

Evaluation of Skills in Arthroscopic Training Based on Trajectory and Force Data

Yasutaka Tashiro MD, Hiromasa Miura MD, PhD,
Yoshitaka Nakanishi PhD, Ken Okazaki MD, PhD,
Yukihide Iwamoto MD, PhD

Received: 6 February 2008 / Accepted: 20 August 2008 / Published online: 13 September 2008
© The Association of Bone and Joint Surgeons 2008

Abstract Objective evaluation of surgical skills is essential for an arthroscopic training system. We asked whether a quantitative assessment of arthroscopic skills using scores, time to completion, instrument tip trajectory data, and force data was valid. We presumed more experienced surgeons would perform better on a simulated arthroscopic procedure than novices, therefore validating the quantitative assessment. Surgical trainees ($n = 12$), orthopaedic residents ($n = 12$), and experienced arthroscopic surgeons ($n = 6$) were tested on a Sawbones[®] knee simulator. Subjects performed a joint inspection and probing task and a partial meniscectomy task. The trajectory data were measured using an electromagnetic motion tracking system and the force data were measured using a force sensor. The experienced group performed both tasks with higher scores and more quickly than the less experienced groups. The path length of the probe and the scissors was substantially shorter and the probe velocity was considerably faster in the experienced group. The trainee group applied substantially stronger forces to the joint during the joint inspection and probing task. Our data suggest a

performance assessment using an electromagnetic motion tracking system and a force sensor provides an objective means of evaluating surgical skills in an arthroscopic training system.

Introduction

Arthroscopic surgery has become more common among patients and orthopaedic surgeons in recent years. This surgical technique offers benefits of less trauma, reduced pain, and quick recovery for patients. Diagnostic accuracy and therapeutic efficacy are major characteristics of this technique. However, arthroscopic procedures demand highly developed psychomotor skills, because observation of the field is limited and the ability to manipulate instruments is less than in traditional open surgery. Surgeons must be able to perceive a three-dimensional environment from a two-dimensional camera image and handle the equipment skillfully. Therefore, it is important for arthroscopic trainees to improve their surgical skills and understand their skill levels objectively to safely and effectively perform arthroscopic surgery.

Arthroscopic training and skill assessment with cadavers, animals, or direct observation are limited because of ethical restrictions and low availability [3, 10, 16]. Therefore, numerous investigators have developed virtual reality simulators for arthroscopy [1, 2, 4, 6, 8, 9, 12, 15]. A knee arthroscopy instructional course has been available at the Kyushu University Training Center for Minimally Invasive Surgery since 2005. A critical challenge for surgical education is to objectively evaluate the skill levels of surgical trainees [2, 8, 9, 12, 15]. Assessing the instrument tip trajectory reportedly is useful in training for laparoscopic surgery [5, 11]. Measurement with a force and

One or more of the authors (YT) has received funding from a Grant of Japan Orthopaedics and Traumatology Foundation, Inc. (No.0183). Each author certifies that his institution has approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent was obtained.

Y. Tashiro, H. Miura (✉), K. Okazaki, Y. Iwamoto
Department of Orthopaedic Surgery, Graduate School of
Medical Sciences, Kyushu University, 3-1-1 Maidashi,
Higashi-ku, Fukuoka 812-8582, Japan
e-mail: miura@ortho.med.kyushu-u.ac.jp

Y. Nakanishi
Digital Medicine Initiative, Kyushu University, Fukuoka, Japan

torque sensor is similarly useful to evaluate skills in endoscopic sinus surgery and laparoscopic surgery training [13, 14, 19].

We asked whether quantitative evaluation using these scores (ie, the score representing the number of figures touched in the time limit in a joint inspection and probing task, and the scores from a partial meniscectomy task), time to completion, instrument tip trajectory data, and force data would be useful to distinguish levels of skill in arthroscopic surgery. We hypothesized more experienced surgeons would have better motor performance in a simulated arthroscopic procedure than individuals with less experience with the technique.

Materials and Methods

We identified three groups of volunteers with different stages of arthroscopic skills. Group 1 consisted of surgical trainees who had observed but had not performed any arthroscopies ($n = 12$), Group 2 consisted of orthopaedic residents with 10 to 20 cases of actual arthroscopic operations ($n = 12$), and Group 3 were experienced arthroscopic surgeons ($n = 6$). All performed two kinds of simulated arthroscopic tasks. Quantitative data of scores, time to completion, instrument tip trajectory, and surgical force were obtained and differences among the three groups were examined. The parameters with a difference in accordance with the surgical skills were regarded to be valid in distinguishing levels of arthroscopic surgical skills. A power analysis was performed based on the path length data of the arthroscope in our preliminary research (a difference in the mean = 2200 mm, the standard deviation = 1500 mm, significant level = 0.05, and power = 0.80) to indicate a sample size of 9.7 in each group could address the questions.

We used an electromagnetic motion tracking system to measure the path length and velocity of the arthroscope and probe and scissors. The Aurora measurement system (Northern Digital Inc, Waterloo, Canada) is designed to calculate the position and orientation of sensor coils (Fig. 1). The field generator creates an electromagnetic field with a characterized volume of 500 mm \times 500 mm \times 500 mm. Sensor coils react to the electromagnetic field and produce signals. The sensor interface unit transmits the signals to the system control unit. The system control unit calculates the position and orientation of the sensor coils based on these data and communicates the results to the host computer. The system can determine five degrees of freedom for each sensor coil tool except the rotation around the sensor coil's z-axis. It has a maximum measurement rate of 45 Hz if measured with five or less sensor coils. In this experiment, we attached the sensor coils to the shaft of the operative instrument.



Fig. 1 An electromagnetic measurement system is shown. The field generator creates a measurement volume of 500 mm \times 500 mm \times 500 mm. Sensor coils are attached to the surgical instrument.

We used a force analysis system to estimate the degree of surgical forces administered during operative procedures. The six degrees of freedom force sensor (Nitta Co, Osaka, Japan) provides three force measurements along three orthogonal axes and three torque measurements along the same axes. It is connected to an interface unit, which consists of an analog-to-digital converter and an amplifier. This interface unit is connected to the host computer. The sensor detects values (forces F_x , F_y , F_z ; torques M_x , M_y , M_z) concurrently with a sampling frequency of 8 kHz and outputs the data set every 125 μ s. We fixed the force sensor beneath the knee model in the training center (Fig. 2). Zero point adjustment can be completed in approximately 5 seconds before the measurement. Once surgical forces are loaded on the joint model, the output of the force sensor is digitized and recorded by the host computer and then shown in real time on a computer display. We used only the data sets of three forces in this study because the load center of surgical force, which was located in front of the origin of coordinates, varied according to each procedure.

