0363-5465/100/2828-0850$02.00/0
THE AMERICAN JOURNAL OF SPORTS MEDICINE, Vol. 28, No. 6

Comparison of Score Evaluations and Instrumented Measurement After Anterior Cruciate Ligament Reconstruction

Ralph Hrubesch,*† MD, Christoph Rangger,* MD, Markus Reichkendler,* MD, Romed Franz Sailer,* MD, Wulf Gloetzer,* MD, and Günther Eibl‡

*From the *University Hospital of Traumatology and the ‡Institute for Biostatistics and Documentation, Innsbruck, Austria*

ABSTRACT

Forty-four patients who had undergone unilateral anterior cruciate ligament reconstructions were evaluated retrospectively with seven different scoring systems (International Knee Documentation Committee, Orthopädische Arbeitsgruppe Knie, Lysholm, Feagin and Blake, Zarins and Rowe, Cincinnati, and Marshall scores). The results varied between systems and therefore lacked reliability. Of the 44 patients, 32 were rated as excellent according to the Cincinnati score while only 3 patients were rated as normal according to the International Knee Documentation Committee form. Good and excellent results were found twice as frequently with the Cincinnati and Lysholm scores compared with the scores of Zarins and Rowe or the International Knee Documentation Committee form. Statistical analysis confirmed this observation and revealed significant differences between the scoring systems. Side-to-side differences using the manual maximum displacement test with the KT-1000 arthrometer revealed good correlation with the International Knee Documentation Committee and the Orthopädische Arbeitsgruppe Knie questionnaires. None of the other scoring systems, which do not measure anterior laxity, produced reasonable correlation with instrumented measurements. We found that certain population-specific factors as well as the distribution of single findings can distort the results of scoring systems. To avoid these interference factors, the patient sample should be homogeneous and selected prospectively and there should be agreement about the value of single findings.

† Address correspondence and reprint requests to Ralph Hrubesch, MD, Universitätsklinik für Unfallchirurgie, Anichstrasse 35, A-6020 Innsbruck, Austria.

No author or related institution has received any financial benefit from research in this study.

Recently, numerous scoring systems have been introduced to evaluate treatment after disruption of the ACL. These systems usually use a numeric system to rate findings such as pain, swelling, subjective assessment of function, level of activity, and others. The summarized results are usually classified as excellent, good, fair, or poor, and these final results are based on the computation of distinct numbers of points.

Because certain findings are given more weight in one scoring system than in another, a single pathologic finding can affect the overall outcome. For example, the International Knee Documentation Committee form does not use a numeric classification, rather it includes terms such as normal or abnormal. A poor result in one area reduces the overall score.

Instrumented measurements have been found to be useful in the determination of knee instability,[5,13,14] but a correlation of instrumented measurements with the results of knee scoring scales has not been frequently found.[7,13] The aim of this study was to determine the accuracy and reliability of these different scoring scales designed for knee injuries. The results of the rating systems were compared with each other and with results of instrumented laxity testing.

MATERIALS AND METHODS

From March 1993 through October 1996, 65 patients underwent surgery by the senior surgeon (WG) after disruption of the ACL. Sixty-three patients had unilateral first-time ACL reconstructions, and two patients had revision surgery after disruption of an ACL graft. Patients with bilateral injuries were excluded from the present study. Forty-four patients who met the criteria returned for postoperative assessment. Eighteen women and 26 men, average age 33 years (range, 17 to 57), could be observed for a mean of 19 months (range, 9 to 36) after surgery. Twenty-two patients had an ACL tear in the right knee, and 22

patients had a tear in the left knee. Thirty-three patients had concomitant injuries in the same knee: 28 meniscal tears, 4 medial collateral ligament tears, and 1 lateral collateral ligament tear. Damage of varying extent to the articular surface, due to either traumatic or degenerative causes, was noted in 36 knees. In 17 patients the injury occurred during winter sports, in 12 patients in ball sports, and in 15 patients the injury was not sports related.

