

Responsiveness of the International Knee Documentation Committee Subjective Knee Form

James J. Irrgang,^{*†} PhD, PT, ATC, Allen F. Anderson,[‡] MD, Arthur L. Boland,[§] MD, Christopher D. Harner,^{||} MD, Philippe Neyret,[¶] MD, John C. Richmond,[#] MD, K. Donald Shelbourne,^{††} MD, and the International Knee Documentation Committee

From the [†]Department of Orthopaedic Surgery, University of Pittsburgh, Pittsburgh, Pennsylvania, [‡]Tennessee Orthopaedic Alliance, Nashville, Tennessee, [§]Massachusetts General Hospital, Boston, Massachusetts, ^{||}University of Pittsburgh Medical Center for Sports Medicine, Pittsburgh, Pennsylvania, [¶]Hospital de la Croix Rousse, Lyon-Caluire, France, [#]New England Baptist Hospital, Boston, Massachusetts, and ^{††}Methodist Sports Medicine Center, Indianapolis, Indiana

Background and Purpose: The International Knee Documentation Committee Subjective Knee Form was developed to measure change in symptoms, function, and sports activity in patients treated for a variety of knee conditions. Although previous research has demonstrated reliability and validity of the form, its responsiveness has not been evaluated. The purpose of this study was to determine responsiveness of the International Knee Documentation Committee Subjective Knee Form.

Study Design: Cohort study (diagnosis); Level of evidence, 1.

Methods: Patients who participated in the original validation study for the International Knee Documentation Committee Subjective Knee Form completed the form and a 7-level global rating of change scale that ranged from greatly worse to greatly better after a mean of 1.6 years (range, 0.5-2.3 years). Analyses included calculation of the standardized response mean and mean change in International Knee Documentation Committee Subjective Knee Form score compared to the patient's perception of change on the global rating of change scale. In addition, a receiver operating characteristic curve was plotted to determine the change in score that best distinguished patients who improved from those who did not.

Results: The overall standardized response mean was 0.94, which is considered large. With the exception of those who were slightly worse or unchanged, the mean change in the International Knee Documentation Committee Subjective Knee Form score compared to the patients' perceived global ratings of change was as expected (greatly worse, -15.1; somewhat worse, -8.4; slightly worse, 20.6; no change, 10.7; slightly better, 5.9; somewhat better, 18.1; greatly better, 38.7). The receiver operating characteristic curve analysis revealed that a change score of 11.5 points had the highest sensitivity, and a change score of 20.5 points had the highest specificity to distinguish between those who were or were not improved.

Conclusion: The International Knee Documentation Committee Subjective Knee Form is a responsive measure of symptoms, function, and sports activity for patients with a variety of knee conditions.

Keywords: knee; patient-oriented outcomes; responsiveness

*Address correspondence to James J. Irrgang, PhD, PT, ATC, Room 911, Kaufman Building, Department of Orthopaedic Surgery, University of Pittsburgh, 3471 Fifth Avenue, Pittsburgh, PA 15213 (e-mail: jirrgang@pitt.edu).

Other members of the International Knee Documentation Committee include AOSSM: John Bergfeld, John Feagin, John Fulkerson, Mininder Kocher, and Steven Howell; European Society of Sports Traumatology, Knee Surgery and Arthroscopy: Hans Staubli, Fritz Hefti, Jorgen Hoher, Roland Jacob, and Werner Mueller; Asia Pacific Orthopaedic Society for Sports Medicine: K. M. Chan and Matsahiro Kurosaka.

The International Knee Documentation Committee Subjective Knee Form is available from the American Orthopaedic Society for Sports Medicine and may be accessed at <http://www.sportsmed.org/research/IKDC.asp>.

No potential conflict of interest declared.

