# Rating Systems in the Evaluation of Knee Ligament Injuries

YELVERTON TEGNER, M.D., AND JACK LYSHOLM, M.D., PH.D.

**Many different methods of evaluating disability after knee ligament injury exist. Most of them differ in design. Some are based on only patients' symptoms. Others include patients' symptoms, activity grading, performance in a test, and clinical findings. The rating in these evaluating systems can be either numerical, as in a score, or binary, with yes/no answers. Comparison between a symptom-related score and a score of more complex design showed that the symptom-related score gave a more differentiated picture of the disability. It was also shown that the binary rating system gave less detailed information than a score and that differences in a binary rating can depend on at what level the symptoms are regarded as "significant." A new activity grading scale, where work and sport activities were graded numerically, was constructed as complement to the functional score. When evaluating knee ligament injuries, stability testing, functional knee score, performance test, and activity grading are all important. However, the relative importance varies during the course of treatment, and therefore they should not all be included in one and the same score.**

During recent decades the use of different scoring scales for follow-up study of treatment of knee problems has become popular.[2–7,10,14] Their design varies, but most are based on ratings of both function and clinical findings.

When constructing an evaluation system, two major questions arise. The first concerns whether the functional rating should be based on only the patients' symptoms during different activities or whether a rating of clinical findings should be added. The second question is whether the rating should be numerical, as in a score, or binary, with yes/no answers.

The object of this study is to analyze differences between different types of scores. A new system of activity grading is presented.

## MATERIAL

Seventy-six patients, 55 men and 21 women (mean age, 27 years), were included. All had anterior cruciate ligament injury diagnosed by clinical examination under anesthesia and by arthroscopy or arthrotomy.

## METHODS

All patients were assessed with a rating scale (Score I, Table 1) that takes up symptoms during daily activities—a modification of the score presented by Lysholm and Gillquist[7] in 1982. Score I is a discrete rating scale in which the patient can achieve a maximum score of 100. Forty-seven of the patients were also assessed with a rating scale (Score II, Table 2) presented by Marshall *et al.*[10] in 1977, which covers the symptoms, activity grading, results of a simple functional test, and clinical findings; many of the items are graded in a binary way, *i.e.*, the symptoms are evaluated in an all-or-none fashion, and the maximum score is 50 points.

From the Sports and Trauma Research Group, Department of Orthopaedic Surgery, University Hospital, Linköping, Sweden.

This study was supported by grants from Östergötlands Läns Landsting, the University of Linköping, the Research Council of the Swedish Sports Association, and the Tore Nilsson foundation.

Reprint requests to Yelverton Tegner, M.D., Department of Orthopaedic Surgery, University Hospital, S-581 85 Linköping, Sweden.

Received: June 8, 1984.

TABLE 1. Lysholm Knee Scoring Scale

| | |
|---|---|
| Limp (5 points) | |
| None | 5 |
| Slight or periodical | 3 |
| Severe and constant | 0 |
| Support (5 points) | |
| None | 5 |
| Stick or crutch | 2 |
| Weight-bearing impossible | 0 |
| Locking (15 points) | |
| No locking and no catching sensations | 15 |
| Catching sensation but no locking | 10 |
| Locking | |
| Occasionally | 6 |
| Frequently | 2 |
| Locked joint on examination | 0 |
| Instability (25 points) | |
| Never giving way | 25 |
| Rarely during athletics or other severe exertion | 20 |
| Frequently during athletics or other severe exertion (or incapable of participation) | 15 |
| Occasionally in daily activities | 10 |
| Often in daily activities | 5 |
| Every step | 0 |
| Pain (25 points) | |
| None | 25 |
| Inconstant and slight during severe exertion | 20 |
| Marked during severe exertion | 15 |
| Marked on or after walking more than 2 km | 10 |
| Marked on or after walking less than 2 km | 5 |
| Constant | 0 |
| Swelling (10 points) | |
| None | 10 |
| On severe exertion | 6 |
| On ordinary exertion | 2 |
| Constant | 0 |
| Stair-climbing (10 points) | |
| No problems | 10 |
| Slightly impaired | 6 |
| One step at a time | 2 |
| Impossible | 0 |
| Squatting (5 points) | |
| No problems | 5 |
| Slightly impaired | 4 |
| Not beyond 90° | 2 |
| Impossible | 0 |

Comparison between a binary system and a score system was done by converting Score I values for different items (pain, swelling, instability) to a binary system. The results were analyzed with a Venn diagram,[15] which is a way of graphically visualizing two or more answers in a nomial system.

