The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 SEPTEMBER 11, 2008 VOL. 359 NO. 11

A Randomized Trial of Arthroscopic Surgery for Osteoarthritis of the Knee

Alexandra Kirkley, M.D.,* Trevor B. Birmingham, Ph.D., Robert B. Litchfield, M.D., J. Robert Giffin, M.D., Kevin R. Willits, M.D., Cindy J. Wong, M.Sc., Brian G. Feagan, M.D., Allan Donner, Ph.D., Sharon H. Griffin, C.S.S., Linda M. D'Ascanio, B.Sc.N., Janet E. Pope, M.D., and Peter J. Fowler, M.D.

ABSTRACT

BACKGROUND

The efficacy of arthroscopic surgery for the treatment of osteoarthritis of the knee is unknown.

METHODS

We conducted a single-center, randomized, controlled trial of arthroscopic surgery in patients with moderate-to-severe osteoarthritis of the knee. Patients were randomly assigned to surgical lavage and arthroscopic débridement together with optimized physical and medical therapy or to treatment with physical and medical therapy alone. The primary outcome was the total Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score (range, 0 to 2400; higher scores indicate more severe symptoms) at 2 years of follow-up. Secondary outcomes included the Short Form-36 (SF-36) Physical Component Summary score (range, 0 to 100; higher scores indicate better quality of life).

RESULTS

Of the 92 patients assigned to surgery, 6 did not undergo surgery. Of the 86 patients assigned to control treatment, all received only physical and medical therapy. After 2 years, the mean (±SD) WOMAC score for the surgery group was 874±624, as compared with 897±583 for the control group (absolute difference [surgery-group score minus control-group score], –23±605; 95% confidence interval [CI], –208 to 161; P=0.22 after adjustment for baseline score and grade of severity). The SF-36 Physical Component Summary scores were 37.0±11.4 and 37.2±10.6, respectively (absolute difference, –0.2±11.1; 95% CI, –3.6 to 3.2; P=0.93). Analyses of WOMAC scores at interim visits and other secondary outcomes also failed to show superiority of surgery.

CONCLUSIONS

Arthroscopic surgery for osteoarthritis of the knee provides no additional benefit to optimized physical and medical therapy. (ClinicalTrials.gov number, NCT00158431.)

From the Fowler Kennedy Sport Medicine Clinic (A.K., T.B.B., R.B.L., J.R.G., K.R.W., S.H.G., L.M.D., P.J.F.); the Faculty of Health Sciences (T.B.B.); Robarts Clinical Trials, Robarts Research Institute (C.J.W., B.G.F., A.D.); and the Departments of Surgery (A.K., R.B.L., J.R.G., K.R.W., P.J.F.), Medicine (B.G.F., J.E.P.), and Epidemiology and Biostatistics (B.G.F., A.D.) — all at the University of Western Ontario; and St. Joseph's Health Care (J.E.P.) — in London, ON, Canada. Address reprint requests to Dr. Litchfield at the Fowler Kennedy Sport Medicine Clinic, 3M Centre, University of Western Ontario, London, ON N6A 3K7, Canada, or at rlitchf@uwo.ca; or to Dr. Feagan at Robarts Clinical Trials, Robarts Research Institute, University of Western Ontario, 100 Perth Dr., London, ON N6A 5K8, Canada, or at bfeagan@robarts.ca.

*Deceased.

This article (10.1056/NEJMoa0708333) was updated on November 10, 2009, at NEJM.org.

N Engl J Med 2008;359:1097-107.

Osteoarthritis of the knee is a degenerative disease that causes joint pain, stiffness, and decreased function.[1-3] Treatment is multidisciplinary and involves physical therapy, medication, and surgery. Arthroscopic surgery, in which an arthroscope is inserted into the knee joint, allows for lavage, a procedure that removes particulate material such as cartilage fragments and calcium crystals. It also allows for débridement, whereby articular surfaces and osteophytes can be surgically smoothed. The goal of this procedure is to reduce synovitis and eliminate mechanical interference with joint motion.

Although arthroscopic surgery has been widely used for osteoarthritis of the knee, scientific evidence to support its efficacy is lacking.[4] No benefit of surgery was shown in a large-scale, randomized, controlled trial reported in the literature.[5] However, the methods used in that study have been questioned,[6-11] and the authors' conclusion that arthroscopic surgery is ineffective for the treatment of moderate-to-severe osteoarthritis of the knee has not been generally accepted.[12-14] Accordingly, the procedure remains widely used.[15] We conducted a randomized, controlled trial to compare optimized physical and medical therapy alone with arthroscopic treatment in addition to optimized physical and medical therapy.