The International Knee Documentation Committee (IKDC) Subjective Knee Form was developed as a region-specific patient-oriented outcome measure of symptoms, function, and sports activity. It was well tested and found to be a reliable and valid instrument that can be used to measure symptoms, function, and sports activity in patients with a variety of knee impairments, including ligament and meniscal injuries, patellofemoral pain, articular cartilage lesions, and osteoarthritis.⁸

The next step in testing was to determine responsiveness of the IKDC Subjective Knee Form. Responsiveness, the ability of an outcome measure to detect clinically relevant change in the patient's status, is an important psychometric property. A responsive clinical outcome measure will reflect improvement

as a patient's condition improves and deterioration as the patient's condition worsens.

Demonstrating responsiveness requires evidence that the measure accurately detects change when it has occurred. This evidence is provided by linking the change in outcome score to ways used to determine that change, in fact, has occurred—called constructs of change.¹ The focus of the construct of change can be at either the group or individual level. The scores being compared may include scores within patients, between patients, or between groups. A within-patients analysis is performed by comparing scores before and after treatment within an individual patient or group of patients. A between-patients analysis entails a comparison of health status between patients at a single point in time. A between-groups analysis compares the within-group change over time between 2 or more groups. Other aspects of the construct of change that need to be considered involve determining and analyzing the kinds of change being quantified in the responsiveness study. These include (1) change that is considered greater than measurement error, (2) change that is observed over time before and after a treatment, (3) change in those who have improved according to a criterion (gold-standard) measure of change, or (4) change in those who have had a major improvement according to a criterion measure of change.^{1,2}

The purpose of this study was to demonstrate responsiveness of the IKDC Subjective Knee Form. To provide evidence for responsiveness of the IKDC Subjective Knee Form, a within-patients analysis was conducted to determine the ability of the IKDC Subjective Knee Form to detect observed change over time before and after treatment. A between-groups analysis was also performed to determine the ability of the IKDC Subjective Knee Form to detect differences between a group of patients who improved from those who did not improve. For the between-groups analysis, the construct of change was the patients' ratings of their perceptions of change since the initiation of treatment. The between-groups analysis of responsiveness was evaluated at both the group and individual patient level of analysis.

METHODS

Subjects

Patients included in this responsiveness study participated in the original validation study for the IKDC Subjective Knee Form. The sample in the original validation study included 590 patients from 7 orthopaedic sports medicine practices, of whom 533 had less than 10% missing data for the IKDC Subjective Knee Form, that is, those who answered at least 16 of the 18 questions. Follow-up responses were received from 271 patients, of whom 251 answered at least 16 of the 18 questions. The final sample for the analysis of responsiveness included a total of 207 patients with a variety of knee problems who had IKDC Subjective Knee Form scores at baseline and final follow-up. The mean length of follow-up for these patients was 19.0 months (SD, 2.9 months; range, 6-28.0 months). Mean age of these patients was 40.5 years (SD, 16.7 years; range, 12.5-81.3 years), and 47.1% were male.

TABLE 1
Frequency of Diagnoses for Patients in Sample
to Provide Evidence for Responsiveness

Diagnosis	Number ^a
Ligament injury	
ACL	50
Posterior cruciate ligament	1
Medial collateral ligament	11
Nonspecified sprain	2
Dislocation	3
Meniscal injury	
Medial meniscus	43
Lateral meniscus	9
Unspecified	7
Osteoarthritis	76
Patellofemoral injury	
Patellofemoral pain	19
Patellofemoral instability	3
Osteochondritis dissecans	6
Fracture	1
Overuse injury	
Bursitis	2
Tendinitis	2
Nonspecified joint pain	14
Other	
Derangement	1
Rheumatoid arthritis	1
Not recorded	9

^aNumber does not total 207 because some patients had more than 1 diagnosis.

The majority of the patients were white (95.9%), 83.3% graduated from high school, and 49.3% graduated from college. In this sample, 73.0% participated in sports activities, 18% were competitive athletes, and 55% were recreational athletes. The remaining individuals did not participate in sports. Diagnosis was established by the physician based on a history, physical examination, and diagnostic studies. Diagnoses and surgical procedures provided to the patients are summarized in Tables 1 and 2.

