International Orthopaedics (SICOT) (1988) 12: 291–297

International
Orthopaedics

A radiological scoring scale for evaluation of post-traumatic osteoarthritis after knee ligament injuries

P. Kannus[1, 4, 5], M. Järvinen[2, 5], and T. Paakkala[1, 3]

[1] Department of Clinical Sciences, University of Tampere, SF-33520 Tampere, Finland
[2] Department of Surgery, University Central Hospital of Tampere, Finland
[3] Department of Radiology, University Central Hospital of Tampere, Finland
[4] Research Unit for Sport and Physical Fitness, Jyväskylä, Finland
[5] Tampere Research Station of Sports Medicine, Tampere, Finland

Summary. *We have developed a radiological scoring scale for the evalutation of post-traumatic osteoarthritic changes after a knee ligament injury. It is based on radiographs taken from the knee at the time of injury, and from both knees at the follow-up examination. Ten criteria were used to devise and list the scale. The method has been used on 60 patients operated on primarily after an acute knee ligament injury. The mean age of the patients when injured was 36.6 years, and they were re-examined at an average of three and a half years after the operation. The scale includes adjustments so that the final scores measure only the pathological changes caused by the ligament injury. Reproducibility of the scale is excellent and the scores correlate well with the clinical results. The registration form is simple, may be completed easily and revised when necessary. The numerical data are readily prepared for statistical and computer analyses. Ninety five per cent of all the changes caused by the ligament injuries fitted the scale, showing success in the selection of the variables. The scale is recommended as a method of improving the uniformity and reliability of the assessment of post-traumatic osteoarthritic changes after knee ligament injury.*

Résumé. *Nous avons mis au point un système de cotation afin d'évaluer les modifications arthrosiques après traumatisme d'un ligament du genou. Il est basé sur les radiographies du genou faites au moment de l'accident et des deux genoux au moment de l'examen de contrôle. On a eu recours à dix critères pour établir l'échelle de cotation. La méthode a été utilisée chez 60 sujets opérés d'emblée d'une lésion récente d'un ligament du genou. L'âge moyen des patients au moment de l'accident était de 26,6 ans et ils ont été revus en moyenne trois ans et demi après l'opération. L'échelle comporte des ajustements de sorte que la cotation définitive mesure seulement les modifications pathologiques liées à la lésion ligamentaire. La reproductibilité de cette échelle est excellente et les cotations sont parfaitement corrélées avec l'évaluation clinique. La fiche est facile à remplir, et peut être mise à jour si nécessaire. Les données numériques se prêtent aisément à l'analyse statistique par ordinateur. 95% de toutes les modifications causées par les traumatismes ligamentaires s'ajustent à la cotation, ce qui démontre la qualité du choix des critères. Ce système de cotation peut être recommandé comme un moyen d'améliorer l'objectivité et la fiabilité de l'estimation des manifestations arthrosiques après une lésion traumatique d'un ligament du genou.*

Key words: *Injury, Knee ligaments, Post-traumatic osteoarthritis, Reproducibility, Scoring scale, Validity*

Offprint requests to: P. Kannus

Introduction

The lack of comparable rating systems to measure the results after operative or conservative treatment of knee ligament injuries makes assessment of papers on the subject difficult. Noyes et al. [21] note that knee rating systems have little uniformity, employ different objective signs and various subjective symptoms. Kettelkamp and Thompson

[16] have proposed seven basic criteria for designing an assessment scale:

(1) It must be based on important measurable characteristics of the knee.

(2) Arbitrary assignment of point values should be avoided as far as possible.

(3) The variables listed must be easily quantified by a physician without requiring complicated instruments.

(4) The registration form must be simple and allow rapid completion.

(5) The variables included in the form should be suitably arranged for statistical and computer analyses.

(6) The total points derived from the scoring scale should be related to the clinical results.

(7) The sale must permit alteration in content when necessary.

These requirements were proposed for the evaluation of the results after knee osteotomies and arthroplasties. We have included three more for the special scale for the evaluation of post-traumatic osteoarthritis after knee ligament injuries.

(8) The scale must show reproducibility, i.e. the variation between different examiners should be minimal.

(9) It must concentrate on the pathological changes seen typically after knee ligament injuries.

(10) It should only taken account of the changes which are caused by that particular injury.