METHODS

PATIENTS

We conducted the trial between January 1999 and August 2007 at the Fowler Kennedy Sport Medicine Clinic, University of Western Ontario, London, Ontario, Canada. The investigators who assessed outcomes were unaware of treatment assignments. The protocol was approved by the institutional review board of the University of Western Ontario. All patients gave written informed consent.

Eligible patients were 18 years of age or older with idiopathic or secondary osteoarthritis of the knee[16,17] with grade 2, 3, or 4 radiographic severity, as defined by the modified Kellgren–Lawrence classification.[18-20] Patients were excluded if they had large meniscal tears ("bucket handle" tears), as detected by clinical examination[21,22] or, in a minority of cases, by magnetic resonance imaging. Other exclusion criteria were inflammatory or postinfectious arthritis, previous arthroscopic treatment for knee osteoarthritis, more than 5 degrees of varus or valgus deformity, previous major knee trauma, Kellgren–Lawrence grade 4 osteoarthritis in two compartments (the medial or lateral compartments of the tibiofemoral joint or the patellofemoral compartment) in persons over 60 years of age, intraarticular corticosteroid injection within the previous 3 months, a major neurologic deficit, serious medical illness (life expectancy of less than 2 years or high intraoperative risk), and pregnancy. Patients who were unable to provide informed consent or who were deemed unlikely to comply with follow-up were also excluded.

BASELINE STUDIES

Patients referred to any of seven orthopedic surgeons were assessed for eligibility. The trial coordinator and one of two surgeons independently reviewed the diagnosis of osteoarthritis. Disagreements regarding eligibility, degree of malalignment (i.e., degree of varus or valgus deformity), and the Kellgren–Lawrence grade were resolved by consensus. Baseline scores on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC),[23,24] the Short Form-36 (SF-36) Physical Component Summary,[25] the McMaster–Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR),[26,27] and the Arthritis Self-Efficacy Scale (ASES)[28] and standard-gamble[29] utility scores were obtained. An orthopedic surgeon performed a detailed examination of the knee and documented the range of motion, the presence of an effusion, and the results of meniscal and stability tests.

STUDY TREATMENT

The patients were randomly assigned, with the use of a computer-generated schedule, to receive optimized physical and medical therapy alone (control group) or to receive both optimized physical and medical therapy and arthroscopic treatment. The randomization was stratified according to surgeon and disease severity (defined according to the Kellgren–Lawrence grade). To minimize the risk of predicting the treatment assignment of the next eligible patient, randomization was performed in permuted blocks of two or four with random variation of the blocking number. Both for patients assigned to surgery and for those assigned to control treatment, the date of treatment initiation was defined as the next available date of surgery.

Arthroscopic treatment was performed within 6 weeks after randomization with the patient under general anesthesia and with the use of a tour-

niquet and a thigh holder. The orthopedic surgeon evaluated the medial, lateral, and patellofemoral joint compartments, graded articular lesions according to the Outerbridge classification,[30] irrigated the compartment with at least 1 liter of saline, and performed one or more of the following treatments: synovectomy; débridement; or excision of degenerative tears of the menisci, fragments of articular cartilage, or chondral flaps and osteophytes that prevented full extension. Abrasion or microfracture of chondral defects was not performed.

Optimized physical and medical therapy was initiated within 7 days after surgery and followed an identical program in both groups. Physical therapy was provided for 1 hour once a week for 12 consecutive weeks. The intervention was standardized and based on a review of the literature and a formal survey of university physical therapists.[31] Information regarding a home exercise program that emphasized range-of-motion and strengthening exercises was provided to all patients. Individualized exercises were recommended on the basis of the severity of osteoarthritis, the patient's age, and the patient's specific needs. Instruction was also provided regarding activities of daily living, walking, use of stairs, and methods of treatment involving cold and heat. The patients were asked to perform the exercises twice daily and once on the day of a scheduled physical-therapy session. After the patients had completed 12 weeks of supervised activity, they continued an unsupervised exercise program at home for the duration of the study. The patients received additional education from attendance at local Arthritis Society workshops, from a copy of *The Arthritis Helpbook*[32] that was provided to them, and from an educational videotape.