Procedure

Patients were contacted by mail to complete the IKDC Subjective Knee Form approximately 18 months after completion of the initial survey. A cover letter that explained the purpose of the study and provided instructions for completion and return of the form was sent to all patients. A stamped addressed envelope was provided to return the completed form. To maximize response rate, a follow-up letter and IKDC Subjective Knee Form were sent to patients who did not respond within 3 weeks.

Patients were asked to rate their changes in knee function on a 7-level global rating scale. The 7 levels were greatly worse, somewhat worse, slightly worse, no change, slightly better, somewhat better, and greatly better compared with the first time they completed the form. The responses to this global rating of change question were used to determine if

TABLE 2
Surgical Procedures Performed on Patients in
Sample to Provide Evidence for Responsiveness

Surgical Procedure	Number ^a
Ligament surgery	
ACL reconstruction	22
Posterior cruciate ligament reconstruction	1
Meniscal procedure	
Medial meniscectomy	19
Lateral meniscectomy	10
Medial meniscus repair	1
Lateral meniscus repair	0
Meniscal repair, unspecified	1
Patellar procedure	
Lateral release	4
Medial reefing	1
Distal realignment	1
Articular cartilage procedure	
Chondroplasty	17
Cartilage cell transplant	1
Total knee arthroplasty	3
Diagnostic arthroscopy	7
Other surgical procedures ^b	8
No surgical procedure	155

^aNumber does not total 207 because some patients had more than 1 surgical procedure.

^bOther surgical procedures include debridement/manipulation, excision of plica, bone grafting, and open reduction and internal fixation.

the patient perceived his or her condition to be worse, the same, or better than it was before treatment. The responses were used as the criterion measure of change from the patient's perspective in the analysis of responsiveness.

Data Management and Analyses

All data were entered into a computerized database for subsequent analyses. All analyses were performed with SPSS version 10.1 for a personal computer (SPSS Science Inc, Chicago, Ill). Descriptive statistics were calculated for all variables. When calculating the change in the IKDC Subjective Knee Form scores, we subtracted the baseline score from the follow-up score so that positive change scores reflected improvement.

To demonstrate responsiveness of the IKDC Subjective Knee Form, 3 different constructs of change were used, and a separate analysis was performed for each.

Group-Level Analysis of Within-Patient Observed Change From Before to After Treatment

The first construct for change was a group-level analysis of the within-patient observed change from before to after treatment. For this analysis, we used the baseline and follow-up scores to calculate an effect size and standardized response

mean. The effect size was calculated by dividing the mean change of the IKDC Subjective Knee Form score by the SD of the baseline scores,¹⁰ and the standardized response mean was calculated by dividing the mean change in the IKDC Subjective Knee Form score by the SD of the change in scores.¹¹ It was hypothesized that overall, patients would improve after treatment, and this would be demonstrated by a relatively large effect size and standardized response mean.

Group-Level Analysis of Between-Groups Difference

The second construct of change was a group-level analysis of the between-groups difference of the change scores from baseline to follow-up. For this analysis, the construct of change was the patient's global rating of change since the initiation of treatment. The patient's responses to the global rating of change were used to create the groups. The change in the IKDC Subjective Knee Form scores from before to after treatment was compared with the patient's perception of change from greatly worse, somewhat worse, slightly worse, no change, slightly better, somewhat better, or greatly better. A 1-way analysis of variance (ANOVA) was used to determine if the differences between groups were statistically significant. Post hoc testing was performed to detect significant pairwise differences between groups. It was hypothesized that the change in IKDC Subjective Knee Form scores from baseline to follow-up would be related to the patient's global perception of change. Specifically, the change in IKDC Subjective Knee Form scores for those who were greatly better would be greater than for those who were somewhat better or less, the change in scores for those who were somewhat better would be greater than for those who were slightly better or less, and so forth.