After orientation and 5 minutes of practice, subjects performed two kinds of tasks designed to assess the core psychomotor skills needed in arthroscopy. The tasks were

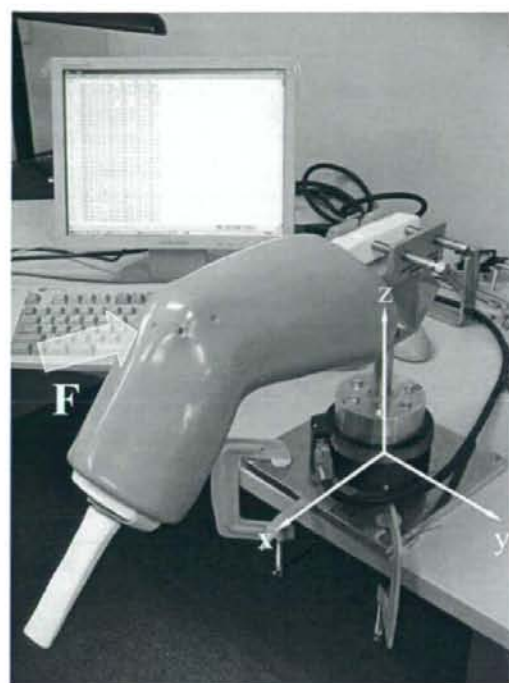


Fig. 2 A six degrees of freedom force sensor is fixed beneath the knee model. This sensor detects surgical forces loaded on the joint, providing the data set of forces F_x , F_y , and F_z and torques M_x , M_y , and M_z .

performed using the Sawbones[®] knee simulator (Pacific Research Laboratories Inc, Vashon, WA). Task 1 was a joint inspection and probing task. Subjects handled the scope and probe to touch the figures located variously in

the joint. A total of 10 figures were placed. This task was limited to 5 minutes. Task 2 was a partial meniscectomy task. The amount of meniscus to be removed was indicated with a line in advance and subjects performed the resection along the line. An excessive resection was indicated with a red line. Task 2 had a limit of 6 minutes. The score in Task 1 was the number of figures touched within the time limit. Scores in Task 2 were graded as excellent, good, fair, or poor for each skill regarding the adequacy of the amount and smoothness of the margin in resection. The computer calculated the following parameters: time to complete each task, trajectory data of the tip of the surgical instruments, and force data. We assessed three indices of force exerted: (1) peak value; (2) average value; and (3) integration of the absolute values.

To determine the differences in the continuous variables (such as scores in Task 1, time to completion, instrument tip trajectory data, and force data) among the three groups, we used a one-way factorial ANOVA with Fisher's protected least significant difference post hoc test. Differences in the discrete variables (scores in Task 2) among the three groups were determined by the Kruskal-Wallis test, Mann-Whitney U test, and Bonferroni correction. Differences in the motor performance on a simulated arthroscopic procedure, in accordance with subjects' surgical skills, was regarded as validation of the quantitative evaluation method.

Results

The expert group obtained better scores ($p = 0.014$) than the resident group, and the resident group obtained better scores ($p = 0.003$) than the trainee group in Task 1 (Table 1). In Task 2, the expert and the resident groups

Table 1. Results of Task 1

Parameter	Group 1 12	Group 2 12	p Value	Group 3 6	p Value
Score	2.5 ± 2.6	6.3 ± 3.6	0.003 ^{*,†}	10 ± 0	0.000 ^{*,†} ; 0.014 ^{†,‡}
Time (seconds)	300 ± 0	287.8 ± 22.9	0.292 [*]	198.8 ± 54.8	0.000 ^{*,†} ; 0.000 ^{†,‡}
Path length of the scope (mm)	7978 ± 2604	6348 ± 1239	0.043 ^{*,‡}	3815 ± 912	0.000 ^{*,†} ; 0.012 ^{†,‡}
Velocity of the scope (mm/second)	26.6 ± 8.7	22.2 ± 4.7	0.111 [*]	21.5 ± 4.7	0.130 ^{*,†} ; 0.828 [†]
Path length of the probe (mm)	1262 ± 442	996 ± 420	0.109 [*]	485 ± 134	0.001 ^{*,†} ; 0.015 ^{†,‡}
Velocity of the probe (mm/second)	13.6 ± 9.5	18.1 ± 8.5	0.201 [*]	27.1 ± 4.1	0.003 ^{*,†} ; 0.040 ^{†,‡}
Peak force (N)	15.6 ± 6.2	18.2 ± 8.7	0.366 [*]	11.7 ± 4.3	0.279 ^{*,†} ; 0.074 [†]
Average force (N)	3.4 ± 1.0	3.3 ± 1.1	0.652 [*]	2.4 ± 0.5	0.049 ^{*,†} ; 0.102 [†]
Integration of force (N second)	1030 ± 310	924.2 ± 285.9	0.357 [*]	455 ± 139	0.000 ^{*,†} ; 0.002 ^{†,‡}

Group 1 = surgical trainees; Group 2 = orthopaedic residents; Group 3 = experienced arthroscopic surgeons; mean ± standard deviation; ^{*}p value for the comparison of Groups 1 and 2; [†]p value for the comparison of Groups 1 and 3; [‡]p value for the comparison of Groups 2 and 3; [§]Significant difference.