After undergoing randomization, the patients reviewed their medical treatment plans with an orthopedic surgeon, and the plans were optimized according to an evidence-based treatment algorithm based on published guidelines[2] that recommended stepwise use of acetaminophen and nonsteroidal antiinflammatory drugs and intraarticular injection of hyaluronic acid. Hyaluronic acid and oral glucosamine were offered to the patients.

The patients were seen in the clinic 3, 6, 12, 18, and 24 months after the initiation of treatment. At each visit, the patients were evaluated by a nurse who was unaware of the treatment assignment. To preserve blinding, each patient wore a neoprene sleeve over the knee so that the study nurse could not identify a surgical scar. Scores on the WOMAC, MACTAR, SF-36, and ASES questionnaires and standard-gamble utility scores were obtained at each visit. Medical treatment was reviewed at each visit, and treatment options were modified according to the algorithm. Records were kept of medical therapies used.

OUTCOME MEASURES

The primary outcome was the WOMAC score at 2 years after the initiation of treatment. The WOMAC is a validated, self-administered instrument specifically designed to evaluate knee and hip osteoarthritis. The WOMAC has subscales for pain, stiffness, and physical function. Total scores can range from 0 to 2400; higher scores indicate increased pain, increased stiffness, and decreased physical function.[23] Patients with moderate-to-severe osteoarthritis of the knee typically have a score of approximately 1000.[24,33] A 20% improvement (typically, a decrease of about 200 points) in the total WOMAC score was considered clinically important.[34-36] We also analyzed the three WOMAC subscales separately. The Physical Component Summary of the SF-36 was used to assess quality of life; scores can range from 0 to 100, with higher scores indicating better quality of life. The MACTAR and ASES are validated questionnaires that assess the symptoms and functional status of patients with osteoarthritis. MACTAR scores can range from 0 to 500; higher scores indicate greater disability. ASES scores can range from 10 to 100; higher scores indicate greater self-efficacy (i.e., perceived ability to cope with the consequences of arthritis). Health-related quality of life was assessed by the standard-gamble utility technique; scores can range from 0.0 (death) to 1.0 (perfect health).[29]

STATISTICAL ANALYSIS

Baseline characteristics were analyzed by descriptive statistics. For the primary analysis, the total WOMAC score at 2 years was compared between the two study groups with the use of analysis of covariance, with adjustment for the baseline score and disease severity (as measured by the Kellgren–Lawrence grade). A two-sided P value of 0.05 was considered to indicate statistical significance. Post hoc analyses of the total WOMAC score were also performed at 3, 6, 12, and 18 months. Missing