Individual-Level Analysis of Between-Groups Difference

The final construct of change used to demonstrate responsiveness of the IKDC Subjective Knee Form involved analysis of the data at the individual level to determine the optimum change in IKDC Subjective Knee Form score that identified a patient who perceived improvement from a patient who was unchanged or worse. In quality of life research, the change score that optimally identified an improved patient has been termed the *minimum clinically important difference*.¹ This analysis allowed us to identify the best change score to determine if an individual patient perceived himself or herself to be improved. For this analysis, the patients' global ratings of change were dichotomized to create 2 groups—one that was improved, consisting of those who perceived their change to be somewhat or greatly better, and one that was not improved, consisting of those who perceived their change to be slightly better or less. Using this dichotomized measure of the patients' global ratings of change, the sensitivity and specificity of each observed value of the IKDC Subjective Knee Form change score were determined. Sensitivity of change was calculated as the proportion of patients who improved according to global rating of change who had a change score at or above the cutoff point for the IKDC Subjective Knee Form change score. Specificity of change was

calculated as the proportion of patients who did not improve according to the patients' global perceptions of change who had a change score below the cutoff point.⁶ Sensitivity and specificity were calculated using multiple IKDC Subjective Knee Form change scores as the cutoff value.

To determine the minimum clinically important difference, that is the change in IKDC Subjective Knee Form score that best distinguished those who were improved from those who were not improved, the sensitivities and specificities for each of the potential change scores were plotted as a receiver operating characteristic curve. A receiver operating characteristic curve is a plot of sensitivity on the vertical axis and 1-specificity on the horizontal axis. The area under the curve is interpreted as the probability of identifying an improved patient on the basis of the IKDC Subjective Knee Form change score from randomly selected pairs of improved and unimproved patients.⁶ The value of the change score that is closest to the upper left-hand corner of the receiver operating characteristic curve is the IKDC Subjective Knee Form change score that best distinguishes an improved patient from one who is not improved. A test that perfectly discriminates between an improved and unimproved patient would have a receiver operating curve that was a right angle curve proceeding up the vertical axis to a sensitivity of 100% with a specificity of 100% (ie, the false-positive rate would be 0%). A number of individuals involved with quality of life research have recommended a receiver operating characteristic curve analysis to determine the minimum clinically important difference.^{1,6,16} In our opinion, this analysis provides the clinician with valuable information when using the IKDC Subjective Knee Form to monitor individual patient progress.

RESULTS

Group-Level Analysis of Within-Subjects Observed Change From Before to After Treatment

The mean baseline IKDC Subjective Knee Form score was 44.8 (SD, 19.4; median, 42.2; range, 9 to 100), and the mean follow-up IKDC Subjective Knee Form score was 66.8 (SD, 22.9; median, 73.6; range, 5 to 100). The mean improvement was 22.0 (SD, 23.4; median, 17.7; range, -59 to 88). The resulting effect size and standardized response mean were 1.13 and 0.94, respectively. The effect size is considered large according to standards described by Cohen,⁵ and the standard response mean is considered large according to Liang et al.¹¹ These results imply that the mean magnitude of the change score was more than 1 SD of the baseline scores and slightly less than 1 SD of the change scores.

Group-Level Analysis of Between-Groups Difference

The IKDC Subjective Knee Form change scores by the patients' global ratings of change are reported in Table 3. With the exception of those that were slightly worse or unchanged, the mean IKDC Subjective Knee Form change score by the patients' global ratings of change was as expected. Because of the heterogeneity of the variances between the groups, we used the Welch approximation of the 1-way ANOVA,¹² and the

resulting F test was significant ($F_{6,24} = 32.6, P < .001$), indicating the mean change score was different between the groups. Based on the sum of squares from the ANOVA, 37.1% of the variance in the IKDC Subjective Knee Form change score was accounted for by the patients' global ratings of change. Post hoc testing was performed with the Scheffé test for the Welch model.¹² The results indicated that the mean IKDC Subjective Knee Form change score for those who rated themselves to be greatly improved was significantly greater than the mean change score for those who were somewhat improved, slightly improved, unchanged, somewhat worse, and greatly worse. In addition, those who rated themselves greatly worse had change scores that were lower than scores of those who rated themselves slightly worse, unchanged, slightly better, somewhat better, and greatly better. Post hoc testing failed to reveal significant differences for all other pairwise comparisons between groups; thus, our hypothesis that the IKDC Subjective Knee Form change scores from baseline to follow-up would be related to the patients' global perceptions of change was only partially supported.