Numerous clinical knee scoring scales and modifications have been developed and used [3, 6, 7, 11, 16–19, 21–23]. Evaluation of the long term results of knee ligament injuries should include both clinical examination and radiographic assessment in order to reveal post-traumatic changes in the injured knee. No such system is currently available. The three common methods [2, 5, 15] for recording radiographic findings are too crude and not suitable for use following knee ligament injury.

We present a new radiological scoring scale which meets the criteria mentioned above.

Method

Anteroposterior films with the affected knee in extension and lateral views with the knee in 30 degrees of flexion were taken in the supine position at the time of initial presentation of the injury. At follow-up, films of both knees were obtained in the same positions.

Five variables are common to each of the three most used classifications namely osteophyte formation, subchondral sclerosis, flattening of the femoral condyles, narrowing of the joint spaces, and subchondral cysts. We included two more particularly suitable for the special scale of knee ligament injuries, angular deformity and ligament calcification (Table 1).

Calculation of the points scores

Osteophytes. Osteophyte formation was recorded by the method described by Hernborg and Nilsson [8]. Osteophytes were measured by mean of a caliper, and the size was defined as the largest perpendicular distance from the edge of the cortex to the outer margin of the osteophyte. Osteophyte formation was measured at seven sites in the knee joint (Table 1) and each graded from 3 to 0 according to size; 3 = no osteophyte formation, 2 = small (1–3 mm) osteophytes, 1 = moderate (4–6 mm), and 0 = large (>6 mm).

The maximum point value of the osteophyte variable was $7 \times 3 = 21$.

Subchondral sclerosis. Subchondral sclerosis was defined as an area with an increased density of cancellous bone adjacent to an articular surface. The amount of sclerosis was measured at five sites (Table 1), and rated on the scale 3 = no subchondral sclerosis, 2 = mild sclerosis (area involved < one-third), 1 = moderate (area involved between one-third and two-thirds), and 0 = severe (area involved > two-thirds). The maximum point value of the subchondral sclerosis variable was $5 \times 3 = 15$.

Flattening of femoral condyles. The amount of flattening of the femoral condyles was measured separately on the medial and lateral sides and rated on the scale 3 = no flattening of femoral condyle, 2 = mild flattening, 1 = moderate, 0 = severe. The maximum point value of the flattening variable was $2 \times 3 = 6$.

Subchondral cysts. The subchondral cysts were measured as general cyst formation and rated 6 = no subchondral cyst formation, 4 = small amount of cyst formation (random cysts just visible), 2 = moderate (clearly visible cyst in most of the subchondral area), and 0 = severe (large cysts in every subchondral area). The maximum value of the subchondral cysts variable was 6.

Ligament calcification. The amount of ligament calcification was measured from six sites in the knee (Table 1) and rated 3 = no ligament calcification, 2 = mild calcification (just visible), 1 = moderate (clearly visible), and 0 = severe (large calcified area around the ligament). The maximum value of the ligament calcification variable was $6 \times 3 = 18$.

Narrowing of the joint spaces. The maximum breadth of the medial and lateral tibiofemoral joint spaces was measured in all films and was marked on the scale in millimetres. The measurement was done on all films of the same patient at the same anatomical point of the joint space. The point score was the measurement in millimetres.

Angular deformity. In all AP films the knee angle was measured in degrees between the long axis of the tibia and the femur. The point score for angular deformity was the degree measurement.

Calculation of the final scoring points

In order that the scale should take into account only the changes caused by the injury, the primary changes seen in the

Table 1. Registration form of radiological knee scoring scale

Variable	Radiographs			
	Primary pretreatment score	Primary follow-up score		Final score
	Injured knee	Injured	Uninjured knee	
Osteophytes				
med. femoral condyle	____	____	____	____
med. tibial condyle	____	____	____	____
lat. femoral condyle	____	____	____	____
lat. tibial condyle	____	____	____	____
med. tibial eminence	____	____	____	____
lat. tibial eminence	____	____	____	____
patella	____	____	____	____
Subchondral sclerosis				
med. femoral condyle	____	____	____	____
med. tibial condyle	____	____	____	____
lat. femoral condyle	____	____	____	____
lat. tibial condyle	____	____	____	____
patella	____	____	____	____
Flattening of femoral condyles				
medial	____	____	____	____
lateral	____	____	____	____
Subchondral cysts	____	____	____	____
Ligament calcification				
med. femoral condyle	____	____	____	____
med. tibial condyle	____	____	____	____
lat. femoral condyle	____	____	____	____
lat. tibial condyle	____	____	____	____
tibial eminence	____	____	____	____
fibula	____	____	____	____
Narrowing of joint spaces				
medial tibiofemoral	____	____	____	____
lateral tibiofemoral	____	____	____	____
Angular deformation				
valgus	____	____	____	____
varus	____	____	____	____
Total final score				____