Surgery was performed within 1 week after trauma in 12 patients, within 6 weeks in 14, and later than 6 weeks after injury in 18 patients (mean, 7.5 weeks). In all patients, the indication for surgery was established and arthroscopically assisted intraarticular reconstruction of the ACL with an ipsilateral autogenous patellar tendon graft was performed. The specific indication for surgery was disabling instability of the knee. On the femoral side, the transplant was fixed with an Endobutton (Acufex Microsurgical Inc., Mansfield, Massachusetts). On the tibial side, the transplant was attached to sutures on resorbable anchors that were introduced transcortically distal to the tunnel orifice. Fourteen medial and 10 lateral meniscal tears were resected, and 1 medial and 1 lateral tear were repaired with resorbable anchors. Two minor lesions of the medial meniscus and four lesions of the collateral ligaments were left to heal. In one patient the medial collateral ligament was repaired. After surgery the patients were mobilized with crutches in a knee brace with toe-touch weightbearing. The brace allowed 70° of flexion and full extension and was removed after 6 weeks. Aggressive physical therapy was started on the 2nd day after surgery and continued up to 12 weeks after surgery. The postoperative physical therapy protocol included passive motion therapy, particularly extension, isokinetic muscle training, and walking exercises without the brace. Recreational sports were allowed after 6 weeks and full activity was allowed 1 year after surgery.

At follow-up, patients were tested by one examiner (RH) who had not been involved in the initial treatment. Evaluation included assessment of knee joint function, subjective assessment according to different scoring systems, and instrumented measurements using the KT-1000 arthrometer (MEDmetric Corp., San Diego, California).[5,14] All rating systems were used separately within one session following the authors' guidelines. Manual laxity testing, as described for the single scoring systems, was performed before instrumented measurements. Finally, the data were entered into a computer and processed using the Microsoft Excel program (Microsoft Corporation, Redmond, Washington).

Seven questionnaires from the literature were used:

1. The International Knee Ligament Standard Evaluation Form (modified in 1993–1994) according to the International Knee Documentation Committee (IKDC)[8]
2. The Orthopädische Arbeitsgruppe Knie (OAK) score by the Swiss Orthopedic Society[11]
3. The Lysholm and Gillquist score[9,15]
4. The Feagin and Blake score[7]
5. The Zarins and Rowe score[16]
6. The Cincinnati score (Noyes)[12]
7. The Marshall score[10]

Comparison of Grading and Categories Among the Scoring Systems

When comparing scoring systems, it appears necessary to find an overall valid syntax. Each of the seven studied systems uses a different mode of presenting the results in terms of grading, categorization, and designation.

Most systems have in common the conversion of functional and clinical findings into numeric values of different levels. However, in the IKDC form, findings are graded into four categories (A, B, C, D) without numeric assessment. The lowest grade within a group determines the group grade. The poorest group grade among the first four groups determines the final result.[8]

The numeric systems, except the Feagin and Blake scale (maximum 30 points) and the Marshall scale (maximum 50 points), use 100 points for assessment. These scores use different numeric ranges for the outcome categories and in some the categories themselves are different (see Table 1). For our comparison, the terminology of the OAK and Lysholm scores was used (excellent, good, fair, and poor), and these categories were based on the numeric value for the particular scale. The Feagin and Blake and Zarins and Rowe scores use numeric scores only, so, for comparative reasons, grading into categories was done according to the terminology of the OAK score.

To allow comparison between numeric systems with different levels of maximum points, as with the systems of Feagin and Blake or Marshall, a calibration to a maximum value of 100 was performed. Thus, the results of these two systems, when necessary, were calculated as percentages.

Comparison of the IKDC Form with Other Scoring Systems

For the IKDC categories, each category (normal, nearly normal, abnormal, and severely abnormal) was compared with the numeric results and categories of the other systems.