Individual-Level Analysis of Between-Groups Difference

The receiver operating characteristic curve is presented in Figure 1. The area under the curve is 0.78 (95% confidence interval, 0.72-0.85), which is significantly greater than an area of 0.5. This indicates that the IKDC Subjective Knee Form change score is significantly better than chance of identifying an improved patient from randomly selected pairs of improved and unimproved patients. There were 2 points on the curve that were closest and of equal distance to the left upper corner of the curve. These 2 points corresponded to the IKDC Subjective Knee Form change scores of 11.5 and 20.5. The sensitivity and specificity of the change score of 11.5 were 0.82 and 0.64, respectively. For the change score of 20.5, the sensitivity and specificity of change were 0.64 and 0.84, respectively.

DISCUSSION

Analysis of all 3 constructs of change used in this study provided evidence to support responsiveness of the IKDC Subjective Knee Form to measure change in symptoms, function, and sports activity in patients treated for a variety of knee conditions.

Group-Level Analysis of Within-Subjects Observed Change From Before to After Treatment

For the first construct of change, it was hypothesized that patients with a variety of knee impairments would improve with treatment. The results supported this hypothesis; the mean IKDC Subjective Knee Form score was 44.8 at presentation and 66.8 at follow-up. The mean IKDC Subjective Knee Form change score was 22.0. Both the effect size and standardized response mean were large according to standards described by Cohen⁵ and Liang et al,¹¹ respectively. The effect size relates the mean IKDC Subjective Knee Form change score to the variability of the baseline scores, and the

TABLE 3

Descriptive Statistics of the IKDC Subjective Knee Form Change Scores by the Patients' Perceived Ratings of Change^a

Rating of Change	No. of Subjects	Mean	SD	Median	Minimum Score	Maximum Score
Greatly worse	3	-15.1	4.9	-16.3	-19	-10
Somewhat worse	7	-8.4	26.1	0.0	-59	24
Slightly worse	15	20.0	19.2	16.8	-8	61
No change	28	10.7	21.4	7.9	-51	62
Slightly better	22	5.6	11.1	7.0	-18	24
Somewhat better	53	17.8	14.4	17.4	-13	44
Greatly better	75	38.6	21.8	35.0	-3	88
Total	203 ^b	22.0	23.5	17.7	-59	88

^aIKDC, International Knee Documentation Committee. Change scores were calculated by subtracting the baseline score from the final score. Positive scores indicate improvement over time.

^bFour subjects did not provide a global rating of change, and the change scores for these subjects are not included.

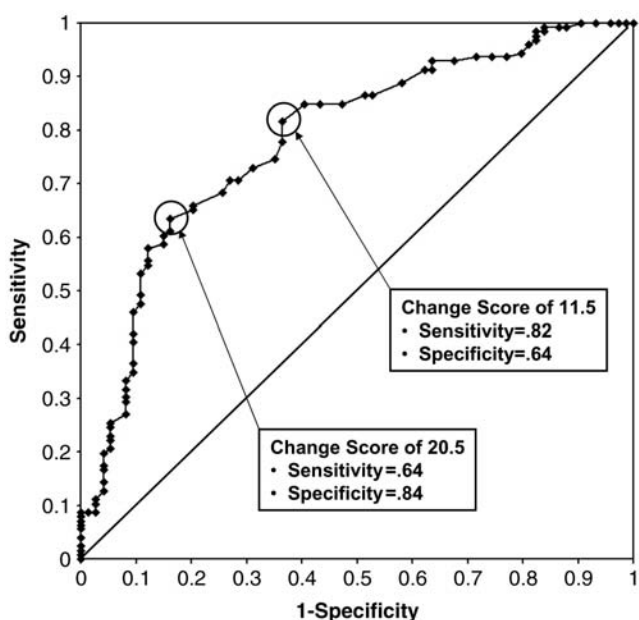


Figure 1. Receiver operating characteristic curve.

standardized response mean relates the change score to variability of the change scores. The standardized response mean is closely linked to the paired *t* test. The effect size and standardized response mean can be related to statistical power.⁵ All things being equal, greater statistical power is achieved and fewer subjects are required to reach statistical significance when the effect size and standardized response mean are large.