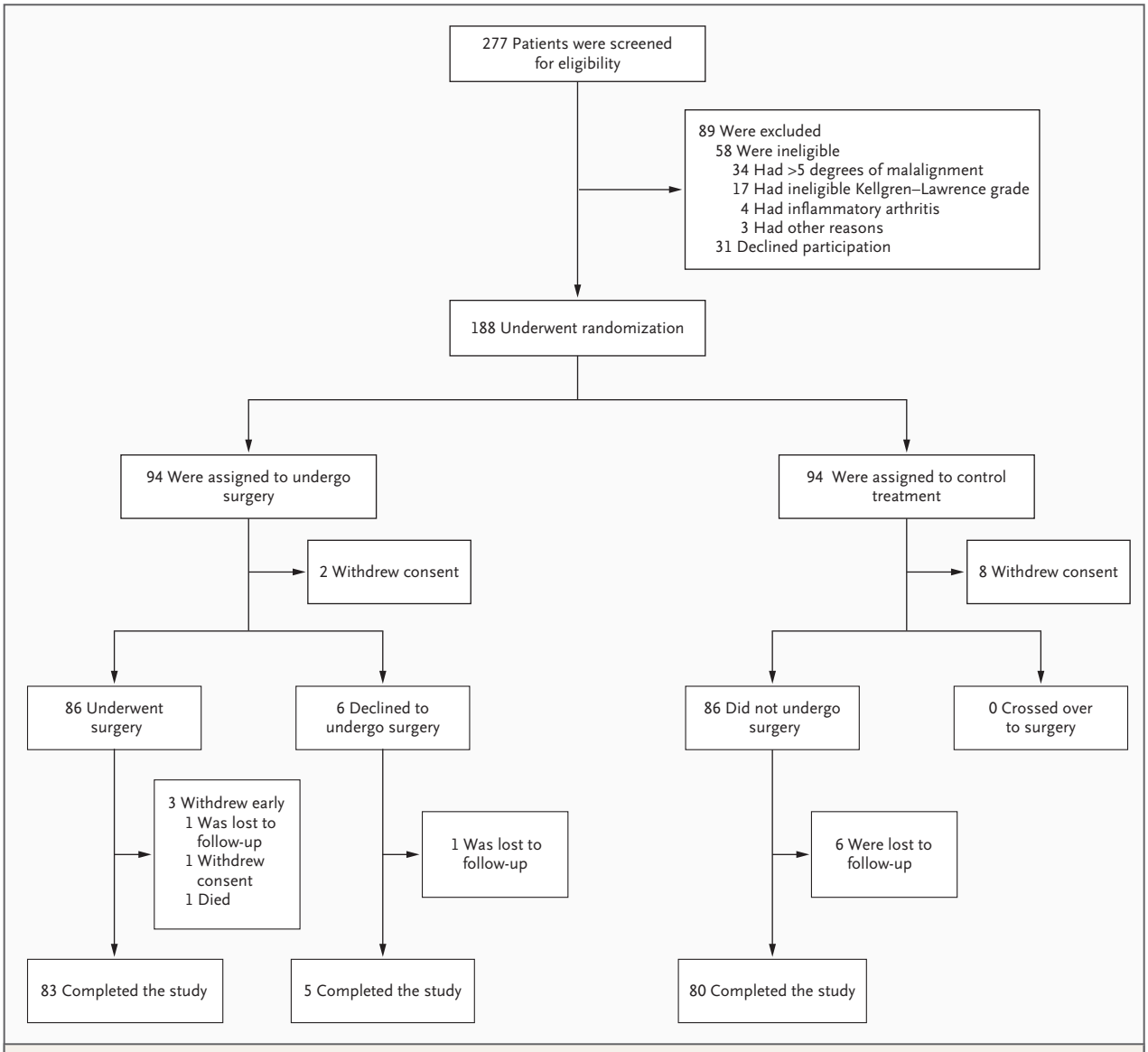


Figure 1. Enrollment of Patients and Completion of the Study.

values were not imputed. A similar approach was used to analyze the scores on the WOMAC subscales, the SF-36 Physical Component Summary, MACTAR, and ASES. Two prespecified subgroup analyses were performed. Patients with less severe disease (Kellgren–Lawrence grade 2) and patients reporting mechanical symptoms of catching, locking, or both catching and locking of the knee were hypothesized to derive greater benefit from surgery. The proportions of patients who received physical therapy and the average number of visits were compared with the use of the chi-square test and Student's t-test, respectively. The proportions of patients who received the various algorithm-specified medical therapies were compared with the use of the score-type test for simultaneous marginal homogeneity.[37] Statistical comparisons were made with the use of SAS software, version 8.2.[38] All analyses were performed according to the intention-to-treat principle.

On the basis of the results of a previous study,[33] the standard deviation of the total WOMAC score was estimated to be 452. Assignment of 186 patients to treatment would provide 80% statistical power to detect a 200-point difference between the two treatment groups, with allowance

Table 1. Baseline Characteristics of the Patients.*

Characteristic	Arthroscopic Surgery (N=92)	Control (N=86)	P Value†
Age — yr	58.6±10.2	60.6±9.9	0.19
Male sex — no. (%)	38 (41)	28 (33)	0.23
Weight — kg	91.3±17.3	84.9±17.9	0.02
Height — cm	170.4±9.7	167.6±10.2	0.07
Body-mass index‡	31.6±6.7	30.2±6.3	0.15
Duration of osteoarthritis symptoms in study knee — mo	47.1±69.4	40.1±72.6	0.52
Kellgren–Lawrence grade — no. (%)§			0.83
2	42 (46)	36 (42)	
3	45 (49)	46 (53)	
4	5 (5)	4 (5)	
Anatomical alignment — degrees¶	1.2±3.4	1.2±3.9	0.88
Symptoms of catching or locking — no. (%)	48 (52)	38 (44)	0.29
Joint effusion — no. (%)	56 (61)	53 (62)	0.92
Positive McMurray test — no. (%)‖	1 (1)	1 (1)	1.0
Pain with forced flexion — no. (%)	62 (67)	56 (65)	0.75
Tenderness at the tibiofemoral joint line — no. (%)	81 (88)	77 (90)	0.75
Magnetic resonance imaging performed — no. (%)	15 (16)	10 (12)	0.37
WOMAC**			
Total score	1187±483	1043±542	0.06
Pain dimension	239±105	214±122	0.14
Stiffness dimension	117±50	103±48	0.05
Physical function dimension	830±355	726±397	0.07
SF-36 Physical Component Summary††	33.8±7.6	33.9±8.6	0.93