injured knee and the follow-up changes in the uninjured knee had to be excluded.

Firstly the primary points for the injured knee before treatment were compared with the primary follow-up points of the same knee for every variable and every anatomical component of the knee (see Table 1). Point reduction indicated development or advancement of arthritic change. If the point score had decreased, the score at follow-up in both knees was compared, and every variable where point reduction had occurred in the injured knee was assessed for both. If the points scores were identical, or the uninjured side had a lower score, no development or advancement of the pathological process differing from the normal background process had occurred in the injured knee during the follow-up period. If the score for the injured knee was lower, increased degenerative change had occurred.

Narrowing of the joint spaces. The classification described by Johnson et al. [12, 13] was used, in which 12 = normal joint space with no obliteration, 6 = narrowed by 50% or less, 3 = narrowing by more than 50% and 0 = obliterated. The normal joint space was taken to be the breadth of the pretreatment level, normally over 4 mm in healthy adults [1].

Comparison was first made to see if there was any narrowing of the joint spaces between the primary and the follow-up films of the injured knee. If there was no change the patient got the maximum 12 points. If narrowing was present the follow-up films of both knees were also compared. If the appearance of the two knees was identical no narrowing had occurred secondary to the injury, and the maximum 12 points was allotted, but if narrowing had occurred the score was calculated by subtracting the score in the injured knee from the uninjured, thus measuring the decrease in joint space due to injury.

Angular deformity. The normal angulation of the knee was defined as between 4 and 9 degrees of valgus [20]. It was called valgus deformity if the angle was more than 9 degrees of valgus, mild varus deformity if the angle was 0–3 degres of valgus, and severe varus deformity if the knee was in the varus position. The angular deformity was rated on the scoring scale in three classes [2, 9, 10, 12]: I = 10 (normal angulation), II =

5 (valgus or mild varus deformity), and III = 0 (severe varus deformity).

Any change in the knee angle between the primary and the follow-up films of the injured knee was measured and assessed as to whether the change was substantial enough to move the patient to a lower class. If the class did not change or had moved towards a higher level, the patient got the maximum of 10 whatever his original class, since no prognostically important change had occurred because of the injury. If a change in the knee angle of the injured knee had occurred which would qualify for a lower class, the uninjured knee was similarly assessed, and any angular change in this knee taken into account in classifying the injured side.

Final score

The final score is the sum of the scores of the seven variables (Table 1). If the injured knee has no sign of post-traumatic changes of osteoarthritis caused by the ligament injury, it gets the maximum 100, which is classified as excellent; 95–99 is classified as good (mild osteoarthritic changes because of the injury), 90–94 as fair (moderate changes), and under 90 as poor (severe changes).

Patient assessment

The knee scoring scale was applied to 60 patients operated on primarily after an acute knee ligament injury. There were 30 men and 30 women, an the mean age at the time of operation was 36.6 years, ranging from 8 to 63. None had previous knee injuries or operations. Primary films of the injured knee were taken at the time of the injury, and a radiological re-examination was performed on an average of 3.5 years (2–7 years) after operation.

50% cent of the patients were operated on in the first two days, and 98% within two weeks of injury. Rupture of the anterior cruciate and medial collateral ligaments was found in 36% of the patients, an isolated medial collateral ligament injury in 31%, rupture of the posterior cruciate ligament alone or in combination with other ligaments in 18%, and other isolated or combined ligament injuries in 15%. The patients with other simultaneous injuries such as knee fractures, were not included in the study.

Validity of the scale

In order to check the validity of the scale subjective, clinical, functional and a combined assessments were made at the time of the radiological review. Each was measured separately against the radiological scale in order to find out if the radiological scores would differ from each other in different healing groups.