Table 2. Results of Task 2

Parameter	Group 1 12	Group 2 12	p Value	Group 3 6	p Value
Time (seconds)	360 ± 0	339.6 ± 34.0	0.212*	298.5 ± 75.7	0.004**†; 0.045†‡
Path length of the scope (mm)	2257 ± 1066	4727 ± 1028	0.014*†‡	4894 ± 1717	0.021**†‡; 0.785†
Velocity of the scope (mm/second)	9.3 ± 3.0	14.1 ± 3.7	0.006*†‡	16.7 ± 4.7	0.001**†‡; 0.174†
Path length of the scissors (mm)	12845 ± 3032	10829 ± 1522	0.042*†‡	9661 ± 1848	0.010**†‡; 0.321†
Velocity of the scissors (mm/second)	35.7 ± 8.4	32.0 ± 4.1	0.175*	33.4 ± 6.1	0.496**†; 0.658†
Peak force (N)	17.9 ± 6.6	26.3 ± 6.8	0.004*†‡	22.7 ± 6.0	0.154**†; 0.285†
Average force (N)	4.0 ± 1.4	4.4 ± 1.0	0.394*	4.4 ± 1.0	0.427**†; 0.921†
Integration of force (N-second)	1433 ± 493	1497 ± 395	0.741*	1372 ± 571	0.798**†; 0.600†

Group 1 = surgical trainees; Group 2 = orthopaedic residents; Group 3 = experienced arthroscopic surgeons; mean ± standard deviation; *p value for the comparison of Groups 1 and 2; **p value for the comparison of Groups 1 and 3; †p value for the comparison of Groups 2 and 3; ‡Significant difference.

obtained better scores than the trainee group in adequacy of the resection ($p = 0.003$ for Group 1 vs 3; $p = 0.029$ for Group 1 vs 2) and smoothness of the margin ($p = 0.011$ for Group 1 vs 3; $p = 0.004$ for Group 1 vs 2), but we observed no difference in scores between the expert and resident groups (Fig. 3).

The expert group completed the procedures more quickly than the two less experienced groups in Task 1 ($p = 0.000$ for Group 1 vs 3; $p = 0.000$ for Group 2 vs 3) (Table 1) and Task 2 ($p = 0.004$ for Group 1 vs 3; $p = 0.045$ for Group 2 vs 3) (Table 2). We observed no

difference in the mean time needed for each task between the resident and trainee groups.

In Task 1, the path length of the scope was shorter ($p = 0.012$) in the expert group than in the resident group and shorter ($p = 0.043$) in the resident group than in the trainee group, whereas the velocity of the scope was similar ($p = 0.180$) among the three groups. The path length of the probe was shorter ($p = 0.001$ for Group 1 vs 3; $p = 0.015$ for Group 2 vs 3) (Fig. 4A) and the velocity of the probe was faster ($p = 0.003$ for Group 1 vs 3; $p = 0.040$ for Group 2 vs 3) in the expert group than in the two less experienced groups (Table 1). In Task 2, the path lengths of the scope were longer ($p = 0.021$ for Group 1 vs 3; $p = 0.014$ for Group 1 vs 2) and the velocities of the scope were faster ($p = 0.001$ for Group 1 vs 3; $p = 0.006$ for Group 1 vs 2) in the expert and resident groups than in the trainee group (Fig. 4B). The path lengths of the scissors in the expert and resident groups were shorter ($p = 0.010$ for Group 1 vs 3; $p = 0.042$ for Group 1 vs 2) than that of the trainee group. The velocities of the scissors were similar ($p = 0.389$) among the three groups (Table 2).

The integrated force values applied by the expert group were lower ($p = 0.000$ for Group 1 vs 3; $p = 0.002$ for Group 2 vs 3) in Task 1 than those applied by the two less experienced groups (Table 1; Fig. 5). In Task 2, the peak force applied by the resident group was higher ($p = 0.004$) than that of the trainee group (Table 2). However, the average ($p = 0.613$) and integration ($p = 0.862$) of the force values were similar among the three groups (Table 2).

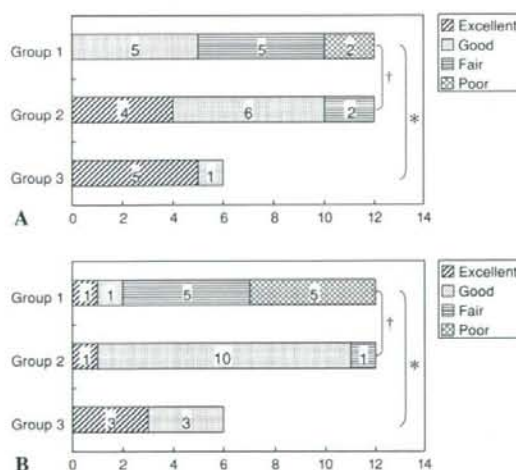
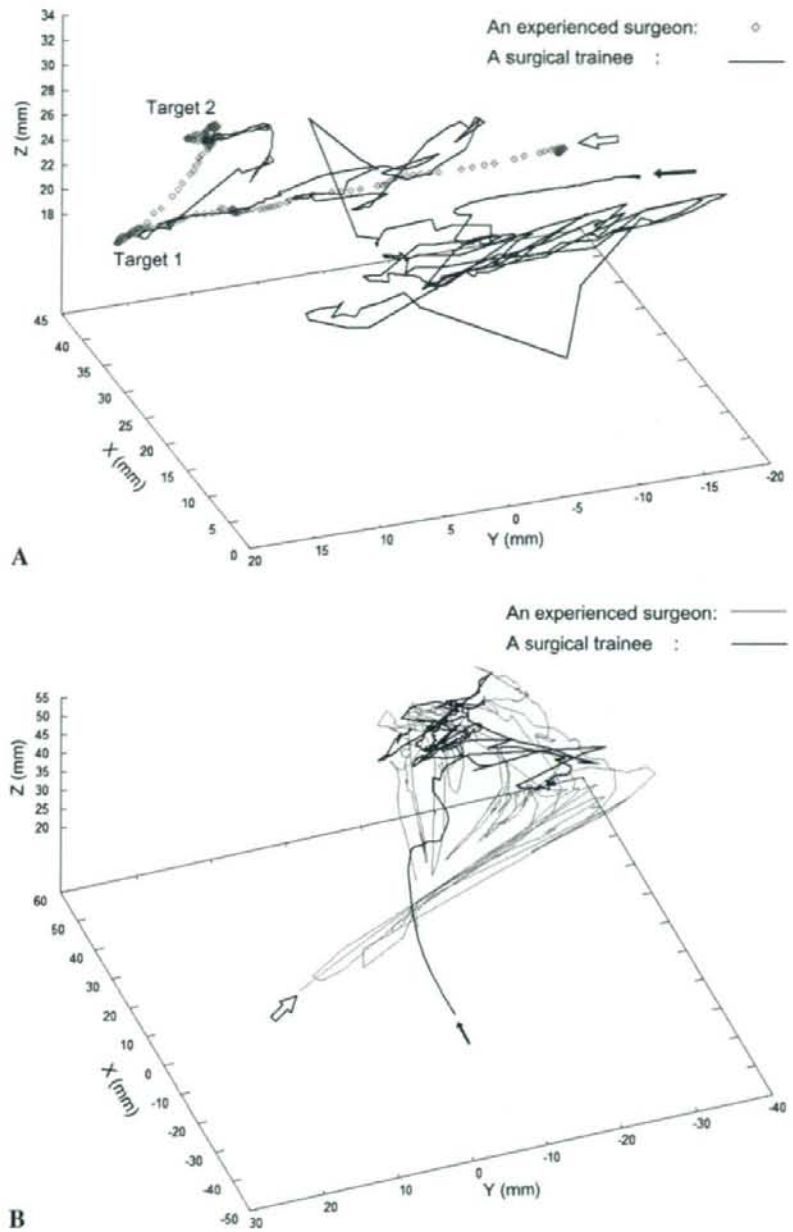