Group-Level Analysis of Between-Groups Difference

The second construct of change used the patients' global ratings of change in knee function as a criterion (gold-standard) measure of change. It was hypothesized that the IKDC Subjective Knee Form change scores would be dependent on

the patients' global ratings of change. With the exception of those who rated themselves slightly worse or unchanged, the mean change in the IKDC Subjective Knee Form score was as expected. Statistically, those who rated themselves as greatly better had a greater mean IKDC Subjective Knee Form change score than did those who rated themselves somewhat better, slightly better, unchanged, somewhat worse, or greatly worse. In addition, those who rated themselves greatly worse had a mean IKDC Subjective Knee Form change score that was significantly lower than that of those who rated themselves slightly worse, unchanged, slightly better, somewhat better, or greatly better. There were no other significant differences in the IKDC Subjective Knee Form change scores between those who were somewhat better, slightly better, unchanged, slightly worse, or somewhat worse. Analysis of the sum of squares from the ANOVA indicated that only 37.1% of the variability in the IKDC Subjective Knee Form change scores was accounted for by the patients' global ratings of change. Thus, our hypothesis that the IKDC Subjective Knee Form change scores from baseline to follow-up would be related to the patients' global ratings of change was only partially supported. Failure to fully support this hypothesis may be owing to problems with the construct of change that used the patients' retrospective global ratings of change as the criterion measure of change or may reflect lack of responsiveness of the IKDC Subjective Knee Form itself.

Patients who rated themselves as slightly worse or unchanged were evaluated in greater detail to determine the reason for the discrepancy between their global ratings of change and the change in IKDC Subjective Knee Form scores. Fifteen patients reported that their conditions were slightly worse. Eleven of these patients had final IKDC Subjective Knee Form scores that were less than 60. Of the remaining 4 patients, 2 had final IKDC Subjective Knee Form scores between 70 and 79, 1 had a score between 80 and 89, and 1 had a score greater than 90. A similar pattern was noted for patients who reported that their conditions were unchanged. Of the 28 patients who reported their conditions were unchanged, 16 had final IKDC Subjective Knee Form scores less than 60, 2 had scores between 60 and 69, 5 had scores between 70 and 79, 1 had a score between 80 and 89, and 4

had final IKDC Subjective Knee Form scores greater than 90. Although the mean change in IKDC Subjective Knee Form scores for those who rated themselves as slightly worse or unchanged was 20.0 and 10.7, respectively, many of these individuals still had relatively low final IKDC Subjective Knee Form scores.

It appears that although these patients may have improved on the basis of the change in IKDC Subjective Knee Form scores, many still perceived themselves to be limited. It is possible that they rated the change in their conditions relative to their status before injury as opposed to the first time that they completed the IKDC Subjective Knee Form. Although patients may have improved after they completed the IKDC Subjective Knee Form the first time, they may have rated their status as slightly worse or unchanged relative to their status before injury. A case that illustrates this point is a patient who was injured in March 1999. He had an IKDC Subjective Knee Form Score of 10 in April 1999 and a final score of 46 in November 2000. Although the improvement in the IKDC Subjective Knee Form score was 36 points, it is likely he rated himself to be slightly worse in comparison to his status before injury rather than to the first time he completed the form in April 1999. Future research to determine responsiveness needs to account for this problem when using the patient's retrospective global rating of change as the criterion measure of change. In addition, alternative methods to determine the criterion measure of change need to be investigated. One such method proposed by Beaton et al³ used the patient's ability to cope with the problem as the measure of change. If the patient was unable to cope with the problem at baseline but was able to do so at follow-up, then change had occurred.