Forty-three patients with anterior cruciate injury filled in a questionnaire in which they graded certain activities according to how troublesome they were to perform. Based on this, a new numerical activity grading scale was constructed as a complement to the functional score (Table 3). The activity levels for all 76 patients were determined and analyzed in relation to Score I.

The reproducibility of Score I was determined. The intrapersonal coefficient of variation was estimated by letting the same orthopedic surgeon determine the score for 15 patients twice with an interval of two weeks. To establish the interpersonal variation, an orthopedic surgeon and a physiotherapist determined the score for the same 15 patients on one and the same occasion.

## STATISTICAL ANALYSIS

Student's *t*-test, the chi-square test, and the Pearson correlation coefficient were used. Significance levels refer to two-tail tests.

## RESULTS

There was significant correlation between Scores I and II ($r = 0.78$; $p < .001$; Fig. 1). With Score I, however, a greater proportion of patients had excellent/good knee function than with Score II (Fig. 1).

Score I and a binary (yes/no) rating of answers concerning "instability" are compared in Figure 2. All patients with a total score of less than 65 points complained of instability during sports, and almost all during daily activities. Most patients (81%) achieving 65 to 83 points had problems during sports, but only one-third during daily activities. Of patients achieving 84–90 points, only 8% had difficulties during daily activities, but 70% during sports. Similar figures were obtained for pain and swelling.

Venn diagrams analyzing instability, pain, and swelling are shown in Figures 3 and 4. If "symptom" is defined as a problem arising during strenuous activities (Fig. 3), only seven individuals will be listed as having no problems. Conversely, if the limit for problems is

TABLE 2. Marshall Scoring Scale

| | |
|---|---|
| Pain | 0 = Yes 1 = No |
| Swelling | 0 = Yes 1 = No |
| Stair difficulty | 0 = Yes 1 = No |
| Clicking/numbness | 0 = Yes 1 = No |
| Giving way | 0 = Regularly upon daily activities<br>1 = With stress upon daily activities<br>2 = With stress only<br>4 = Normal, none |
| Return to sports/work | 0 = No return<br>1 = Return to different<br>2 = Return to original with limitations<br>3 = Full return |
| Functional tests | |
| Duck walk | 0 = Cannot perform<br>1 = Can perform but with discomfort<br>2 = Can perform |
| Run in place | 0 = Cannot<br>1 = Can |
| Jump on one leg | 0 = Cannot perform<br>1 = Can perform but with discomfort<br>2 = Can perform |
| Half squat | 0 = Cannot<br>1 = Can |
| Full squat | 0 = Cannot<br>1 = Can |
| Specific knee examinations | |
| Tenderness | 0 = Yes 1 = No |
| Joint effusion | 0 = Yes 1 = No |
| Swelling (soft tissue) | 0 = Yes 1 = No |
| Crepitations | 0 = Yes 1 = No |
| Muscle power | 0 = Very weak<br>1 = Diminished flexion and extension<br>2 = Diminished flexion or extension<br>3 = Normal |
| Thigh sizes | 0 = >2 cm difference<br>1 = 1–2 cm difference<br>2 = Equal |
| Range of motion | 0 = <90°<br>1 = Limited flexion and extension<br>2 = Limited flexion or extension<br>3 = Normal |
| Stability | |
| LCL | 0 = Gross instability<br>2 = Instability in flexion and extension<br>3 = Moderate instability in flexion<br>4 = Mild instability in flexion<br>5 = Normal |
| MCL | 0 = Gross instability<br>2 = Instability in flexion and extension<br>3 = Moderate instability in flexion<br>4 = Mild instability in flexion<br>5 = Normal |
| ACL | 0 = Severe in neutral and rotation (Pivot shift, Slocum, Jerk test)<br>2 = Severe in neutral<br>3 = Moderate jog<br>4 = Slight jog<br>5 = Normal |
| PCL | 0 = Severe in neutral and rotation<br>2 = Severe in neutral<br>3 = Moderate jog<br>4 = Slight jog<br>5 = Normal |

set at daily activities (Fig. 4), more individuals will be regarded as symptom-free ($\chi^2 = 37.95$, $p < .001$).

The activity scale is graded from 0 to 10 (Table 3) and covers activities in daily life and recreational and competitive sports. Activity levels 5–10 can be achieved only if the patient takes part in recreational or competitive sports. Significant differences in scores at different activity levels were obtained (Fig. 5). The mean score for patients at activity levels 5–10 was 83 ± 10, and that for patients at activity level 0 was 53 ± 16 ($p < .001$). Seventeen percent of patients in activity levels 0–3 had a score above 83.