Fig. 3A–B (A) Scores for adequacy of the resection in Task 2, the meniscectomy task, are shown. The experienced surgeons (Group 3) and the orthopaedic residents (Group 2) showed better (*: $p = 0.003$; †: $p = 0.029$) performance than the surgical trainees (Group 1). (B) Scores for smoothness of the margin in Task 2 show Groups 2 and 3 obtained better (*: $p = 0.011$; †: $p = 0.004$) scores than Group 1.

Discussion

Advances in arthroscopic surgical technique have revolutionized the diagnostic accuracy and therapeutic efficacy

Fig. 4A–B (A) The probe tip trajectories of an experienced surgeon (dotted line) and a surgical trainee (full line) for Task 1 are shown. The probe path length of the experienced surgeon is considerably shorter and smoother than that of the trainee. (B) The scope tip trajectories of an experienced surgeon (thin line) and a surgical trainee (thick line) in Task 2 are shown. The experienced surgeon manipulated the scope more for better observation.



for intraarticular disorders [12, 17]. Although arthroscopic procedures demand highly developed psychomotor skills, the opportunity for orthopaedic trainees to learn those skills is limited and a means of objectively evaluating the skill levels of surgical trainees has not been established. This

study was designed to clarify whether quantitative evaluation of arthroscopic surgical skills using scores, time to completion, instrument tip trajectory data, and force data was valid. We hypothesized more experienced surgeons therefore would obtain higher scores, complete tasks more

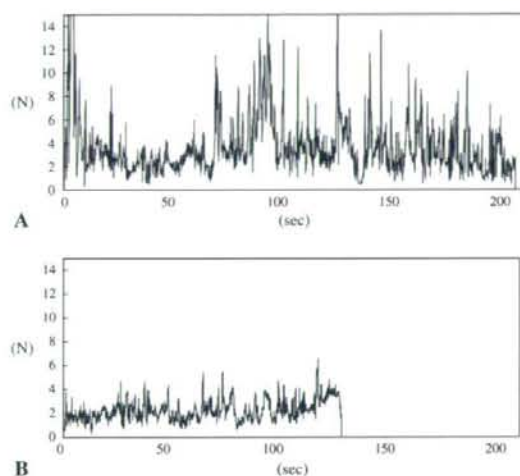


Fig. 5A–B (A) The time-series data of the output of a force sensor for a trainee are shown. Many large spikes are seen. (B) The force data of an experienced surgeon show time to completion is shorter and the forces are lower than those for the trainee.

quickly, and manipulate surgical devices more smoothly without loading any excessive forces on tissues in a simulated arthroscopic procedure than individuals with less experience, and this difference thus would validate the objective evaluation method.

One of the limitations of this study was simulated arthroscopic procedures were performed with Sawbones[®] physical knee models. Fresh cadavers would be more desirable to train and assess surgical skills of trainees outside the operating room, however, ethical restrictions and limited availability make use of cadavers difficult. To assess improvement of actual surgical skills of a person, a skill evaluation would be needed after he or she had gained a certain amount of real surgical experiences. However, the differences in motor performance depending on the real surgical experience suggest the evaluation method is valid and reliable. We had subjects perform only two kinds of tasks to assess their arthroscopic skills, although there are various situations and numerous techniques in arthroscopic surgery. We believe a simplified method would be more suitable for skill evaluation in surgical training than a complex and time-consuming one. A convenient means of discriminating surgical skills is needed in surgical education.

The score on Task 1 increased according to the surgical experience of the subjects. The score on Task 2 discriminated the more experienced two groups from the less experienced group. Previous studies suggest scoring surgical performance on simulated tasks is useful to assess surgical skills [7, 11, 15].

The experienced group performed both tasks more quickly than the two less experienced groups. Time to completion has been a popular parameter to evaluate surgical skills in arthroscopic and endoscopic training [2, 8, 12, 15, 18], although it cannot clarify the details of psychomotor skills in which the trainee lacks. These results of score and time data support the hypothesis that individuals with more experience in performing actual arthroscopic procedures also will perform better in simulated tasks.

An analysis of the trajectory data suggests experts manipulate the probe more smoothly and quickly than less experienced surgeons. Compared with experienced surgeons, the trainees had more unnecessary movements of the scissors. Similar results were reported in an evaluation study with a virtual reality simulator for shoulder arthroscopy. In that study, the probe path length of specialists was less than half as long as that of novices with no prior experience, and also shorter than that of residents with some experience [2]. Gomoll et al. also reported the probe velocity of specialists was almost double that of novices [2]. Another study with a shoulder arthroscopy simulator reported probe path lengths of experienced surgeons were shorter than those of trainees and residents [12]. A shorter path of the scope during a joint inspection and probing task, depending on surgical experience, suggests smoothness in handling the scope by more skillful surgeons. A validation study for a knee simulator suggested more experienced surgeons had substantially shorter scope path lengths during a loose body searching task, which required similar techniques as our joint inspection and probing task [9]. In Task 2, however, the positional data of the scope did not match this tendency. This suggests an increased trajectory of the scope does not mean more unnecessary movement in a task such as a meniscectomy, which demands different procedures from the joint inspection task. In fact, with Task 2, experienced surgeons tended to change their vantage points with freedom to make the meniscectomy work easier (Fig. 4B).