Individual-Level Analysis of Between-Groups Difference

The third construct of change used a receiver operating characteristic curve to determine the change in the IKDC Subjective Knee Form score necessary to distinguish a patient who perceived improvement from a patient who did not perceive improvement. The results revealed 2 changes in IKDC Subjective Knee Form scores that were of equal distance from the left-hand corner of the receiver operating characteristic curve. A change score of 11.5 had a sensitivity of 0.82 and specificity of 0.64, and a change score of 20.5 had a sensitivity of 0.64 and specificity of 0.84 for detecting a patient who perceives improvement. The cutoff score of 11.5 should be used to maximize the sensitivity of change, and a change score of 20.5 should be used to maximize the specificity of change.

To apply and interpret the information related to the sensitivity and specificity of change, it is useful to use the acronyms "SnOut" and "SpIn" popularized by Sackett et al.¹⁴ The acronym SnOut indicates that for a measure that has high sensitivity, a negative test result effectively rules out the condition, which in this case means that the individual does not perceive himself or herself to be improved. Thus, if a patient has a change score less than 11.5, it is not likely that this individual would perceive his or her condition to be improved. The acronym SpIn means that for a measure with

high specificity, a positive test result effectively rules in the condition, which in this case means that the individual perceives himself or herself to be improved. Thus, if an individual's change score is greater than 20.5, it is likely that this individual would perceive his or her condition to be improved. Clinicians can apply these results to individual patients to determine if the patient perceives improvement. In our opinion, these cutoff scores will place an individual patient's change in IKDC Subjective Knee Form score in a more meaningful context.

The IKDC Subjective Knee Form cutoff score of 11.5 identified in this study corresponds closely to the minimal detectable change for the IKDC Subjective Knee Form. Previously, the test-retest reliability coefficient for the IKDC Subjective Knee Form was reported to be 0.94. The standard error of measurement (SEM) calculated with the test-retest reliability coefficient was 4.6, resulting in a minimal detectable change for the IKDC Subjective Knee Form score of ± 12.8 . Consequently, changes in the IKDC Subjective Knee Form score of less than 12.8 points may represent sampling error due to the occasion of testing.¹ The close agreement of these results provides additional evidence that a change in the IKDC Subjective Knee Form score of 11.5 is necessary to distinguish between those who have improved and those who have not improved.

Evaluation of baseline and follow-up IKDC Subjective Knee Form scores for 2 of the most common diagnostic groups, those with an ACL injury and those with osteoarthritis, provides additional evidence of validity and responsiveness of the IKDC Subjective Knee Form. Fifty patients had an ACL injury. The mean IKDC Subjective Knee Form scores for these patients were 49.5 (SD, 22.0; median, 44.6; range, 15 to 88) at baseline and 77.2 (SD, 18.1; median, 80.4; range, 20 to 100) at follow-up. The mean change in score for these patients was 27.7 (SD, 27.7; median, 26.1; range, -59 to 83). Seventy-six patients had knee osteoarthritis. The mean IKDC Subjective Knee Form Scores were 37.3 (SD, 16.3; median, 34.7; range, 10 to 83) at baseline and 51.5 (SD, 19.7; median, 50.2; range, 5 to 91) at follow-up. The mean change in score for those with osteoarthritis was 14.2 (SD, 16.8; median, 12.5; range, -19 to 57). As one would expect, patients with knee osteoarthritis had lower IKDC Subjective Knee Form scores than did patients with an ACL injury. This was evident at both baseline and follow-up. One would also expect a greater change in the IKDC Subjective Knee Form score for those undergoing treatment for an ACL injury as opposed to treatment for knee osteoarthritis. These data also provide researchers with an estimate of the treatment effect for clinical studies designed to evaluate treatment of ACL injuries and osteoarthritis, which in turn could be used to estimate sample size.

Limitations of Study

A limitation of this study was the use of the patients' global ratings of change to provide evidence for responsiveness. Although a retrospective global rating of change has been used in several responsiveness studies,^{7,9,15} Norman¹³ criticized its use as the criterion measure of change, stating that the reliability and validity of the retrospective global rating

of change have not been well established. Furthermore, the retrospective judgment of change is psychologically challenging for a patient, and the global rating of change is not independent of the patient's responses on the functional status scale. Consequently, use of the patient's global rating of change may cause errors in measurement of responsiveness.