Comparison with Instrumented Laxity Measurements

A side-to-side difference (involved minus noninvolved side), as measured with the manual maximum displacement test, was compared with the results of the seven rating systems.

Statistical Analysis

Statistical analysis was performed with the Statistical Program for Social Sciences (SPSS, Chicago, Illinois). Comparison of categories was assessed with the chi-square and Kendall's tau test. P values of <0.05 were considered significant. Correlation between the numeric scoring systems and the KT-1000 arthrometer measurements was determined with Spearman's test.

RESULTS

The correlation between outcomes on the seven scales and the KT-1000 arthrometer measurements is shown in Table 1.

TABLE 1
Correlation Between the Seven Rating Systems and KT-1000 Arthrometer Measurements

Scoring scale	Grading	Categories[a]	No. of patients	Arthrometer side-to-side difference (mm)[b]		Correlation *P* (Spearman)
				Mean value	Range	
IKDC		Normal	3	1.0	1	
		Nearly normal	15	2.5	0–7	
		Abnormal	24	3.7	0–9	
		Severely abnormal	2	6.0	6	
OAK		Excellent	9	1.9	0–4	0.319
		Good	22	3.0	0–7	
		Fair	10	4.3	2–9	
		Poor	3	6.0	5–7	
Lysholm	100–96 pts.	Excellent	10	3.5	1–7	0.146
	95–84 pts.	Good	29	3.3	0–9	
	83–65 pts.	Fair	5	4.0	0–7	
	less than 65 pts.	Poor	0			
Feagin and Blake[c]	100%–91%		11	1.5	0–3	0.530
	90%–81%		15	2.9	0–7	
	80%–71%		12	4.9	1–9	
	less than 70%		6	4.2	2–7	
Zarins and Rowe	100–91 pts.		8	1.8	0–4	0.410
	90–81 pts.		12	3.3	0–9	
	80–71 pts.		15	3.2	1–5	
	less than 70 pts.		9	4.8	0–7	
Cincinnati	100–86 pts.	Category I	32	2.9	0–9	0.426
	85–66 pts.	Category II	10	3.7	1–7	
	65–51 pts.	Category III	1	7.0	7	
	less than 51 pts.	Category IV	1	6.0	6	
Marshall	50–41 pts.	Excellent/good	17	2.9	0–9	0.363
	40–36 pts.	Fair+	17	3.2	0–7	
	35–31 pts.	Fair−	8	3.8	2–7	
	less than 31 pts.	Poor	2	6.0	6	

[a] For the Feagin and Blake and Zarin and Rowe scales categories were assigned according to the OAK scale terminology.
[b] Arthrometer testing was performed as manual maximum test involved minus noninvolved knee.
[c] Percentage of maximum points.

Comparison of the IKDC Score Outcome with the Other Scoring Systems

The three patients who were rated normal according to the IKDC scale had excellent results in the other scoring systems (Fig. 1). They were graded as follows in the other scoring systems: OAK score, one patient with 99 points and two patients with 95 points; Lysholm score, two patients with 100 points and one patient with 98 points; Feagin and Blake, two patients attained 100% and one patient attained 97% of 30 points; Zarins and Rowe, one patient had 100 points and two patients had 95 points; Cincinnati score, all three patients were graded with 100 points; Marshall score, patients obtained 92%, 94%, and 96% of 50 points.

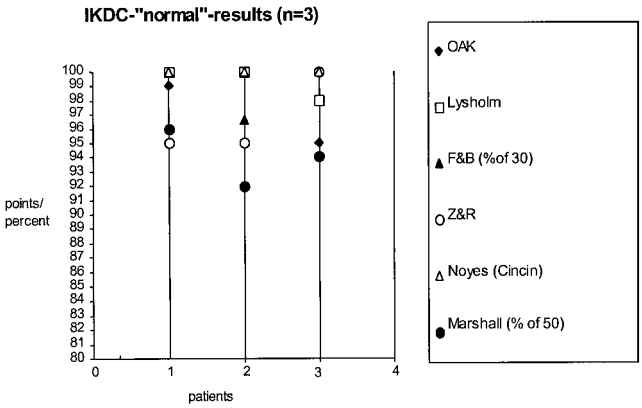

Figure 1. Results on all scoring systems for the patients rated "normal" on the IKDC scoring system. F&B, Feagin and Blake; Z&R, Zarins and Rowe.