Force analysis is reportedly valuable to assess interference between surgical tools and tissues [13, 14, 19]. A previous study using a force sensor for an endoscopic sinus surgery training system proposed the peak as the instantaneous strong force, the average as the average force on the tissue, and the integration as the total force on the tissue during the tasks [19]. We found the integration of force by the two less experienced groups to be higher than values of experienced surgeons in the joint inspection and probing task. These forces were believed unnecessary, because the experienced surgeons used no such excessive forces. The use of excessive force might cause damage to the articular cartilage and soft tissues. However, force indices for Task 2 showed no differences between experienced surgeons and the two less experienced groups. The difficulty in

distinguishing harmful from harmless contact would be the reason for this result. In one study, high force magnitudes were applied by novice surgeons in comparison to expert surgeons while performing tissue manipulation and vice versa in tasks involving tissue dissection [14]. Even useful manipulation of tools, such as repetitively cutting the meniscus, was detected as a harmful contact with the tissue, because the force sensor fixed beneath the joint model sensed the entire power loaded to the joint. A study evaluating a virtual reality simulator for shoulder arthroscopy also reported some of the useful probe contacts were potentially harmful collisions [15]. A force analysis system is required to improve the discrimination to resolve this problem.

It is important to assess the skill levels of surgical trainees. An objective evaluation system helps surgeons understand their degree of improvement and which technique they need to improve. We found score and time to task completion clearly discriminated the performance of the operators depending on their surgical experience. The path lengths of the probe and scissors by the experienced surgeons were shorter than those of the less experienced. The force analysis system detected excessive interference between the surgical tools and tissues for the joint inspection and probing task. Our data suggest this quantitative method of evaluating arthroscopic surgical skills using scores, time to completion, an electromagnetic motion tracking system, and a force sensor distinguish the level of surgical skill in an arthroscopic training system.

Acknowledgments We thank Shuichi Matsuda, MD, PhD (Department of Orthopaedic Surgery, Kyushu University), Makoto Hashizume, MD, PhD, and Kazuo Tanoue, MD, PhD (Department of Advanced Medical Initiatives, Faculty of Medical Sciences, Kyushu University), for assistance with this study, and we thank Munechika Misumi (Department of Medical Informatics, Kyushu University Hospital) for help with the statistical analysis.

References

- Cannon WD, Eckhoff DG, Garrett WE Jr, Hunter RE, Sweeney HJ. Report of a group developing a virtual reality simulator for arthroscopic surgery of the knee joint. *Clin Orthop Relat Res* 2006;442:21–29.
- Gomoll AH, O'Toole RV, Czarnecki J, Warner JJ. Surgical experience correlates with performance on a virtual reality simulator for shoulder arthroscopy. *Am J Sports Med*. 2007;35:883–888.
- Grechenig W, Fellingner M, Fankhauser F, Weiglein AH. The Graz learning and training model for arthroscopic surgery. *Surg Radiol Anat*. 1999;21:347–350.
- Heng PA, Cheng CY, Wong TT, Wu W, Xu Y, Xie Y, Chui YP, Chan KM, Leung KS. Virtual reality techniques: application to anatomic visualization and orthopaedics training. *Clin Orthop Relat Res*. 2006;442:5–12.
- Leong JJ, Nicolaou M, Atallah L, Mylonas GP, Darzi AW, Yang GZ. HMM assessment of quality of movement trajectory in laparoscopic surgery. *Comput Aided Surg*. 2007;12:335–346.
- Mabrey JD, Gillogly SD, Kasser JR, Sweeney HJ, Zarins B, Mevis H, Garrett WE Jr, Poss R, Cannon WD. Virtual reality simulation of arthroscopy of the knee. *Arthroscopy*. 2002;18:E28.
- Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84:273–278.
- McCarthy A, Harley P, Smallwood R. Virtual arthroscopy training: do the “virtual skills” developed match the real skills required? *Stud Health Technol Inform*. 1999;62:221–227.
- McCarthy AD, Moody L, Waterworth AR, Bickerstaff DR. Passive haptics in a knee arthroscopy simulator: is it valid for core skills training? *Clin Orthop Relat Res*. 2006;442:13–20.
- Morris AH, Jennings JE, Stone RG, Katz JA, Garroway RY, Hendler RC. Guidelines for privileges in arthroscopic surgery. *Arthroscopy*. 1993;9:125–127.
- Munz Y, Kumar BD, Moorthy K, Bann S, Darzi A. Laparoscopic virtual reality and box trainers: is one superior to the other? *Surg Endosc*. 2004;18:485–494.
- Pedowitz RA, Esch J, Snyder S. Evaluation of a virtual reality simulator for arthroscopy skills development. *Arthroscopy*. 2002;18:E29.
- Rosen J, Hannaford B, MacFarlane MP, Sinanan MN. Force controlled and teleoperated endoscopic grasper for minimally invasive surgery: experimental performance evaluation. *IEEE Trans Biomed Eng*. 1999;46:1212–1221.
- Rosen J, Hannaford B, Richards CG, Sinanan MN. Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. *IEEE Trans Biomed Eng*. 2001;48:579–591.
- Srivastava S, Youngblood PL, Rawn C, Hariri S, Heinrichs WL, Ladd AL. Initial evaluation of a shoulder arthroscopy simulator: establishing construct validity. *J Shoulder Elbow Surg*. 2004;13:196–205.
- Voto SJ, Clark RN, Zuelzer WA. Arthroscopic training using pig knee joints. *Clin Orthop Relat Res*. 1988;226:134–137.
- Yamaguchi K, Ball CM, Galatz LM. Arthroscopic rotator cuff repair: transition from mini-open to all-arthroscopic. *Clin Orthop Relat Res*. 2001;390:83–94.
- Yamaguchi S, Konishi K, Yasunaga T, Yoshida D, Kinjo N, Kobayashi K, Ieiri S, Okazaki K, Nakashima H, Tanoue K, Maehara Y, Hashizume M. Construct validity for eye-hand coordination skill on a virtual reality laparoscopic surgical simulator. *Surg Endosc*. 2007;21:2253–2257.
- Yamauchi Y, Yamashita J, Morikawa O, Hashimoto R, Yokoyama K. An endoscopic sinus surgery training system for assessment of surgical skill. *Stud Health Technol Inform*. 2004;98:416–418.



