

K. Labs · B. Paul

To compare and contrast the various evaluation scoring systems after anterior cruciate ligament reconstruction

Received: 28 August 1995

Abstract In a prospective study we evaluated with eight established scoring systems 56 patients who received a Leeds-Keio ligament as an anterior cruciate ligament replacement. The systems were compared based upon their total results, as well as their subjective, objective and functional criteria. It was found that the diversity of both the results and criteria within the systems made valid comparisons impossible. From this investigation it was concluded that the use of an internationally standardized evaluation system would provide the best possibility for comparison of surgical results. As a result, we recommend the IKDC Score system together with a visual analogue scale to evaluate subjective knee complaints.

Introduction

Evaluate the surgical results of an anterior cruciate ligament (ACL) reconstruction to objectively, several postoperative evaluation systems have been developed. In 1955, O'Donoghue first applied a metric scale as a means of scoring his postoperative results. The goal of this pioneer works was based on a 100-point scoring system which was used to quantify and compare his results (Flandry et al. 1991). This system was found to be inadequate in differentiating between subjective, objective and functional criteria. Subsequently, a multitude of schemes was developed, based upon different documentation criteria (verbal, numerical, binary, visual) with regards to postoperative results (Geens et al. 1974; Larson 1974; Marshall et al. 1977; Aichroth et al. 1978; Oretrop et al. 1979; Turba et al. 1979; Lysholm and Gillquist 1982; Feagin and Blake 1983; Katelkamp and Thompson 1984; Tegner and Lysholm 1985; Zarins and Rowe 1986; Lukianov et al. 1987; Windsor et al. 1988; Flandry et al. 1991; Laboureau

and Cazenave 1991). Current international studies reveal a 70%–85% very good and good result after surgical reconstruction of the ACL when utilizing these various scoring systems (Fujikawa et al. 1989; Shelbourne et al. 1990). The question of comparability of these various scoring systems based upon subjective, objective and functional parameters has not been thoroughly investigated. Comparison studies with established scores have been documented by Tegner and Lysholm (1985), Kannus and Jarvinnen (1987), Schmid and Bandi (1988), Bollen and Seedhom (1991) as well as by Hefti et al. (1991). In all cases the results were not in agreement.

The need for an effective evaluation system was recognized in 1987 by certain members of the European Society for Sports Traumatology, Knee Surgery and Arthroscopy (ESSKA) and the American Orthopaedic Society of Sport Medicine (AOSSM). This group, referred to as the International Knee Documentation Committee, developed the IKDC Score system (Hefti and Müller 1993). Minimal criteria are contained in this system, meaning that certain points are minimal with regard to their value. In our study we investigated earlier evaluation systems and IKDC Score system. The value of their subjective, objective and functional criteria was analyzed as well as their statement ability.

Materials and methods

In our study, patients were examined at the Orthopaedic Department of Humboldt University, Charité Hospital, from June 1990 to February 1993 following an ACL replacement. All patients with chronic knee joint instability had received a Leeds-Keio ligament replacement. The follow-up study was done an average of 24.2 months (19–32 months) postoperatively. This study involved 56 patients (42 men, 14 women) with a mean age of 28.9 years. We utilized the following knee investigation schemes:

Lysholm II Scale, Clancy Score, Gaudernak Score, Marshall Score, Hospital for special surgery Score (HSS), modified AOSSM Score (American Orthopaedic Society for Sports Medicine), OAK Score (Orthopädische Arbeitsgruppe Knie) and IKDC Score (International Knee Documentation Committee). Furthermore, all patients were examined with a instrumented (KT-1000) manual maximum test.