* Plus–minus values are means ±SD.
† P values were calculated by the t-test for continuous variables and the chi-square test for categorical variables.
‡ The body-mass index is the weight in kilograms divided by the square of the height in meters.
§ The Kellgren–Lawrence scale evaluates the radiographic severity of osteoarthritis of the knee. Grade 0 denotes normal; grade 1 doubtful narrowing of the joint space and possible osteophyte lipping (irregular bone formation); grade 2 definite osteophytes and possible narrowing of the joint space; grade 3 multiple moderate-size osteophytes, definite narrowing of the joint space, some sclerosis, and possible deformity of bone contour; and grade 4 large osteophytes, marked narrowing of the joint space, severe sclerosis, and definite deformity of bone contour. Patients with grade 1 osteoarthritis were excluded from the trial.
¶ Anatomical alignment of the lower limb (anatomical axis angle) was assessed from anteroposterior radiographs obtained while the patient was standing. Positive values indicate valgus alignment, and negative values indicate varus alignment.
‖ A McMurray test is positive for a tear in the meniscus if a click is palpable over the medial or lateral tibiofemoral joint line during flexion and extension of the knee during varus or valgus stress.
** The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) comprises three subscales (pain, stiffness, and physical function) composed of 24 questions. Scores can range from 0 to 2400; higher scores indicate more severe disease.
†† The Short Form-36 (SF-36) is a self-administered 36-item questionnaire that assesses the concepts of physical functioning, role limitations due to physical problems, social function, bodily pain, general mental health, role limitations due to emotional problems, vitality, and general health perceptions. Scores can range from 0 to 100; higher scores indicate better quality of life. The SF-36 has become the most commonly used global health-status tool in orthopedic trials.

for up to 15% of patients whose data could not be evaluated.

A single prespecified interim analysis was performed by an external data monitoring board when one third of the patients had completed 2 years of follow-up. This analysis was based on an O'Brien–Fleming boundary[39] that specified a P value of 0.0007 to stop the trial because of su-

Table 2. Use of Medical, Physical, and Surgical Therapy in the Patients.*

Therapy	Arthroscopic Surgery (N=92)	Control (N=86)	P Value†
Medical therapy — no. (%)			0.86
Nonsteroidal antiinflammatory drugs	53 (58)	48 (56)	
Acetaminophen	53 (58)	43 (50)	
Chondroitin sulfate or glucosamine	28 (30)	25 (29)	
Hyaluronic acid injection	39 (42)	33 (38)	
Physical therapy			
Patients participating — no. (%)	88 (96)	77 (90)	0.12
No. of visits by participating patients	9.3±5.1	8.0±5.7	0.13
Use of a brace — no. (%)	3 (3)	5 (6)	0.49
Surgical therapy — no. (%)‡			
Débridement of articular cartilage	83 (97)		
Débridement or partial resection of meniscus	70 (81)		
Repair of meniscus	0		
Excision of osteophytes	8 (9)		
Removal of loose bodies	12 (14)		

* Plus–minus values are means ±SD.
† P values for continuous variables were calculated by the t-test. The chi-square or the two-tailed Fisher's exact test was used for categorical variables, except for medical therapies, for which the score-type test for simultaneous marginal homogeneity was used.
‡ The percentages are based on 86 patients rather than 92 because 6 patients who were assigned to surgery declined the procedure.

periority of treatment and a P value of 0.984 to stop the trial because of futility of treatment. Neither criterion was met.