Fifteen patients were rated nearly normal according to the IKDC form (Fig. 2). They were graded as follows in the other scoring systems: OAK score, range from 83 to 95 points; Zarins and Rowe, range from 74 to 93 points; Cincinnati score, range from 88 to 100 points; Lysholm score, range from 87 to 100 points; Feagin and Blake, range from 80% to 100% of 30 points; Marshall score, range from 66% to 96% of 50 points.

Twenty-four patients were rated abnormal according to the IKDC form (Fig. 3). They were graded as follows in the other scoring systems: OAK score, range from 62 to 98 points; Zarins and Rowe, range from 48 to 90 points; Cincinnati score, range from 64 to 100 points; Lysholm score, range from 77 to 100 points; Feagin and Blake, range from 67% to 100% of 30 points; Marshall score, range from 58% to 88% of 50 points.

Two patients were rated severely abnormal according to the IKDC form (Fig. 4). They were graded as follows in the other scoring systems: OAK score, one patient with 68

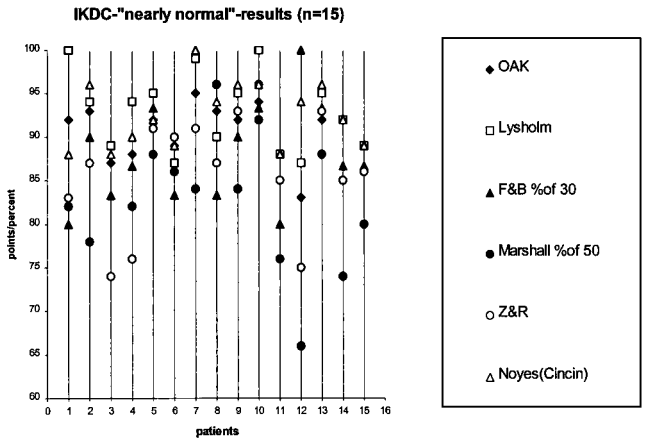

Figure 2. Results on all scoring systems for the patients rated "nearly normal" on the IKDC scoring system. F&B, Feagin and Blake; Z&R, Zarins and Rowe.

points and one patient with 73 points; Lysholm score, one patient with 77 points and one patient with 90 points; Feagin and Blake, one patient attained 50% and one patient 77% of 30 points; Zarins and Rowe, one patient had 47 points and one patient had 62 points; Cincinnati score, one patient was graded with 40 and one patient with 86 points; Marshall score, one patient attained 52% and one patient 72% of 50 points.

Comparison of categories revealed that those rated normal on the IKDC form had similar high outcomes on the other scores. Significant differences were found between IKDC outcome and the other scales for patients who had nearly normal or abnormal ratings on the IKDC form. Two patients who had been rated severely abnormal according to the IKDC form also had poor results on the other scoring systems, except for one patient who was considered excellent on the Cincinnati score and good on the Lysholm and Marshall scores.

Comparison of categories showed significant differences

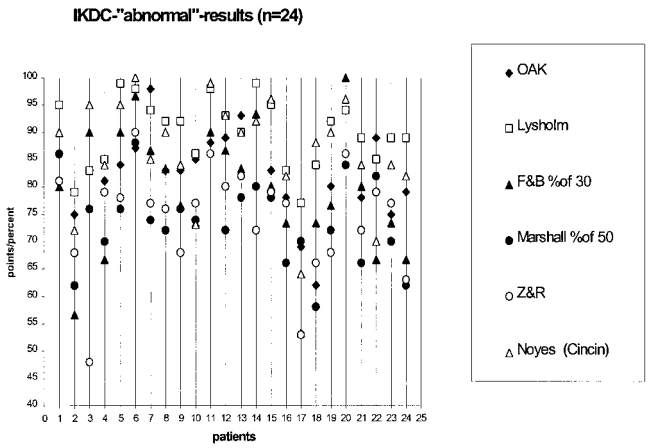

Figure 3. Results on all scoring systems for the patients rated "abnormal" on the IKDC scoring system. F&B, Feagin and Blake; Z&R, Zarins and Rowe.

