arthroscopy*. 2012;28(10):1437-1444. doi:10.1016/j.arthro.2012.02.029
55. Jinks C, Jordan K, Croft P. Measuring the population impact of knee pain and disability with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). *Pain*. 2002;100(1):55-64. doi:10.1016/s0304-3959(02)00239-7
56. Jinks C, Jordan K, Ong BN, Croft P. A brief screening tool for knee pain in primary care (KNEST). 2. Results from a survey in the general population aged 50 and over. *Rheumatology*. 2004;43(1):55-61. doi:10.1093/rheumatology/keg438
57. Brinker MR, Lund PJ, Barrack RL. Demographic biases of scoring instruments for the results of total knee arthroplasty. *J Bone Joint Surg Am*. 1997;79(6):858-865. doi:10.2106/00004623-199706000-00009
58. Ritter MA, Thong AE, Davis KE, Berend ME, Meding JB, Faris PM. Long-term deterioration of joint evaluation scores. *J Bone Joint Surg Br*. 2004;86(3):438-442. doi:10.1302/0301-620x.86b3.14243
59. Paradowski PT, Bergman S, Sundén-Lundius A, Lohmander LS, Roos EM. Knee complaints vary with age and gender in the adult population. Population-based reference data for the Knee injury and Osteoarthritis Outcome Score (KOOS). *BMC Musculoskelet Disord*. 2006;7:38. doi:10.1186/1471-2474-7-38

APPENDICES

APPENDIX 1: INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM: SHORT-FORM 9

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM

SHORT-FORM 9

1. What is the highest level of activity that you can perform without significant knee pain?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework or yard work
 0 ☐ Unable to perform any of the above activities due to knee pain

2. What is the highest level of activity you can perform without significant swelling in your knee?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework, or yard work
 0 ☐ Unable to perform any of the above activities due to knee swelling

3. What is the highest level of activity you can participate in on a regular basis?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework or yard work
 0 ☐ Unable to perform any of the above activities due to knee

4. How does your knee affect your ability to:

		Not difficult at all	Minimally difficult	Moderately difficult	Extremely difficult	Unable to do
a.	Go up stairs	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
b.	Squat	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
c.	Sit with your knee bent	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
d.	Run straight ahead	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
e.	Jump and land on your involved leg	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
f.	Stop and start quickly	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐

APPENDIX 2: INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM: SHORT-FORM 8

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM

SHORT-FORM 8

1. What is the highest level of activity that you can perform without significant knee pain?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework or yard work
 0 ☐ Unable to perform any of the above activities due to knee pain

2. What is the highest level of activity you can perform without significant swelling in your knee?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework, or yard work
 0 ☐ Unable to perform any of the above activities due to knee swelling

3. What is the highest level of activity you can participate in on a regular basis?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging

- 1 ☐ Light activities like walking, housework or yard work
 0 ☐ Unable to perform any of the above activities due to knee

4. How does your knee affect your ability to:

		Not difficult at all	Minimally difficult	Moderately difficult	Extremely difficult	Unable to do
a.	Go up stairs	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
b.	Squat	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
c.	Sit with your knee bent	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
d.	Jump and land on your involved leg	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
e.	Stop and start quickly	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐

APPENDIX 3: INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM: SHORT-FORM 6

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM

SHORT-FORM 6

1. What is the highest level of activity that you can perform without significant knee pain?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework or yard work
 0 ☐ Unable to perform any of the above activities due to knee

- 0 ☐ Unable to perform any of the above activities due to knee pain

2. What is the highest level of activity you can perform without significant swelling in your knee?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework, or yard work
 0 ☐ Unable to perform any of the above activities due to knee swelling

3. What is the highest level of activity you can participate in on a regular basis?

- 4 ☐ Very strenuous activities like jumping or pivoting as in basketball or soccer
 3 ☐ Strenuous activities like heavy physical work, skiing or tennis
 2 ☐ Moderate activities like moderate physical work, running or jogging
 1 ☐ Light activities like walking, housework or yard work
 0 ☐ Unable to perform any of the above activities due to knee

4. How does your knee affect your ability to:

		Not difficult at all	Minimally difficult	Moderately difficult	Extremely difficult	Unable to do
a.	Go up stairs	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
b.	Squat	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐
c.	Sit with your knee bent	4 ☐	3 ☐	2 ☐	1 ☐	0 ☐