RESULTS

Figure 1 shows the disposition of the study participants. Between January 1999 and August 2005, 277 patients were assessed for eligibility. Fifty-eight patients were not eligible and 31 declined participation, resulting in a total of 188 patients who underwent randomization. Ninety-four patients were assigned to receive arthroscopic surgery and optimized physical and medical therapy and 94 to receive physical and medical therapy alone. Ten patients (two in the surgery group and eight in the control group) withdrew consent after randomization. Six patients assigned to surgery elected not to have the procedure; data from these patients were analyzed, according to the intention-to-treat principle, with data from the surgery group. Although the baseline characteristics of the groups were similar (Table 1), patients assigned to surgery had slightly higher total WOMAC scores.

Figure 2 (facing page). Total WOMAC Scores over Time According to Treatment Group.

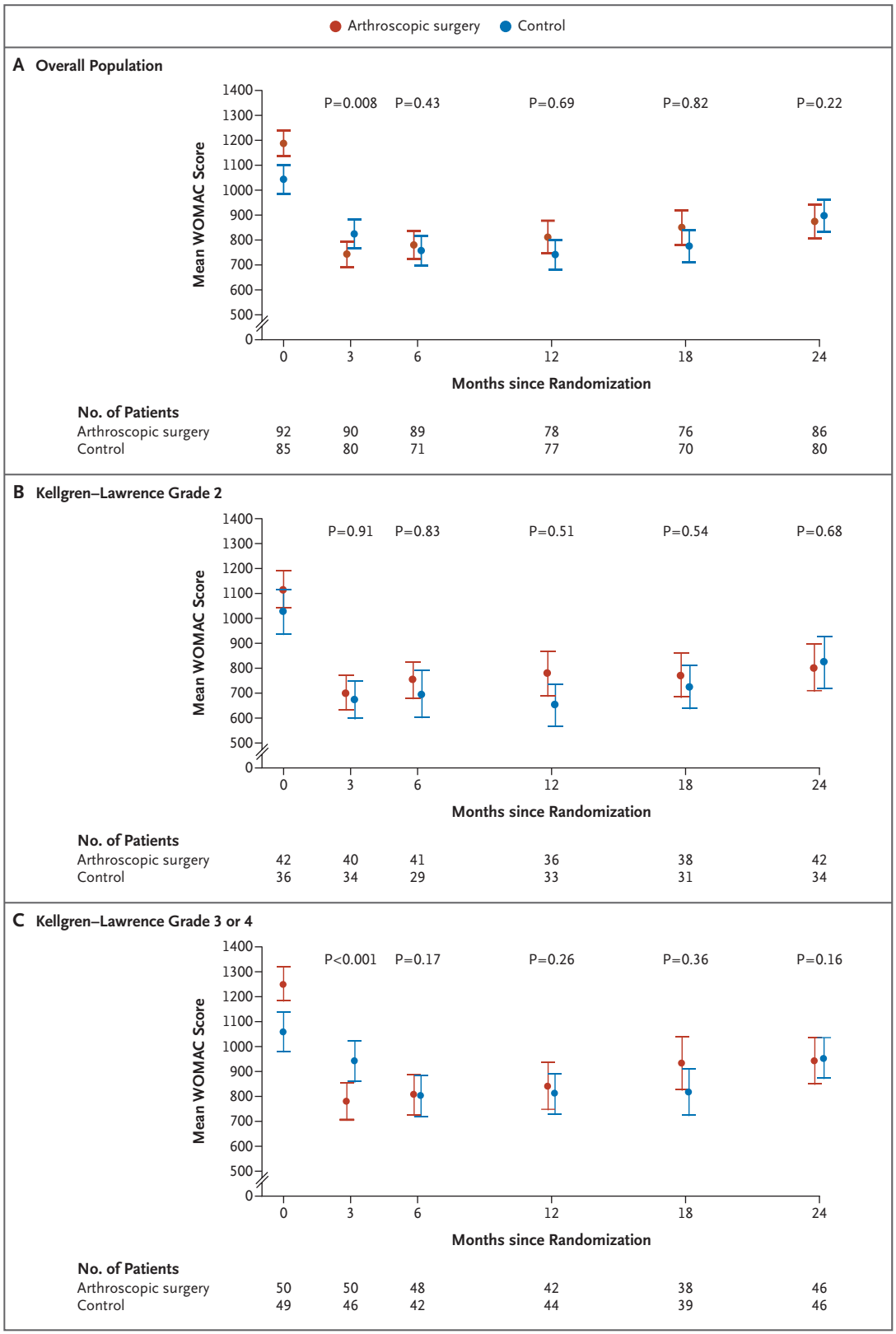
Scores are shown for all patients (Panel A), those with less severe disease (Kellgren–Lawrence grade 2) (Panel B), and those with more severe disease (Kellgren–Lawrence grade 3 or 4) (Panel C). P values for the differences in 24-month scores were generated by analysis of covariance with adjustment for baseline score (Panels A, B, and C) and disease severity (Panel A). Error bars indicate the standard error. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) comprises three subscales (pain, stiffness, and physical function) composed of 24 questions. Total scores can range from 0 to 2400; higher scores indicate more severe disease.