K. Labs (✉) · B. Paul
Department of Orthopaedics, Charité-Hospital,
Humboldt-University, Berlin, Schumannstrasse 20/21,
D-10117 Berlin, Germany

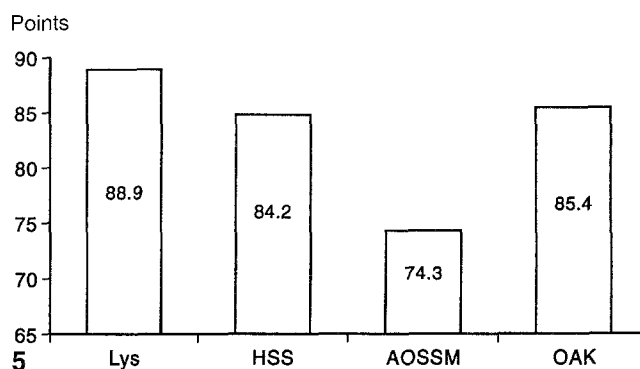
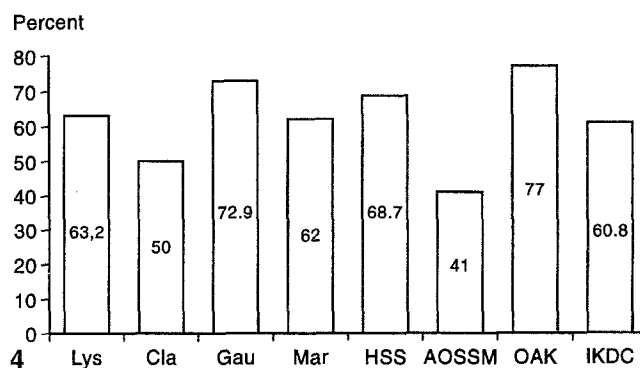
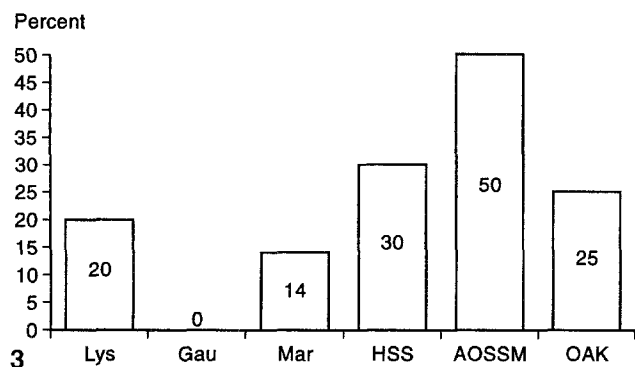
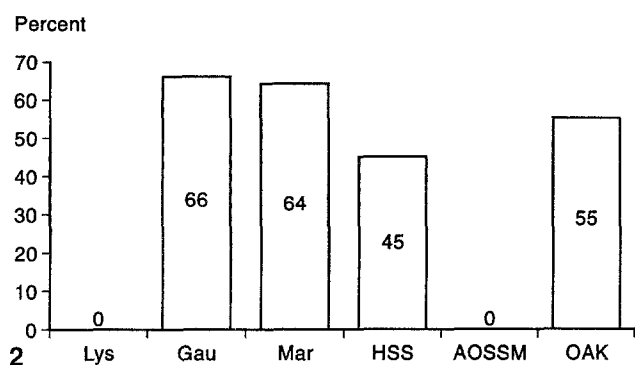
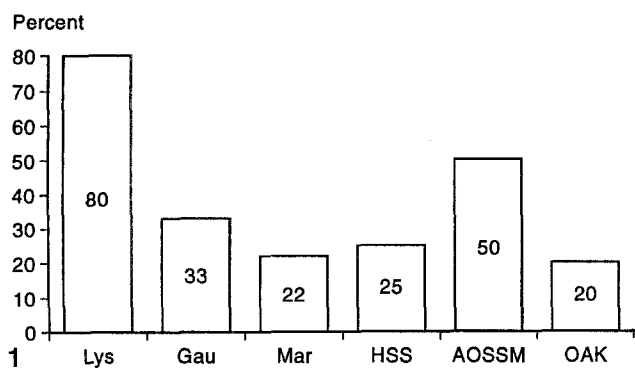


Fig. 1 Maximum percentage for *subjective criteria* among six different scoring systems: *Lys* Lysholm II Score, *Gau* Gaudernak Score, *Mar* Marshall Score, *HSS* Hospital for Special Surgery, *AOSSM* American Orthopaedic Society for Sports Medicine, *OAK* Orthopädische Arbeitsgruppe Knie

Fig. 2 Maximum percentage for *objective criteria*

Fig. 3 Maximum percentage for *functional criteria*

Fig. 4 Overview of very good and good clinical results with the different scoring systems: *Cla* Clancy Score, *IKDC* International Knee Documentation Committee

Fig. 5 Average results for 100-point scales

The following is a concise description of the scores obtained from each system.