Subjective evaluation. The subjective evaluation was by questionnaire. The knee was graded as excellent, good, fair or poor (Table 2).

Clinical evaluation. Both knees were examined as described previously [14], with special emphasis on the stability of the knee joint. The patients were rated as excellent if no more than one pathological change was found when compared to the uninjured knee, good if 2 to 3, fair if 4 to 5, and poor if 6 or more (Table 2).

Functional evaluation. The functional evaluation was based on four different questions, whether the patients had difficulties walking, running or climbing stairs, and whether their physical activity had diminished because of the injury. The functional recovery was rated excellent if the answer was no to all questions, good if there was only one affirmative answer, fair if two or three, and poor if all four activities were impaired (Table 2).

Table 2. Distribution of test population into different healing groups

	Excellent	Good	Fair	Poor	Total
Subjective evaluation	8	19	28	5	60
Clinical evaluation	12	20	22	6	60
Functional evaluation	13	16	22	9	60
Combined evaluation	11	20	24	5	60
Radiological evaluation	13	21	20	6	60

Combined evaluation. A median of the subjective, objective and functional assessments was made (Table 2).

Reproduceability of the scale. All films were analyzed first by two of us (PK and MJ) and then independently by the third (TP). The coefficients of correlation and variation between the examinations in respect of the total scores were calculated. The total scores were also divided into four groups. The test coefficient of reliability (kappa) and the percentage of patients assigned to the same healing groups between the two examinations were also calculated. The distribution of the patients according to this radiological grouping is presented in Table 2.

Statistics

Knee scoring systems are basically ordinal scales but they are widely used as interval scales [11, 16, 18, 19, 21, 22], and thus in this study both nonparametric and parametric statistics were used. All calculations between the differences of averages were done by Student's t-test and between the differences of medians by Mann-Whitney's *U*-test. Between the scores of two repeated examinations the significance of the Pearson correlation coefficient and the Spearman rank correlation coefficient was calculated. The given significance levels refer to two-tailed tests: * means $P < 0.05$, ** means $P < 0.01$, and *** means $P < 0.001$. All data were stored in the Dec 2060 computer of Tampere University using in analyses the program library of BMDP-82 [4].

Results

The mean radiological score of the knees tested was 95.1 ± 5.2 (median 96.5) ranging from 73 to 100. The six most characteristic radiographic changes after the knee ligament injuries were osteophytes in the patella, in the spine of the tibia, and in the medial condyles of the femur and the tibia, narrowing of the medial joint space, and ligament calcification at the medial femoral condyle

(Fig. 1). All these changes were included in the variables of the scoring scale; 95% of all changes caused by the injuries were identified and registered on the scale.

The correlation between the score values, means and medians, and the subjective evaluation is presented in Fig. 2. The only statistically significant step was from good healing to fair ($P < 0.05$). However, the descending trend of the scores was clear from excellent towards poor healing.

The correlation between the scores and the clinical evaluation is presented in Fig. 3. Statistically significant differences were seen between the steps from excellent healing to good and fair healing to poor. In the mean scores of the subgroups good and fair were almost the same (95.2 and 95.4); between the medians there was some, but no significant difference (96.5 and 95.5).

The correlation between the scores and the functional evaluation is presented in Fig. 4. Between each step in the functional healing groups the means and medians of the scores differed significantly from each other.

The correlation between the scores and the combined evaluation is presented in Fig. 5. As in the functional evaluation there was a significant difference between each step in the healing groups.

Between the two independent evaluations of the x-ray films the Pearson correlation coefficient was 0.94 ($P < 0.001$), the Spearman rank correlation coefficient 0.90 ($P < 0.001$), and the coefficient of variation 5.6% (Fig. 6). The test coefficient of reliability (kappa) was 0.70 and the percentage of patients classified as belonging to the same radiological healing groups between the two evaluations was 78% (Table 3).