The intrapersonal coefficient of variation[1]

TABLE 3. Activity Score

10. Competitive sports
    Soccer—national and international elite
9. Competitive sports
    Soccer, lower divisions
    Ice hockey
    Wrestling
    Gymnastics
8. Competitive sports
    Bandy
    Squash or badminton
    Athletics (jumping, etc.)
    Downhill skiing
7. Competitive sports
    Tennis
    Athletics (running)
    Motorcross, speedway
    Handball
    Basketball
  Recreational sports
    Soccer
    Bandy and ice hockey
    Squash
    Athletics (jumping)
    Cross-country track findings both recreational and competitive
6. Recreational sports
    Tennis and badminton
    Handball
    Basketball
    Downhill skiing
    Jogging, at least five times per week
5. Work
    Heavy labor (*e.g.*, building, forestry)
  Competitive sports
    Cycling
    Cross-country skiing
  Recreational sports
    Jogging on uneven ground at least twice weekly
4. Work
    Moderately heavy labor (*e.g.*, truck driving, heavy domestic work)
  Recreational sports
    Cycling
    Cross-country skiing
    Jogging on even ground at least twice weekly
3. Work
    Light labor (*e.g.*, nursing)
  Competitive and recreational sports
    Swimming
  Walking in forest possible
2. Work
    Light labor
  Walking on uneven ground possible but impossible to walk in forest
1. Work
    Sedentary work
  Walking on even ground possible
0. Sick leave or disability pension because of knee problems

was 3% and the interpersonal 4%. The test–retest correlation coefficients were 0.97 and 0.90, respectively.

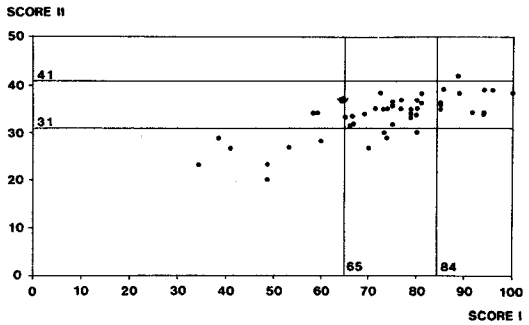


FIG. 1. Relation between Scores I and II. The lines represent the limits between excellent/good and fair (84 resp. 41) and between fair and poor (65 resp. 31). Pearson correlation coefficient = 0.78.

## DISCUSSION

Scores I and II represent two different approaches to the same problem. There is a correlation between the two, indicating that both measure the same thing, namely, knee function. In spite of this correlation, there are differences between the two scales. Patients achieving a high score with Score I tend to achieve too-low values with Score II (Fig. 1), and patients with low Score I values tend to be overestimated with Score II. It seems reasonable to explain these differences by differences in design. A possible explanation would be that Score II includes tests of stability and function. When constructing Score II, Marshall claimed that stability is highly important for successful rehabilitation and return of function.[10] While this is un-

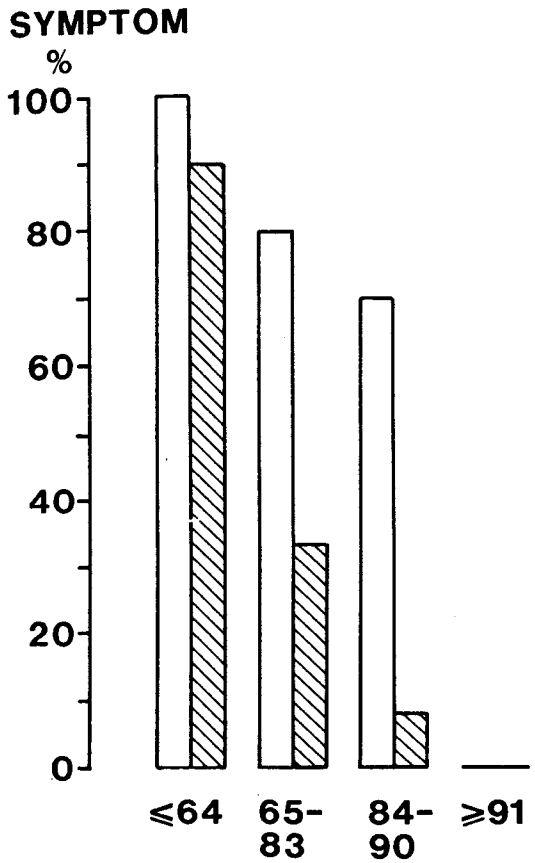


FIG. 2. Proportion with and without symptoms of instability in 76 patients divided into four groups based on their results with Score I. In the group with a score of 91 points or more, none had symptoms of instability.