Varus–valgus laxity correlates with pain in osteoarthritis of the knee

Hiromasa Miura^{a,*}, Shin-ichiro Takasugi^a, Tsutomu Kawano^b, Takashi Manabe^b, Yukihide Iwamoto^b

^a Department of Rehabilitation Medicine, Kyushu University Hospital, Fukuoka, Japan

^b Department of Orthopaedic Surgery, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

ARTICLE INFO

Article history:

Received 2 February 2008

Received in revised form 23 August 2008

Accepted 14 September 2008

Keywords:

Knee
Osteoarthritis
Pain
Risk factor
Varus–valgus laxity

ABSTRACT

Pain during osteoarthritis (OA) of the knee does not necessarily correlate with the severity of the radiographic grade, and the mechanism of pain has not been completely clarified. The purpose of this study was to evaluate risk factors for pain in the knee OA using epidemiologic analyses.

We evaluated 518 out of 4183 people over the age of 40 (156 males and 362 females) from Shinyoshitomi village, Japan. Mean ages were 63.8 years for men and 60.7 years for women. Screening included a physical examination of the knee and a standing AP roentgenogram of the bilateral knee. Radiographic OA was defined as a Kellgren–Lawrence grade 2 or higher. All data were coded and pain risk factors were evaluated using a multiple logistic regression model.

Radiographic OA was observed in 18.4% of men and 26% of women. Of these subjects with OA, 10.9% of men and 32.5% of women complained of knee pain. Seven factors—age, gender, BMI, radiographic grade, varus–valgus laxity, torque of quadriceps muscles, and varus–valgus alignment—were evaluated as potential risk factors for pain. A significant increase in the odds ratio was observed with varus–valgus laxity ($p=0.005$; odds ratio, 3.04). Our results suggest that varus–valgus laxity is a risk factor for pain during knee OA.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

It is well recognized that pain during knee osteoarthritis (OA) does not necessarily correlate with the severity of the radiographic grade of OA [1]. In spite of severe radiographic OA, some patients, more often men than women, show no pain or mild pain [2,3]. A direct relationship between measures of severity of radiographic knee OA and pain or functional disability has been described by some [4], but not all authors [5,6].

There are various conjectures concerning the etiology of knee pain during OA, but the mechanism of pain has not been completely clarified. Intraosseous hypertension is one factor that has been associated with a deep aching bone pain, particularly at rest [7]. Tissues in the knee that possess nociceptors include the joint capsule, ligaments, synovium, bone, and the outer edge of the menisci [8].

As pain in knee OA occurs predominantly during gait and stair climbing, dynamic factors such as instability or thrust may contribute to pain. A relationship between varus–valgus laxity and progression of OA has been shown [9,10]. However, only a few attempts have been made at characterizing the relationship between laxity and pain. Gibson et al. described that no relationship could be demonstrated

between knee pain and joint laxity from their population study in Pakistan [11].

The purpose of this study, therefore, was to evaluate risk factors for pain during knee OA. Towards this end, we investigated the relevance of laxity on pain during knee OA using epidemiologic analyses. The null hypothesis is that the presence of varus–valgus laxity does not influence pain in knee osteoarthritis.

2. Materials and methods

In October 2000 and March 2001, the Survey of Bone and Joint Diseases, which focused on disorders such as knee OA, degenerative spinal disorders, and osteoporosis, was conducted in a rural community. The population of this community was estimated at 4240. A total of 518 subjects (156 men and 362 women), ranging in age between 24 and 87 years, voluntarily participated in this survey which was performed in concert with cancer screening to prevent a self-selection bias of the participants. The mean age was 63.8 years old [standard deviation (SD): 9.9] for men and 60.7 years old (SD: 10.9) for women.

A physical examination of the knee including knee laxity, and varus–valgus deformity was conducted for each subject. Knee laxity was manually evaluated by the one examiner who had more than 20 years of experience in knee surgery. The examiner was blinded to the radiographic data. For assessment of varus–valgus laxity, varus and valgus stress was applied manually with the knee flexed at 20 degrees while subjects were in supine position, and for evaluation of

* Corresponding author. Department of Rehabilitation Medicine, Kyushu University Hospital, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan. Tel.: +81 92 642 5860; fax: +81 92 642 5864.

E-mail address: miura@ortho.med.kyushu-u.ac.jp (H. Miura).

Table 1
Incidence of pain and laxity

	Knees	Pain +	V-v laxity +	A-p laxity +
Men				
OA +	46	5 (10.9%)	6 (13.0%) *	1 (2.2%)
OA -	204	25 (12.3%)	18 (8.8%)	0 (0.0%)
Subtotal	250	30 (12.0%)	24 (9.6%)	1 (0.4%)
Women				
OA +	169	55 (32.5%)	82 (48.5%) *	2 (1.2%)
OA -	481	36 (7.5%)	69 (14.3%)	0 (0.0%)
Subtotal	650	91 (14.0%)	151 (23.2%)	2 (0.03%)
Total	900	121 (13.4%)	175 (19.4%) 3	(0.03%)

V-v: varus-valgus, a-p: antero-posterior, * $p < 0.0001$.

antero-posterior laxity, the Lachman test and posterior drawer test were performed. The joint laxity in one knee was compared with that of the opposite side. An AP standing radiogram of both knees was taken. A physical test of isometric quadriceps strength measured by a hand-held dynamometer was performed on each subject.

Of these parameters, the AP standing radiogram of the knees, manual joint test for varus-valgus and antero-posterior laxity, varus-valgus deformity, measurement of quadriceps torque, and assessment of current gait pain in the knees were used for this study. Two orthopaedic surgeons with more than 10 years of experience, who were blinded to the data of the physical examination, did the grading of the X-rays using the Kellgren-Lawrence radiographic classification. Knee OA was defined as a Kellgren-Lawrence score greater than grade 2 [12]. Knee laxity was graded manually as -, ±, +, ++. Laxity more than + was defined as "loose" and laxity less than ± was defined as "normal". The varus-valgus deformity was graded as -, +.