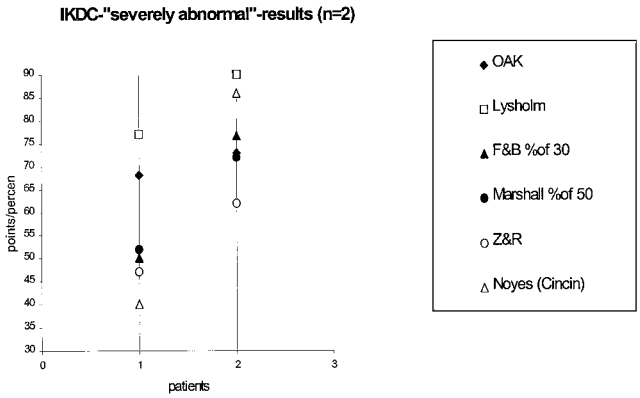

Figure 4. Results on all scoring systems for the patients rated "severely abnormal" on the IKDC scoring system. F&B, Feagin and Blake; Z&R, Zarins and Rowe.

in five patients. Four patients who were rated abnormal according to the IKDC form obtained excellent results on the Cincinnati score (three patients) and the Marshall score (one patient), while they were rated poor on the scores of Zarins and Rowe (three patients) and Feagin and Blake (one patient). One patient who was rated severely abnormal according to the IKDC form achieved an excellent result on the Cincinnati score but a poor result on the score of Zarins and Rowe.

Table 2 presents a comparison of categories according to the IKDC form, with the average number of points/percentage achieved in the other scoring systems.

Comparison of KT-1000 Arthrometer Measurements and Rating Systems

Table 1 demonstrates the distribution of mean laxity testing results with the KT-1000 arthrometer in comparison with categories of the studied rating systems. Patients classified as normal on the IKDC form had a small side-to-side difference on KT-1000 arthrometer measurements, while patients with nearly normal, abnormal, and severely abnormal results demonstrated an increasing side-to-side difference on KT-1000 arthrometer measurements. The correlation between the OAK score system and KT-1000 arthrometer measurements was similar. Associa-

TABLE 2
Average Points or Percentage Achieved in the Numeric Scoring Systems According to IKDC Outcome

Scoring scale	IKDC rating			
	Normal	Nearly normal	Abnormal	Severely abnormal
OAK	96.3	89.7	82.2	70.5
Lysholm	99.3	92.9	90	83.5
Feagin and Blake[a]	98.9	88.6	79	63.3
Zarins and Rowe	96.7	86.2	74.2	54.5
Cincinnati	100	92.5	86.6	63
Marshall[b]	94	82	74.1	62

[a] Based on a percentage of 30 points.
[b] Based on a percentage of 50 points.

TABLE 3
Statistical Analysis of Correlation for the Different Scoring Scales Used in This Study

Scoring scale	OAK	Lysholm	Feagin and Blake	Zarins and Rowe	Cincinnati	Marshall
			Chi-square test			
IKDC	$P < 0.001$	0.011	0.012	$P < 0.001$	$P < 0.001$	0.007
OAK		0.006	$P < 0.001$	$P < 0.001$	$P < 0.001$	$P < 0.001$
Lysholm			0.001	0.002	0.001	0.005
Feagin and Blake				$P < 0.001$	$P < 0.001$	$P < 0.001$
Zarins and Rowe					0.016	$P < 0.001$
Cincinnati						$P < 0.001$
			Kendall's tau test			
IKDC	0.691	0.383	0.531	0.682	0.480	0.490
OAK		0.461	0.588	0.693	0.439	0.600
Lysholm			0.525	0.526	0.478	0.535
Feagin and Blake				0.614	0.623	0.546
Zarins and Rowe					0.496	0.676
Cincinnati						0.538

tions between KT-1000 arthrometer measurements and the results of other scoring systems are definitely incoherent.