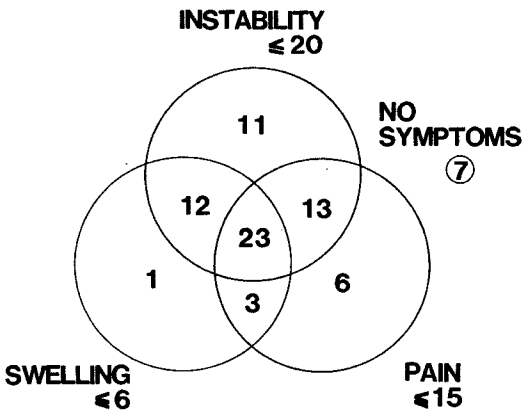


FIG. 3. Venn diagram showing the distribution of patients with problems during *strenuous activities* in the circles.

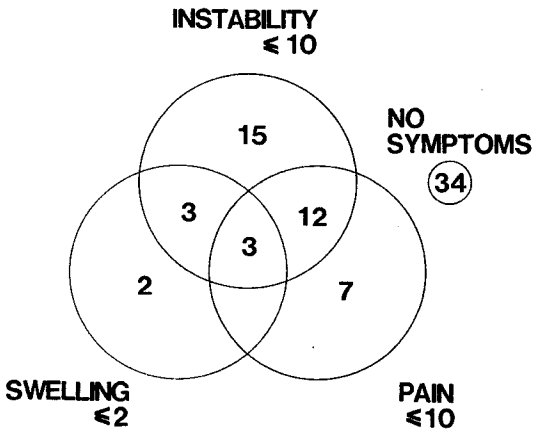


FIG. 4. Venn diagram showing the distribution of patients with problems during *daily activities* in the circles.

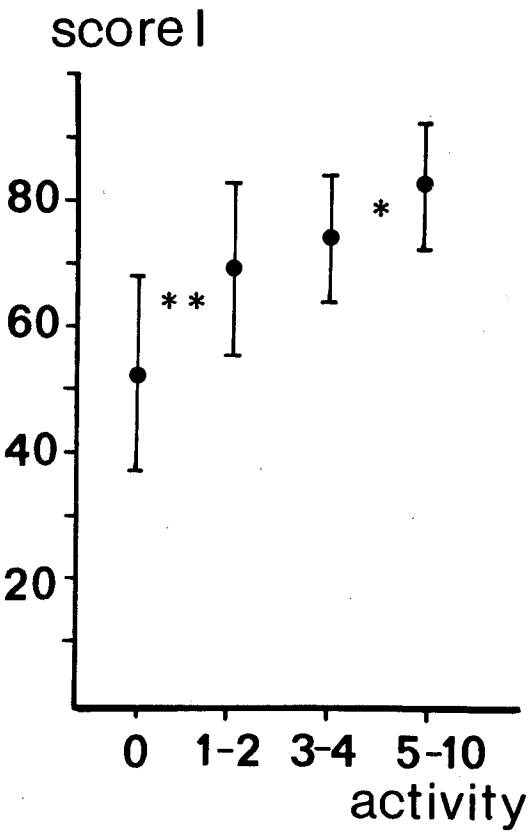


FIG. 5. Relation between Score I and activity level. Significant differences in score existed between patients at activity level 0 and activity levels 1–2 ($p < .01$) and between patients with activity levels 3–4 and 5–10 ($p < .05$).

deniable, other factors such as muscular strength and meniscus status also influence the outcome.[11] The authors have shown earlier[12] that after a period of strength training only, many patients can regain acceptable knee function in daily life and sports activities; similar results have been obtained by Noyes *et al.*[11]

In the Marshall system, the same number of points are deducted for instability owing to injuries to the lateral and medial collateral ligaments and to the anterior and posterior cruciate ligaments. Lysholm *et al.*[8] showed that different instabilities result in different degrees of reduction in the knee score; in other words, different ligament injuries influence function in different ways. Clinical findings should therefore not be included in a rating scale for knee function but should be recorded separately.

Score I has been compared earlier to a score of similar design, the Larson score.[6] It was shown that it gave a better picture of the disability in the patients with a knee instability than the Larson score.