Chi-square test and a multiple logistic regression model were used for statistical analysis. The level of absolute statistical significance was set at $p < 0.05$. We examined the association between pain and various factors using bivariate regression analyses. A statistical significance level of $p < 0.10$ was used as a screening cutoff.

3. Results

A total of 900 knees from 450 subjects (125 men and 325 women) with valid knee X-ray data were analyzed. Radiographic knee OA was observed in 46 knees (18.4%) of men and in 169 knees (26%) of women. In men, knee pain was reported in 25 knees (12.3%) out of the 204 knees without radiographic OA as well as in five knees (10.9%) out of the 46 knees with radiographic OA. In women, knee pain was reported in 36 knees (7.5%) out of the 481 knees without radiographic OA and in 55 knees (32.5%) out of the 169 knees with radiographic OA. The incidence of reported knee pain associated with radiographic OA was significantly greater in women than in men ($p < 0.01$).

Varus-valgus laxity was observed in 24 knees (9.6%) of men and 151 knees (23.2%) of women in total. Of these, varus-valgus laxity was observed in six knees (13%) of men and 82 knees (48.5%) of women with actual radiographic OA. The incidence of varus-valgus laxity associated with radiographic OA was significantly greater in women than in men ($p < 0.0001$). Antero-posterior laxity was observed only in one knee of men and two knees of women. No significant gender difference in antero-posterior laxity was found (Table 1).

Seven factors—age, gender, body mass index (BMI), radiographic grade, varus-valgus laxity, torque of quadriceps muscles, and varus-valgus alignment—were evaluated by a multiple logistic regression model as possible risk factors for knee pain. A significant increase in the odds ratio of knee pain was only observed with varus-valgus laxity [$p < 0.005$; odds ratio, 3.04 (95% CI: 1.41–6.54)] (Table 2).

Table 2
Risk factors for pain in knee OA

	p value	Odds ratio (95% CI)
Age	0.682	1.01 (0.96–1.06)
Gender	0.246	0.48 (0.14–1.65)
BMI	0.228	1.08 (0.95–1.23)
Radiographic grade	0.267	1.34 (0.80–2.25)
Varus-valgus laxity	0.005	3.04 (1.41–6.54)
Torque of quadriceps	0.316	1.00 (0.99–1.00)
Varus-valgus alignment	0.058	2.00 (0.98–4.10)

Table 3
Risk factors for varus-valgus laxity

	p value	Odds ratio (95% CI)
Age	0.066	1.05 (1.00–1.10)
Gender	0.007	0.18 (0.05–0.61)
BMI	0.414	1.06 (0.93–1.21)
Radiographic grade	0.006	2.14 (1.24–3.70)
Torque of quadriceps	0.190	1.00 (0.99–1.00)
Varus-valgus alignment	<0.0001	6.10 (3.12–11.93)

Six factors—age, gender, BMI, radiographic grade, torque of quadriceps muscles and varus-valgus alignment—were evaluated as possible risk factors for varus-valgus laxity. A significant increase in the odds ratio of such laxity was observed with female gender, radiographic grade, and varus-valgus alignment (Table 3).

4. Discussion

Static factors such as knee alignment affect progression and functional decline in knee OA [13,14]. However, several authors suggest that varus-valgus laxity is related to the development and progression of knee OA. Wada et al. reported that varus-valgus laxity tends to increase with the severity of disease [9], and Sharma et al. demonstrated that a portion of increasing laxity of the knee precedes disease development and may in fact predispose individuals to disease [10]. Moreover, dynamic factors such as varus thrust are important in the progression of OA. In fact, Chen et al. have asserted that knees with a thrust are a subset of varus-aligned knees that are at particularly high risk for progression of OA [15]. This study shows a relationship between varus-valgus malalignment and varus-valgus laxity. Repetitive stretching of the collateral ligament due to malalignment increases capsulo-ligamentous laxity.

On the other hand, the relationship between OA and pain has not been clearly defined. It is clinically well recognized that even severe OA of the knee is not necessarily associated with severe knee pain [1]. Ciuttini et al. [5] reported that the presence of osteophytes predicts pain more accurately than a radiographic grade of severity, while the presence or absence of joint space narrowing in antero-posterior, lateral and skyline radiological views of the knee does not correlate with knee pain. Bruyere et al. [6], on the other hand, have suggested that radiographic and clinical progression of disease is significantly associated with knee pain. The clinical relevance of this association, however, is questionable.

Because varus-valgus laxity is thought to be intimately related to the onset and progression of knee OA [9,10,15], we hypothesized that pain in knee OA occurs when varus-valgus laxity is present. In this study, we demonstrate that varus-valgus laxity is a risk factor for knee pain in OA and that female gender, the grade of OA and varus-valgus alignment are risk factors for varus-valgus laxity. To our knowledge, this is the first report on the relationship between varus-valgus laxity and pain in knee OA.

Knee varus-valgus laxity in individuals with knee OA is most likely a multi-factorial problem that can result from such factors as increased capsulo-ligamentous laxity, structural damage to the knee, and altered lower extremity muscular strength and neuromuscular control [16]. Varus-valgus laxity increases with age and is greater in women than men [10]. This may well be the reason why men do not report knee pain as frequently as women, even in the setting of severe OA. Notably, antero-posterior laxity is not affected by age, and there is no difference in antero-posterior laxity between the OA and control group [10]. In agreement with our own results, Wada et al. report that antero-posterior laxity with OA steadily declines along with the Kellgren-Lawrence grade [9].

Various conservative treatments are effective in relieving the pain of knee OA. Interestingly, these treatments are all aimed at decreasing varus-valgus laxity. Thus, their effectiveness would seem to support our finding that varus-valgus laxity is fundamentally related to the

presence of pain in knee OA. Valgus bracing, for example, is effective in reducing external adduction movement and appears to subsequently provide pain relief [17–20]. The lateral-wedge insole has also been proven effective for the conservative treatment of medial compartment knee OA via a reduction in external varus movement and medial compartment load [21,22]. In addition, several reports describe the effectiveness of quadriceps exercise in pain relief. In our study, we did not find a direct association between quadriceps torque and the presence of pain. Quadriceps exercise may nonetheless effect pain relief through an improvement in instability. High tibial osteotomy is also considered effective for pain relief and functions by reducing peak adduction movement of the knee [23].