1. The Lysholm II Score is a 100-point score with eight subcriteria, including 3 functional and 5 subjective. Altogether, 50% of the total score is based on the symptoms of pain and instability. Documentation by the patient her- or himself is made possible and applied categorically. The score has a clear numerical differentiation (Lysholm and Gillquist 1982).

2. The Clancy Score utilizes a verbal estimate for subjective, objective and functional criteria. The results are divided into excellent, good, fair or poor. The worst level determined the final result. It is simple and fast. The pitfall of this system is that no objective results are obtained, and it is possible to overemphasise one particular question complex (Clancy 1988).

3. The Gaudernak Score consists of three question groups: subjective appraisal, mobility and stability. There is only one possible answer per complex. The score is very easy to fill out, but the statements are inaccurate. This scoring system does not consider function and neglects the status of the knee joint (Gaudernak 1984).

4. The Marshall Score is a 50-point score with four subcategories. Each individual subcategory contributes unequally to the total result: subjective parameters (22%), functional test (14%), knee status (24%) and ligament laxity (40%). This system deals partially with a binary answer system, and the functional statements are underemphasised. The score is short and very practicable (Marshall et al. 1977).

5. The HSS Score is a 100-point score, which evolved from the Marshall Score. Subjective parameters are greatly reduced. Special criteria for sport ability are utilized. The objective ligament stability test can reach a maximum of 45% within the total result. Particularly interesting is the appraisal of the patient's pain symptom, which when present, is deducted from the total result (maximum 10%) (Windsor et al. 1988).

6. The AOSSM modified from Buckley et al. (1989) a 100-point score, of which 50% is for subjective symptoms and 50% for function. For the first time the patients have been given the opportunity to document their own functional result.

7. The OAK Score is a 100-point system with problem criteria having a respective proportional influence in the total result: A discomfort/swelling (20%); B mobility/power (15%); C stability (40%); and D function (25%). This score is very clear and concise, and problems are recorded immediately. The lowest level of a category is partially included into the final result (Hefti et al. 1991).

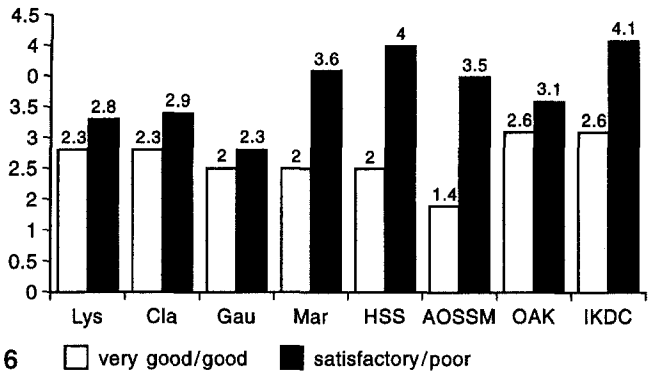
8. The results of the IKDC Score are documented via verbal descriptions of normal, nearly normal, abnormal and severely abnormal based on eight investigation criteria. For evaluation there are four problem areas (subjective assessment, symptoms, range of motion and ligament examination). These are supplemented by four areas that are documented but not included in the evaluation (compartmental findings, donor site pathology, X-ray findings and functional tests). All parameters are of equal value and add up to 100%. The worst single value determines the final result (Hefti and Müller 1993).