A knee performance test[13] is an objective way of measuring performance (running time, hop length, etc.) in controlled sportslike activities, whereas a score can be used to assess symptoms in various activities of sports and daily life. The value of the knee function test lies in monitoring rehabilitation.[9,13] A score is less useful for this purpose because rehabilitation commonly involves restrictions of activity. However, in a situation with no such restriction, the score evaluates knee function "seven days a week," but the performance test evaluates knee function only in a short test situation. Therefore, with no restriction of activity, the score probably gives the better picture of knee function. With this in mind, the authors prefer, contrary to Marshall, not to include a function test in the knee score.

In a binary (yes/no) system, differences in results can be obtained depending on the activity level at which a symptom is regarded as "significant." If a symptom is defined as such occurring only in daily life, many patients will be regarded as symptom-free even though they have considerable problems on strenuous activity. A binary rating thus depends on the activity level at which symptoms are regarded significant. Conversely, with a scoring scale, the more often symptoms arise and the lower the load causing them, the lower will be the score. A more differentiated picture of the disability can thus be obtained with a scoring scale.

However, if the limits are set correctly, a Venn diagram of binary ratings is useful because it shows the combination of two or more symptoms.

Terms such as "return to sports" are often used in the evaluation of different treatments of knee ligament injuries. Because different sports and activities put different strains on the knee, such terms lack meaning. It is better to grade different activities in a standardized way on a numerical scale. The preinjury, present, and desired activity levels are readily defined with such a scale. Almost 20% of patients with an activity level between 0 and 3 had a high score (>83), which indicates that limitations in knee function may be masked by an involuntarily low activity level. The activity scale is thus a valuable complement to the functional score.

A functional score, activity grading, test of function, and static stability grading are all important in the evaluation of knee ligament injuries before, during, and after treatment and rehabilitation. However, the relative importance of each part of the evaluation system can vary during the course of treatment and during the follow-up period.

ACKNOWLEDGMENTS

The authors wish to thank Mr. Eric Leander for statistical advice and Dr. Marcia Skogh for revising the English text.

REFERENCES

1. Armitage, P.: Statistical Methods in Medical Research. London, Blackwell Scientific Publications, 1983.

2. Arnold, J. A., Coker, T. P., Heaton, L. M., Park, J. P., and Harris, W. D.: Natural history of anterior cruciate tears. Am. J. Sports Med. 7:305, 1979.
3. Feagin, J. A., Jr., and Blake, W. P.: Postoperative evaluation and result recording in the anterior cruciate ligament reconstructed knee. Clin. Orthop. 172:143, 1983.
4. Geens S., Clayton, M. L., Leidholt, J. D., Smyth, C. J., and Bartholomew, B. A.: Synovectomy and débridement of the knee in rheumatoid arthritis. J. Bone Joint Surg. 51A:626, 1969.
5. Kettelkamp, D., and Thompson, C.: Development of a knee scoring scale. Clin. Orthop. 107:93, 1975.
6. Larson, R.: Rating sheet for knee function, 1972. *In* Smillie, I.: Disease of the Knee Joint. Edinburgh, Churchill Livingstone, 1974, p. 29.
7. Lysholm, J., and Gillquist, J.: Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. Am. J. Sports Med. 10:150, 1982.
8. Lysholm, J., Tegner, Y., and Gillquist, J.: Functional importance of different clinical findings in the unstable knee. Acta Orthop. Scand. 55:472, 1984.
9. Lysholm, J., Tegner, Y., Odensten, M., Nordin, M., and Gillquist, J.: Is the rehabilitation after knee ligament surgery effective enough? Berlin, European Society of Knee Surgery, 1984.
10. Marshall, J., Fetto, J., and Botero, P.: Knee ligament injuries: A standardized evaluation method. Clin. Orthop. 123:115, 1977.
11. Noyes, F. R., Matthews, D. S., Mooar, P. A., and Grood, E. S.: The symptomatic anterior cruciate-deficient knee. Part II. J. Bone Joint Surg. 65A:163, 1983.
12. Tegner, Y., Lysholm, J., Gillquist, J., and Öberg, B.: Two-year follow-up on conservative treatment of knee ligament injuries. Acta Orthop. Scand. 55: 176, 1984.
13. Tegner, Y., Lysholm, J., Lysholm, M., and Gillquist, J.: A performance test to monitor rehabilitation and for evaluation of anterior cruciate ligament injuries. Am. J. Sports Med. (In press).
14. Turba, J., Walsh, M., and McLeod, W.: Long-term results of extensor mechanism reconstruction: A standard for evaluation. Am. J. Sports Med. 7:91, 1979.
15. Venn, J.: Symbolic Logic. London, MacMillan, 1894.