There are some limitations to our study. First, varus–valgus laxity of the knee was not quantitatively evaluated. Cushnaghan et al. demonstrate the poor reliability of physical examination in the assessment of laxity even with a single examiner (within-observer agreement: 0.55) [24]. However, computerized systems to measure knee laxity in the frontal plane (such as the Genucom) are no longer available. Even with the use of such a system, it is difficult to evaluate laxity using any special device on a large population of subjects in a short period of time. Yet another limitation of this study is posed by the fact that we measured passive knee laxity, not dynamic instability, i.e., varus thrust. Passive knee laxity and functional knee instability are not synonymous. Passive laxity is a clinical sign that indicates either lack of tension in capsuloligamentous structures of a joint, or the degree of joint looseness on the passive motion testing of the joint [25]. However, in a knee with varus–valgus laxity, thrust is likely to occur, especially in the knee with static malalignment.

In conclusion, our study presents the novel finding that varus–valgus laxity is a risk factor for pain in the knee OA in a Japanese population. As such, it is entirely conceivable that Japanese patients with severe radiographic OA may have mild or no pain if their varus–valgus laxity is well controlled. This raises the potential for the development of new treatments for knee OA in Japanese patients that focus on improving varus–valgus stability.

5. Conflict of interest statement

No author had an association or financial involvement with any organization or commercial entity having a financial interest in or financial conflict with the subject matter or research presented in the manuscript.

Acknowledgement

The Japan Foundation for Aging and Health supported this survey.

References

- [1] Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. *J Rheumatol* 2000;27:1513–7.
- [2] Keeffe FJ, Lefebvre JC, Egert JR, Affleck G, Sullivan MJ, Caldwell DS. The relationship of gender to pain, pain behavior, and disability in osteoarthritis patients: the role of catastrophizing. *Pain* 2000;87:325–34.
- [3] Tsai YF. Gender differences in pain and depressive tendency among Chinese elders with knee osteoarthritis. *Pain* 2007;130:188–94.
- [4] Lethbridge-Cejku M, Scott Jr WW, Reichle R, Ettinger WH, Zonderman A, Costa P, et al. Association of radiographic features of osteoarthritis of the knee with knee pain: data from the Baltimore Longitudinal Study of Aging. *Arthritis Care Res* 1995;8:182–8.
- [5] Cicuttini FM, Baker J, Hart DJ, Spector TD. Association of pain with radiological changes in different compartments and views of the knee joint. *Osteoarthritis Cartil* 1996;4:143–7.
- [6] Bruyere O, Honore A, Rovati LC, Giacovelli G, Henrotin YE, Seidel L, et al. Radiologic features poorly predict clinical outcomes in knee osteoarthritis. *Scand J Rheumatol* 2002;31:13–6.
- [7] Simkin PA. Bone pain and pressure in osteoarthritic joints. *Novartis Foundation Symposium* 260, East Hanover (NJ): Novartis; 2004. p. 179–90.
- [8] Felson DT. The sources of pain in knee osteoarthritis. *Curr Opin Rheumatol* 2002;31:13–6.
- [9] Wada M, Imura S, Baba H, Shimada S. Knee laxity in patients with osteoarthritis and rheumatoid arthritis. *Br J Rheumatol* 1996;35:560–3.
- [10] Sharma L, Lou C, Felson DT, Dunlop DD, Kirwan-Mellis G, Hayes KW, et al. Laxity in healthy and osteoarthritic knees. *Arthritis Rheum* 1999;42:861–70.
- [11] Gibson T, Hameed K, Kadir M, Sultana S, Fatima Z, Syed A. Knee pain amongst the poor and affluent in Pakistan. *Br J Rheumatol* 1996;35:146–9.
- [12] Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;16:494–502.
- [13] Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *JAMA* 2001;286:188–95.
- [14] Brouwer GM, van Tol AW, Bergink AP, Belo JN, Bernsen RM, Reijnen M, et al. Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. *Arthritis Rheum* 2007;56:1204–11.
- [15] Chang A, Hayes K, Dunlop D, Hurwitz D, Song J, Cahue S, et al. Thrust during ambulation and the progression of knee osteoarthritis. *Arthritis Rheum* 2004;50:3897–903.
- [16] Fitzgerald GK, Piva SR, Irrgang JJ. Reports of joint instability in knee osteoarthritis: its prevalence and relationship to physical function. *Arthritis Rheum* 2004;51:941–6.
- [17] Draganich L, Reider B, Rimington T, Piotrowski G, Mallik K, Nasson S. The effectiveness of self-adjustable custom and off-the-shelf bracing in the treatment of varus gonarthrosis. *J Bone Jt Surg [Am]* 2006;88:2645–52.
- [18] Pollo FE, Otis JC, Backus SI, Warren RF, Wickiewicz TL. Reduction of medial compartment loads with valgus bracing of the osteoarthritic knee. *Am J Sports Med* 2002;30:414–21.
- [19] Lindenfeld TN, Hewett TE, Andriacchi TP. Joint loading with valgus bracing in patients with varus gonarthrosis. *Clin Orthop* 1997;344:290–7.
- [20] Self BP, Greenwald RM, Pfister DS. A biomechanical analysis of a medial unloading brace for osteoarthritis in the knee. *Arthritis Care Res* 2000;13:191–7.
- [21] Crenshaw SJ, Pollo FE, Calton EF. Effects of lateral-wedged insoles on kinetics at the knee. *Clin Orthop* 2000;375:185–92.
- [22] Yasuda K, Sasaki T. The mechanics of treatment of the osteoarthritic knee with a wedged insole. *Clin Orthop* 1987;215:162–72.
- [23] Wada M, Imura S, Nagatani K, Baba H, Shimada S, Sasaki S. Relationship between gait and clinical results after high tibial osteotomy. *Clin Orthop* 1998;354:180–8.
- [24] Cushnaghan J, Cooper C, Dieppe P, Kirwan J, McAlindon T, McCrae F. Clinical assessment of osteoarthritis of the knee. *Ann Rheum Dis* 1990;49:768–70.
- [25] Noyes FR, Grood ES, Torzilli PA. Current concepts review. The definitions of terms for motion and position of the knee and injuries of the ligaments. *J Bone Jt Surg [Am]* 1989;71:465–72.