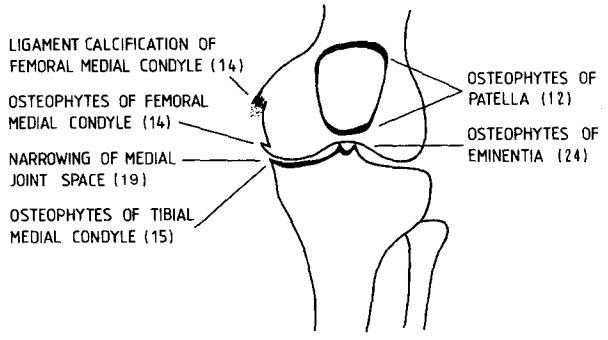

Fig. 1. Frequency of most common radiological changes caused by knee ligament injuries

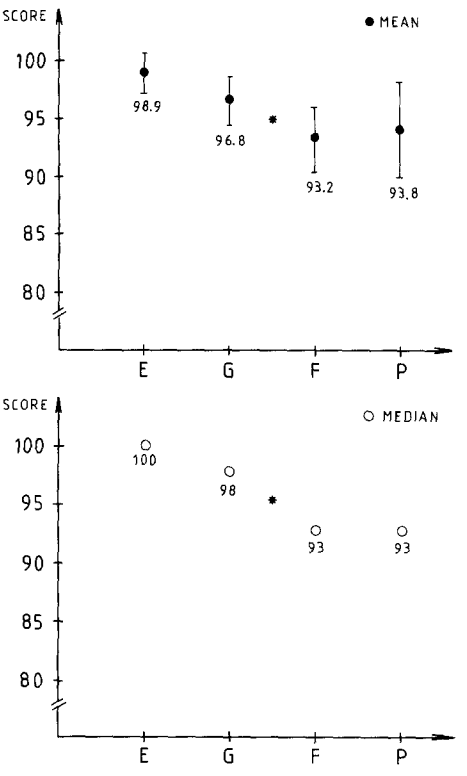

Fig. 2. Correlation between radiological scores and subjective evaluation. ● Mean ± SD; ○ median; *E* excellent; *G* good; *F* fair; *P* poor

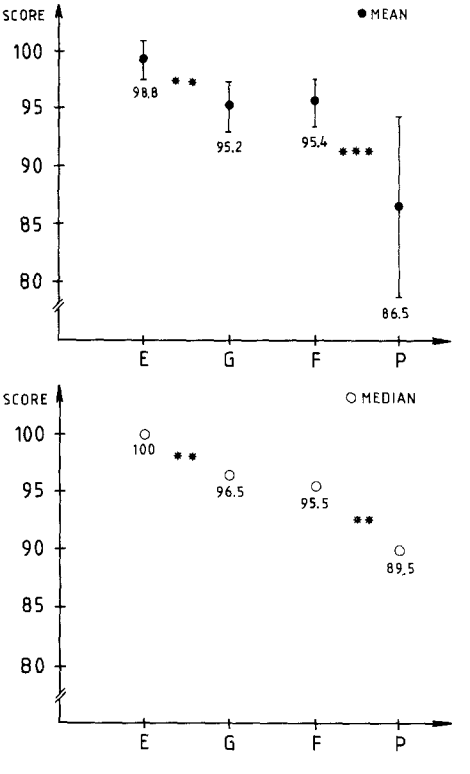

Fig. 3. Correlation between radiological scores and clinical evaluation. ● Mean ± SD; ○ median; *E* excellent; *G* good; *F* fair; *P* poor

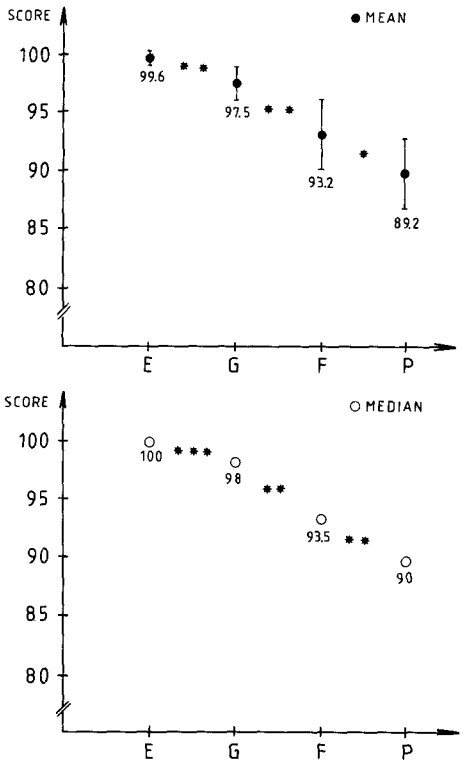

Fig. 4. Correlation between radiological scores and functional evaluation. ● Mean ± SD; ○ median; *E* excellent; *G* good; *F* fair; *P* poor