Statistical Analysis

The chi-square test and Kendall's tau test did not reveal any correlation between the numeric scoring systems used in this study. The results of individual scores were significantly different from each other (Table 3).

DISCUSSION

Several scoring systems have been introduced to evaluate knee joint function after ligament injuries. Assessment of pre- and postoperative status serves as a basis for deciding the therapeutic procedure. The objective of these questionnaires is to achieve objective and comparable results. Nevertheless, our evaluation of 44 patients after unilateral ACL reconstruction using 7 different scoring systems revealed significant differences. Additionally, instrumented measurements showed poor correlation with most of the scores.

Classification of results is based on objective and subjective parameters. These fare differently in the numeric scoring systems. Pain accounts for 30% in the Lysholm score, 20% in the Cincinnati score, 13.3% in the Feagin and Blake score, 10% in the Zarins and Rowe score, 5% in the OAK score, and 2% in the Marshall score. Differences were also registered in the assessment of return to preoperative activity. In the Lysholm score, return to walking, running, and jumping accounts for 70%; in the Cincinnati score, function and overall activity level account for 30%; in the Feagin and Blake score, activity level accounts for 16.6%; in the OAK score, work and sports account for 10%; in the Zarins and Rowe score, athletics accounts for 10%; and in the Marshall score, return to sports or work accounts for 6%.

Single findings, which may be rated high or low, can distort the overall result.[8] The various aspects are evaluated separately in the OAK score, whereas other scoring systems merely summarize points.[7,9,10,12,15,16] The latter approach may be a source of inaccuracy, especially if the results are poor, as was seen in five patients in the present study who were simultaneously rated from excellent to poor in different scoring systems. Patients with increased anterior laxity, for instance, will have fewer points in a scoring system that assesses this defect. For example, the Cincinnati score does not take anterior laxity into account and therefore produces better results, especially in patients who do not suffer subjectively from their anterior laxity. On the other hand, the Zarins and Rowe score does not elucidate subjective problems such as trouble with jumping and twisting activities.

Beside these influencing factors, which are inherent to the systems, there may be others that distort the overall results in a way that makes fixed correlations between scoring systems impossible. Specific patient samples may produce different result levels because of the influence of population-specific factors such as sex, age, activity level, or education.[6,12] Therefore, the attempt to compare treatment procedures on the basis of a patient sample that is not selected prospectively is inadmissible in scientific terms. It is neither reliable nor reproducible.

Many researchers have experienced differing results with the application of various single scoring systems. Consequently, the use of two or more systems provides comparable results. However, not every author always uses the same two or three rating systems. This has led to the necessity of trying to translate the results from one system to another so that comparisons can be made between studies. As one solution, some authors have tried to find a fixed mathematical proportion between scoring systems and establish correlation coefficients to facilitate comparison of results.[2,15] Coefficients such as 0.78 (comparing Lysholm and Marshall[15]) and 0.87 (comparing Lysholm and Cincinnati[2]) have been established. Within our sample, we found correlation coefficients of 0.85 (Lysholm to Marshall) and 0.97 (Lysholm to Cincinnati) when comparing these studies. Bollen and Seedhom[2] examined 41 patients and determined that the Cincinnati score led to a better outcome than did the Lysholm score. This agrees

with our observations, in spite of the fact that the factor of correlation within our sample was about 10% higher than that found by Bollen and Seedhom. Tegner and Lysholm[15] calculated a correlation coefficient comparing the Lysholm and the Marshall score in 76 patients. They considered differences in results due to emphasis on criteria like stability or muscular strength and meniscus status. The proportion we found between the two systems in the present study was, again, different.