Results

Figures 1–3 show graphically the proportional distribution of individual criteria and the total result. These figures do not include the IKDC and the Clancy scoring systems because each single criterion may influence the total result of 100%. The value of the subjective criteria comprises a very high percentage with regards to the Lysholm Score, followed by the AOSSM Score. Other scores have concentrated between 1/5 and 1/3 of their point values on the subjective criteria. Based upon objective parameters the Gaudernak and the Marshall Scores have the highest influence on the total results. It is striking that in the Lysholm and AOSSM Scores no consideration of objective parameters is included in the total result. With regard to functional parameters the AOSSM Score exceeds all other scoring systems with 50% of its total result based upon function. One can also compare the share of very good and good results of the OAK and Gaudernak Scores, which have a very high percentage, versus the low percentage from the AOSSM and Clancy Scores. Surprisingly, we noted a low percentage (63.2%) of very good and good results with the Lysholm Score, which had been previously described as being much higher (Fig.4). The maximum achievable score of 100 points in the Lysholm, HSS, AOSSM and OAK Scores demonstrates on average clear differences (Fig.5).

Interestingly, the findings reveal that the HSS Score has a slighter lower median value but a high proportional share of very good and good results than the Lysholm Score. Similarly, the OAK Score demonstrates the highest share of very good and good results (77%), but its maximum point level is slightly lower when compared with the Lysholm scoring system. The assignment of very good and good results to the manual maximum test for knee laxity revealed a much higher proportion of satisfactory and poor results in all systems evaluated (Fig.6). No agreement was found between the laxity measurements and the assignment of laxity and proportional share by satisfactory and poor results (Fig.7). There was a relatively high share of satisfactory and poor results in the Clancy Score but a lower median value in the manual maximum test in contrast to the slightly lower share of satisfactory

Millimeter side difference



Millimeter and percent

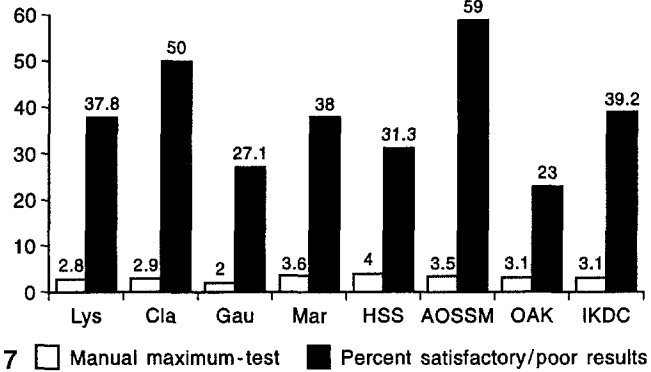


Fig.6 Relationship between manual maximum test (KT-1000) and clinical result

Fig.7 Relationship between knee ligament laxity (KT-1000) and satisfactory or poor results

Table 1 Case report 1 (woman, 31 years old) with a overview of all results

Test/score	Value
Manual maximum	- 1 mm
Lysholm	Poor
Clancy	Fair
Gaudernak	Excellent
Marshall	Satisfactory
HSS	Poor
AOSSM	Poor
OAK	Satisfactory
IKDC	Abnormal

Table 2 Case report 2 (man, 29 years old) with a overview of all results

Test/score	Value
Manual maximum	10 mm
Lysholm	Very good
Clancy	Good
Gaudernak	Excellent
Marshall	Very good
HSS	Normal
AOSSM	Normal
OAK	Very good
IKDC	Abnormal

and poor results in the OAK Score, which reveals a higher laxity. Two case reports give an overview of all scores used and the manual maximum test (Tables 1 and 2). With regard to one patient with a clinically stable knee joint, all scoring systems revealed a satisfactory or poor result with the exception of the Gaudernak Score. Another discrepancy was found in the manual maximum test where a patient with a 10-mm difference had predominantly very good and good results, the one exception being the IKDC system. The final result of the IKDC system radically and solely through the existing ligament laxity achieved an abnormal result.