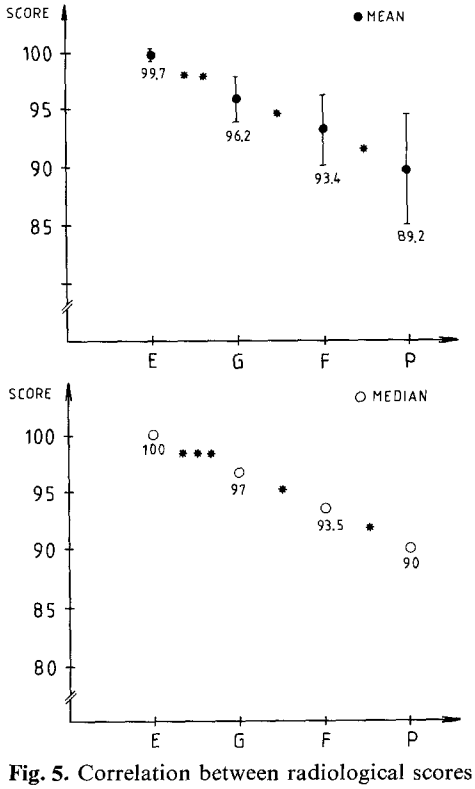

Fig. 5. Correlation between radiological scores and combined evaluation. ● Mean ± SD; ○ median; *E* excellent; *G* good; *F* fair; *P* poor

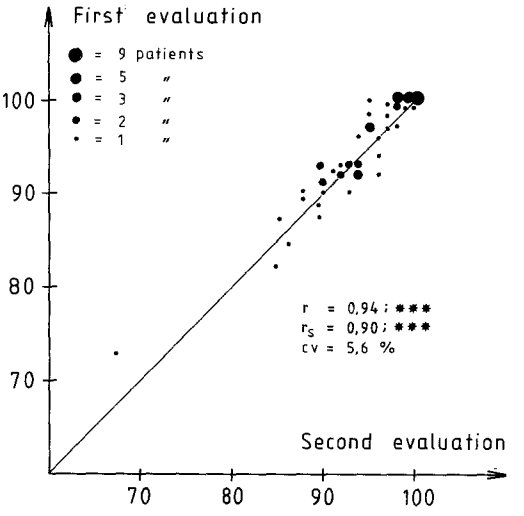

Fig. 6. Correlation of radiological scores between two independent (interpersonal) evaluations. *r* Pearson correlation coefficient; *rs* Spearman rank correlation coefficient; *cv* coefficient of variation

Discussion

A radiological scoring scale is a new approach to the evaluation of post-traumatic osteoarthritic changes in the knee joint after ligamentous injury. The scale met the criteria listed in the introduction.

(1) The selection of variables was based on the recognised methods of assessment of knee radiographs [2, 5, 15].

(2) The assignment of point values to each variable and between variables was made according to their frequency and importance, adjusting the maximum to 100. The three most important variables, narrowing of the joint spaces, osteophyte formation and ligament calcification, were the most usual findings in the knees assessed.

Table 3. Distribution of test population into different radiological healing groups between two evaluations

Second evaluation	First evaluation				
	Excellent	Good	Fair	Poor	Total
Excellent	8	2	0	0	10
Good	5	18	2	0	25
Fair	0	1	15	0	16
Poor	0	0	3	6	9
Total	13	21	20	6	60

Test coefficient of reliability (kappa) = 0.70
Percent of same gradation 47/60 = 78%

(3) The selected radiographic variables could be quantified by a physician in his normally equipped office.

(4) The registration form with seven variables was simple enough to be completed rapidly.

(5) All data were numerical and thus immediately prepared for statistical and computer analyses, which were easily accomplished.

(6) The radiological scores correlated excellently with the functional and combined assessments, but to a lesser degree with the subjective and clinical results. However, the trend of the score values was clearly descending from excellent to poor in the subjective and clinical evaluation.

(7) The scale was such that it permitted alteration in content when necessary.