Christel et al.[4] compared 3 scoring systems in 90 patients who had undergone ACL reconstructions with augmentation as described by Marshall and MacIntosh. As in our study, they achieved good and excellent results in 92.5% according to the Lysholm score. However, they had only 82% good results according to the IKDC scale. In the present study, only 41% of patients had normal or nearly normal results on the IKDC score, which is half that found by Christel et al. In a group of 52 patients, Müller et al.[11] achieved good and excellent results after ACL reconstruction in 75% using the OAK score, and in 82.7% using the Marshall score. We had 70% good and excellent results with the OAK score, and 77% with the Marshall score. Buess et al.,[3] using the OAK score, obtained good or excellent results in 93% of 14 patients after acute PCL surgery and in 38% of 16 patients after chronic PCL surgery. The IKDC score, however, revealed nearly normal results in only 29% of the acute and 6% of the chronic cases. They found no patients with normal results. For the patients with reconstruction of acute tears, this signifies a correlation coefficient of 0.31, which markedly differs from our study (0.58). Plancher et al.[13] reviewed 72 patients with 75 ACL reconstructions using 3 different scoring systems. Ninety-three percent were considered normal and nearly normal according to the IKDC score, and all patients had 69 to 100 points with the Lysholm score. The authors gave no information about the distribution of results for the latter system, but there were obvious differences regarding correlation compared with our findings. Demirdjian et al.[6] found differences when investigating 418 healthy knee joints with the Cincinnati and Lysholm knee scoring questionnaires. The "exchange" rate for the two systems was calculated to be 1 for the male subjects, which is comparable with our value (0.97). This matches with our observation that results for patients with a nearly healthy knee joint correlate well in all studied scoring systems.

Among others, the IKDC evaluation form is based on principles from the OAK score. All parameters are weighted equally and group qualifications are used to determine the overall result, which cannot be better than the worst qualification of the first four groups of problems (subjective assessment, symptoms, range of motion, ligament examination).[8] Therefore, in the present study, it is not surprising that a greater agreement was found between results with the IKDC form and the OAK score than between these two scores and the other scoring systems. Nevertheless, these two systems also differ from each other.

A comparison between rating systems seems to be permissible if the results are similar, as in the three patients with normal IKDC results (also see Demirdjian et al.). However, the poorer the results in different scoring systems, the less comparable they are. In this case the influence of single findings, depending on their "value" in different scoring systems, appears to be important. Different patterns of detected defects lead to different result levels. If most of the findings are rated with maximum values, all studied scoring systems produce excellent overall results. However, lower ratings may distort the result, as explained earlier. The relationship found between the Lysholm score and the IKDC form by Christel et al. differs significantly from that registered by us. Considered strictly scientifically, the relation of scoring systems within one patient sample should always be the same, but, depending on different characteristics of patient samples, they change. Correlation coefficients as calculated by Tegner and Lysholm[15] or Bollen and Seedhom[2] appear to be representative only for their specific patient populations. From this perspective, the ratings of excellent, good, fair, and poor become highly unreliable for scientific assessment.

Instrumented measurements have been found to be useful in the determination of knee instability.[1,5,13,14] According to the IKDC form, the examiner is advised to perform additional manual, radiological, or instrumented measurements. Manual testing may involve observer-biased differences.[8] In our patients, KT-1000 arthrometer measurements correlated well with the results of the IKDC form and with the OAK score. Patients who were rated normal demonstrated less side-to-side differences than did patients who were rated worse. Apparently, correlation depends on the fact that testing and rating of laxity is laid down in the instructions for these two systems. Inadequate correlation was found between KT-1000 arthrometer measurements and the other scoring systems.