Discussion

The selection of a knee scoring system based on the total result is of great importance. From this investigation it is obvious that the OAK, Gaudernak and HSS Scores have outstandingly achieved better results. The explanation for this lies in the devaluation of functional parameters and an emphasis on objective criteria. So, for example in the OAK Score, the patient's inability to perform a duck gait will still allow a good total result to be obtained. The Gaudernak Score with its three possible answers cannot describe the complete knee ligament status, and one can achieve reciprocal results in comparison with other evaluation systems. Far worse results were determined with the modified AOSM and Clancy Scores and average results were obtained with the Lysholm, Marshall and IKDC Scores. The AOSM Score with its emphasis on functional criteria reflects the productive power of the knee joint in special activities in combination with the subjective feeling. A ligament examination is completely excluded.

There are limited possibilities with the Clancy Score, and this is comparable to the all or nothing principle. Both objective as well as functional criteria have a high question level and can therefore give a negative total result. The weakness of the Lysholm Score lies in the overemphasis of subjective symptoms. The low median value in our investigation is due to the fact that the investigation was not performed by the physician, and therefore the answers from the patient were not influenced by the physician. This system proved to be of great value because it excluded the physician from the questioning process. The Marshall Score through its partially binary documentation devalues the functional criteria and may result in the loss of accurate information.

The IKDC Score as result of long-term development is unique in its construction and has greater potential than other scoring systems. The smallest deviation from normal shows clearly deleterious results. The amount of very good and good results will decrease consistently with further application internationally in the future. The final result of this evaluation system can be influenced by the first question put to the patient by the physician. We recommend that physicians exhaustively explain to the patient the item being asked. A further criticism involves the

documentation of a 4–5 deg loss of extension, which give a nearly normal result vs a 6–7 deg loss of extension, which gives an abnormal result. It is doubtful whether such slight differences can be measured accurately. Still more impressive are the findings in patients with complete subjective satisfaction and unrestricted functional athletic activity, but a measurable laxity, as in the second case report presented. In that case only an abnormal total result was possible. It is reassuring to note that the IKDC system has included the one-leg hop test within its parameters because of its sensitivity to the functional restoration of the knee joint.

The documentation form is very important. The verbal documentation form has a limited scope, numerical systems make a clear differentiation possible, binary systems can cause loss of information, and visual systems have a high sensitivity. The visual analogue scales (VAS) for pain described by Scott and Huskisson (1976) and a VAS for knee complaints described Flandry et al. (1991) could be come in the future extremely valuable additions to the IKDC Score. An improved statement in association with VAS results may allow for a better subjective assessment of the patient.

In conclusion, it is clear that different evaluation systems are not comparable. There exists individual variation according to subjective, objective and functional parameters. Each follow-up is based upon the physician's interest in the problem, his or her experience and knowledge, the sensitivity of the scores, the patient's compliance, and interpretive influences. Frequently, disagreements exist between the subjective satisfaction and the objective results. The patient may not feel what we measure, and we may measure something which the patient does not feel.

There is a need for an uniform evaluation system which must be imposed universally. With the IKDC Score a corresponding basis is created, and in combination with the VAS Score there is the possibility, for the first time, to achieve valid and comparable results. The use of the IKDC evaluation form produces less favourable results than with other current systems, because a persisting knee problem cannot be hidden behind a high numerical score added for other, unrelated parameters. It is hoped that this new form will enable us to compare treatment methods in various publications with each other.

References

- Aichroth P, Freeman MAR, Smillie IS (1978) A knee function assessment chart. *J Bone Joint Surg [Br]* 60B: 308–309
- Bollen S, Seedhom B (1991) A comparison of the Lysholm and Cincinnati knee scoring questionnaires. *Am J Sports Med* 19 2: 189–190
- Buckley SL, Barrack RL, Alexander AH (1989) *Am J Sports Med* 17: 221–225
- Clancy WR Jr (1988) Acute tears of the anterior cruciate ligament – grading of the results. *J Bone Joint Surg [Am]* 70: 1483–1488
- Feagin JA, Blake WP (1983) Postoperative evaluation and result recording in the anterior cruciate ligament tears: *Clin. Orthop* 172: 143–147