(8) The interpersonal reproducibility of the scale was excellent, probably because the variables in the scale were clear and easy to evaluate, and only the knee joint was assessed. Kellgren and Lawrence (15), when making their radiological assessment of osteoarthrosis of human joints, showed that in the knee joint there was the best correlation between two observers.

(9) Ninety five per cent of all the changes caused by the ligament injuries were recorded by the scale, which therefore concentrated well on the pathological changes seen typically after such injuries.

(10) The method used to calculate the scale enabled assessment to be strictly confined to changes related only to the injury.

We consider that the use of this scale together with subjective, functional [18], clinical [19] and other rating systems will allow a more reliable assessment of overall results.

References

1. Adams ID (1976) Osteoarthrosis and sport. Clin Rheum Dis 2: 523–541
2. Ahlbäck S (1968) Osteoarthrosis of the knee. A radiographic investigation. Acta Radiol Suppl 277: 1–72
3. Aichroth P, Freeman MAR, Smillie JS, Souter WA (1978) A knee function assessment chart. J Bone Jt Surg 60-B: 308–309
4. BMDP statistical software (1981) Dixon WJ (ed) University of California Press, Berkeley
5. Fairbank TJ (1948) Knee joint changes after meniscectomy. J Bone Jt Surg 30-B: 664–670
6. Feagin JA Jr, Blake WP (1983) Postoperative evaluation and results recording in the anterior cruciate ligament reconstructed knee. Clin Ortop 172: 143–147
7. Freeman MA, Todd RC, Cundy AD (1977) Technique for recording the results of knee surgery. Clin Ortop 128: 216–221
8. Hernborg J, Nilsson BE (1973) The relationship between osteophytes in the knee joint, osteoarthritis and aging. Acta Orthop Scand 44: 69–74
9. Hernborg JS, Nilsson BE (1977) The natural course of untreated osteoarthritis of the knee. Clin Ortop 123: 130–137
10. Iseki F, Fujikawa K (1980) Clinical pictures of the osteoarthritis in the knee joint. J Jpn Orthop Ass 54: 563–574
11. Jensen JE, Slocum DB, Larson RL, James SL, Singer KM (1983) Reconstruction procedures for anterior cruciate ligament insufficiency: a computer analysis of clinical results. Am J Sports Med 11: 240–248
12. Johnson RJ, Kettelkamp DB, Clark W, Leaverton P (1974) Factors affecting late results after meniscectomy. J Bone Jt Surg 56-A: 719–729
13. Johnson RJ, Eriksson E, Häggmark T, Pope H (1984) Five to ten-year follow-up evaluation after reconstruction of the anterior cruciate ligament. Clin Ortop 183: 122–140
14. Järvinen M, Kannus P (1985) Clinical and radiological long-term results after primary knee ligament surgery. Arch Orthop Trauma Surg 104: 1–6
15. Kellgren JH, Lawrence JS (1957) Radiological assessment of osteo-arthrosis. Ann Rheum Dis 16: 494–502
16. Kettelkamp DB, Thompson C (1975) Development of a knee scoring scale. Clin Ortop 107: 93–99
17. Larson R (1974) Rating sheet for knee function. In: Smillie I (ed) Diseases of the knee joint. Churchill Livingstone, Edinburgh, pp 29–30
18. Lysholm J, Gillquist J (1982) Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am J Sports Med 10: 150–154
19. Marshall JL, Fetto JF, Botero PM (1977) Knee ligament injuries. A standardized evaluation method. Clin Ortop 123: 115–129
20. McDaniel WC, Dameron TB (1983) The untreated anterior cruciate ligament rupture. Clin Ortop 172: 158–163
21. Noyes FR, McGinniss GH, Grood E (1985) The variable functional disability of the anterior cruciate ligament-deficient knee. Orthop Clin North Am 16: 47–67
22. Tegner Y, Lysholm J (1985) Rating systems in the evaluation of knee ligament injuries. Clin Ortop 198: 43–49
23. Turba JE, Walsh MW, McLeod WD (1979) Long-term result of extensor mechanism reconstruction. A standard for evaluation. Am J Sports Med 7: 91–94

Note added in proof

Case examples on how to use the radiological knee scoring scale are available from Dr. P. Kannus.