CONCLUSIONS

The most important finding of the present study is that rating systems do not produce interchangeable results. The relationship between the results of two scoring systems for one patient sample may be different in another sample. Thus, scoring systems appear to be unreliable tools for scientific assessment. Even more, a comparison between different samples based on rating systems must lack reliability. The major question is which are the most important items in the assessment of the injured knee? No agreement exists on this subject or about the values of different findings. The diversity and differences between patient samples and findings assessed in different scoring systems reduces the credibility of ratings like good or poor for a procedure outcome. Furthermore, the question arises as to whether results from scoring systems are reliable enough to make therapeutic decisions. To minimize the errors in evaluating surgical procedures or treatment modalities according to score systems the aim has to be homogeneity of the patient sample and an agreement about the value of single findings to avoid interference factors, which inevitably lead to wrong results. Consequently patient samples should be selected prospectively.

We found the IKDC evaluation form to have the best correlation with instrumented laxity measurements.

REFERENCES

1. Bach BR Jr, Warren RF, Flynn WM, et al: Arthrometric evaluation of knees that have a torn anterior cruciate ligament. *J Bone Joint Surg 72A:* 1299–1306, 1990
2. Bollen S, Seedhom BB: A comparison of the Lysholm and Cincinnati knee scoring questionnaires. *Am J Sports Med 19:* 189–190, 1991
3. Buess E, Imhoff AB, Hodler J: Knee evaluation in two systems and magnetic resonance imaging after operative treatment of posterior cruciate ligament injuries. *Arch Orthop Trauma Surg 115:* 307–312, 1996
4. Christel P, Djian P, Darman Z, et al: Results of Marshall-MacIntosh reconstruction according to 3 scoring systems (ARPEGE, Lysholm, IKDC). 90 cases reviewed with at least a one-year follow-up [in French]. *Rev Chir Orthop Reparatrice Appar Mot 79:* 473–483, 1993
5. Daniel DM, Malcom LL, Losse G, et al: Instrumented measurement of anterior laxity of the knee. *J Bone Joint Surg 67A:* 720–726, 1985
6. Demirdjian AM, Petrie SG, Guanche CA, et al: The outcomes of two knee scoring questionnaires in a normal population. *Am J Sports Med 26:* 46–51, 1998
7. Feagin JA Jr, Blake WP: Postoperative evaluation and result recording in the anterior cruciate ligament reconstructed knee. *Clin Orthop 172:* 143–147, 1983
8. Hefti F, Müller W: Heutiger Stand der Evaluation von Kniebandläsionen. *Orthopäde 22:* 351–362, 1993
9. Lysholm J, Gillquist J: Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am J Sports Med 10:* 150–154, 1982
10. Marshall JF, Fetto JF, Botero PM: Knee ligament injuries. A standardized evaluation method. *Clin Orthop 123:* 115–129, 1977
11. Müller W, Biedert R, Hefti F, et al: OAK knee evaluation. A new way to assess knee ligament injuries. *Clin Orthop 232:* 37–50, 1988
12. Noyes FR, Barber SD, Mooar LA: A rationale for assessing sports activity levels and limitations in knee disorders. *Clin Orthop 246:* 238–249, 1989
13. Plancher KD, Steadman JR, Briggs KK, et al: Reconstruction of the anterior cruciate ligament in patients who are at least 40 years old. A long-term follow-up and outcome study. *J Bone Joint Surg 80A:* 184–197, 1998
14. Rangger C, Daniel DM, Stone ML, et al: Diagnosis of an ACL disruption with KT-1000 arthrometer measurements. *Knee Surg Sports Traumatol Arthrosc 1(1):* 60–66, 1993
15. Tegner Y, Lysholm J: Rating systems in the evaluation of knee ligament injuries. *Clin Orthop 198:* 43–49, 1985
16. Zarins B, Rowe CR: Combined anterior cruciate-ligament reconstruction using semitendinosus tendon and iliotibial tract. *J Bone Joint Surg 68A:* 160–177, 1986