- Flandry F, Hunt JP, Terry GC, Hughston JC (1991) Analysis of subjective knee complaints using visual analog scales. *Am J Sports Med* 19:112-117
- Fujikawa K, Iseki F, Seedham B (1989) Arthroscopy after anterior cruciate reconstruction with the Leeds-Keio-Ligament: *J Bone Joint Surg [Br]* 4B:566-570
- Gaudernak W (1984) Nachuntersuchungsscore für Kniegelenksverletzungen. *Hefte Unfallheilk* 167:174-178
- Geens S, Clayton ML, Liedholt JD, Smyth CJ, Bartholomev BA (1974) One-hundred point scale for assessment of knee disability. In: Smillie I (ed) *Diseases of the knee joint*. Churchill Livingstone, Edinburgh, pp 28-35
- Hefti F, Müller W (1993) Heutiger Stand der Evaluation von Kniebinnenläsionen. *Orthopäde* 22:351-362
- Hefti F, Drobny T, Hackenbruch W, Kipfer W, Holzach P (1991) Dokumentation und Evaluation von Kniegelenkinstabilitäten. In: Jakob RP, Stäubli HU (eds) *Kniegelenk und Kreuzbänder*. Springer, Berlin Heidelberg New York, pp 138-144
- Kannus P, Jarvinnen M (1987) Conservatively treated tears of the anterior cruciate ligament. Long-term results. *J Bone Joint Surg [Am]* 69:1007-1012
- Katalkamp D, Thompson C (1984) Development of a knee scoring scale. *Clin Orthop* 107:93-96
- Laboureaux JP, Cazenave A (1991) Functional system AEPEGE. *Rev Chir Orthop* 77:92-102
- Larson R (1974) Rating sheet for knee function. In: Smillie I (ed). *Diseases of the knee joint*. Churchill Livingstone, Edinburgh, pp 29-38
- Lukianov AV, Gillquist J, Grana WA, DeHaven KE (1987) An anterior cruciate ligament evaluation format for assessment of artificial or autologous anterior cruciate reconstruction results. *Clin Orthop* 218:167-180
- Lysholm J, Gillquist G (1982) Evaluation of the knee ligament surgery results with special emphasis on a scoring scale. *Am J Sports Med* 10:150-154
- Marshall JL, Fetto JF, Botero PM (1977) Knee ligament injuries. A standardized evaluation method. *Clin Orthop* 123:115-119
- O'Donoghue DH (1955) An analysis of surgical treatment of major injuries to the ligaments of the knee. *J Bone Joint Surg [Am]* 37:1-13
- Oretrop N, Gillquist J, Lijedahl S-O (1979) Long term results of surgery for non-acute anteromedial rotatory instability of the knee. *Acta Orthop Scand* 50:329-332
- Schmid K, Bandi F (1988) Die Verletzungen des medialen Seitenbandes (MSB) und vorderen Kreuzbandes (VKB) am Kniegelenk und ihre operativ-funktionelle Behandlung nach Lemaire. *Unfallchirurgie* 14:118-124
- Scott J, Huskisson EC (1976) Graphic representation of pain. *Pain* 2:175-184
- Shelbourne KD, Whitetaker HJ, Carroll JR, Retting AC, Hirschman LD (1990) Anterior cruciate ligament injury: evaluation of intraarticular reconstruction of acute tears without repair. Two to seven year follow up of 155 athletes. *Am J Sports Med* 18:484-488
- Tegner Y, Lysholm J (1985) Rating system in the evaluation of knee ligament injuries. *Clin Orthop* 118:43-49
- Turba JE, Walsh MW, McLeod WD (1979) Long-term results of extensor mechanism reconstruction. A standard for evaluation. *Am J Sports Med* 7:91-96
- Windsor RE, Insall JN, Wickiewicz TL (1988) The Hospital for Special Surgery knee ligament rating form. *Am J Knee Surg* 1:140-145
- Zarins B, Rowe CR (1986) Knee rating score. *J Bone Joint Surg [Am]* 68:160-177