The difficulty in making a judgment of change from the past may partially explain the change scores for those who rated themselves to be slightly worse or unchanged. Patients may find it hard to differentiate change from the time they first completed a health status measure as opposed to their status before injury. To rate change, patients must quantify their initial and present states and perform "mental subtraction." In addition, they may be unable to recall their initial states when making a rating of change. If this were the case, one would expect the rating of change to be positively correlated with the present state and uncorrelated with the initial state. In our study, this was the case. The correlations between patients' global ratings of change and the final and baseline IKDC Subjective Knee Form scores were 0.59 and 0.09, respectively. Any bias in the patients' judgments at the time of the final observation, such as temporary exacerbation or remission of the disease, may also influence the overall judgments of their present states and perception of change on a global rating scale.

Future research to provide evidence for responsiveness should consider alternative designs that minimize reliance on the patients' retrospective global ratings of change. These designs may contrast groups that are expected to undergo differential rates of change over time, such as groups undergoing treatments that have known or hypothesized differences in effectiveness (eg, debridement vs tibial osteotomy for individuals with grade III medial compartment osteoarthritis associated with varus alignment) or groups that are expected to undergo different rates of change because of the natural or clinical history of the condition (eg, change over 12-month period of time for those who have recently had ACL reconstruction vs those who had ACL reconstruction 3 to 5 years ago).

In summary, the IKDC Subjective Knee Form, a well-standardized outcomes instrument, has been proven to be reliable and valid. The evidence presented in this study also demonstrates that the IKDC Subjective Knee Form has acceptable levels of responsiveness to measure change in symptoms, function, and sports activity over time in patients with a variety of knee impairments. Future research is needed to compare responsiveness of the IKDC Subjective Knee Form to other knee-specific and general measures of health status.

Addendum

The minimal detectable change is calculated by multiplying the SEM by $1.96 \times \sqrt{2}$. In this case, the SEM should be based on the test-retest correlation coefficient. The SEM is multiplied by 1.96, which is the standard normal score that is necessary to create a 2-tailed 95% confidence interval and the $\sqrt{2}$ to take into account that there is error associated with both the first and second scores.⁴

REFERENCES

1. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. *Spine*. 2000;25:3192-3199.
2. Beaton DE, Bombardier C, Katz JN, et al. Looking for important change/differences in studies of responsiveness. *J Rheumatol*. 2001; 28:400-405.
3. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J Hand Ther*. 2001;14:128-146.
4. Christensen L, Mendoza JL. A method of assessing change in a single subject: an alteration of the RC Index. *Behav Ther*. 1986;17:305-308.
5. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Academic Press; 1969.
6. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis*. 1986;39:897-906.
7. Heald SL, Riddle DL, Lamb RL. The Shoulder Pain and Disability Index: the construct validity and responsiveness of a region-specific disability measure. *Phys Ther*. 1997;77:1079-1089.
8. 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;29:600-613.
9. Irrgang JJ, Synder-Mckler L, Wainner RS, Fu FH, Harner CD. Development of a patient-reported measure of function of the knee. *J Bone Joint Surg Am*. 1998;80:1132-1145.
10. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care*. 1989;27:S178-S189.
11. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care*. 1990;28:632-642.
12. Marascuilo LA, Serlin RC. *Statistical Methods for the Social and Behavioral Sciences*. New York, NY: WH Freeman; 1988.
13. Norman GR. Issues in the use of change scores in randomized trials. *J Clin Epidemiol*. 1989;11:1097-1105.
14. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine*. 2nd ed. Boston, Mass: Little, Brown; 1991.
15. Stratford PW, Binkley J, Solomon P, Gill C, Finch E. Assessing change over time in patients with low back pain. *Phys Ther*. 1994;74:528-533.
16. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther*. 1996;76:1109-1123.