STUDY TREATMENT

The use of physical and medical therapy was similar in the two treatment groups. The majority of patients assigned to arthroscopic surgery underwent débridement of articular cartilage or meniscal lesions (Table 2).

PRIMARY OUTCOME MEASURE

Figure 2A shows the changes in the mean total WOMAC scores. At 3 months, scores in the surgery



group had improved to a greater extent than those in the control group, but there were no significant differences between the groups at any visits thereafter. For patients assigned to surgery, the mean (±SD) total WOMAC score at 24 months was 874±624, as compared with 897±583 in the control group (absolute difference [surgery-group score minus control-group score], −23±605; 95% confidence interval [CI], −208 to 161; P=0.22 after adjustment for baseline score and radiographic grade of disease severity). A similar analysis performed in patients with less severe disease (Kellgren–Lawrence grade 2) at baseline also found no significant difference between the treatment groups (Fig. 2B). Likewise, no benefit was conferred by surgery among the subgroup of patients with mechanical symptoms of catching or locking. A post hoc analysis of patients with more severe radiographic disease (Kellgren–Lawrence grade 3 or 4) also found no benefit of surgery (Fig. 2C). We repeated these analyses on the basis of treatment actually received by including the data from the six patients assigned to surgery who elected not to undergo the procedure with the data from the patients in the control group. The results of these analyses were consistent with those of the intention-to-treat analyses.

SECONDARY OUTCOME MEASURES

No significant differences were observed between the treatment groups for any of the secondary outcome measures (Table 3). Specifically, patients assigned to arthroscopic surgery were no more likely to improve with respect to physical function, pain, or health-related quality of life than were those assigned to the control group. After 2 years, the SF-36 Physical Component Summary scores were 37.0±11.4 and 37.2±10.6, respectively (absolute difference, −0.2±11.1; 95% CI, −3.6 to 3.2; P=0.93).

DISCUSSION

This study failed to show a benefit of arthroscopic surgery for the treatment of osteoarthritis of the knee. At the end of 2 years, patients assigned to arthroscopic treatment in addition to optimized physical and medical therapy had no greater improvement in WOMAC scores than did those who received only physical and medical therapy. Patients assigned to surgery did have a greater improvement in WOMAC scores within the first 3 months; however, this transient benefit was anticipated, since sham surgery is associated with a large, short-term placebo effect.[5] WOMAC scores at all other time points did not significantly differ between the groups. In addition to WOMAC scores, a broad range of validated patient-reported outcomes was assessed at multiple time points. None of these instruments identified a benefit of arthroscopic treatment.

These negative results are in agreement with the previously published findings of Moseley and colleagues.[5] That trial, which was conducted by a single surgeon at a Veterans Affairs hospital, was methodologically rigorous, since use of a sham-operation control allowed concealment of the treatment assignment. Nevertheless, several methodologic issues were raised that we believe are addressed in the current study. For example, the outcome measure in the study by Moseley et al., the Knee Specific Pain Scale,[5] was not validated.[9] We used the WOMAC score, a validated instrument that has been widely used in osteoarthritis research, as the primary measure of efficacy. Patients with substantial malalignment (varus or valgus deformity) and those with advanced disease, who might have a poorer response to surgical intervention,[7,11] were included in the earlier trial; we excluded patients with more than 5 degrees of malalignment and stratified the randomization according to both surgeon and Kellgren–Lawrence grade of radiographic severity. Moseley and colleagues evaluated mostly older men who were treated in a Veterans Affairs Medical Center.[6,7] In contrast, our study evaluated a more typical population of both men and women who were treated in a university hospital. Seven experienced arthroscopists performed lavage, débridement, or both at their discretion. Thus, we believe that our results are highly generalizable to usual orthopedic practice.

The results of the present trial, along with the results of Moseley et al., call into question the widespread use of arthroscopic treatment for osteoarthritis of the knee. Although some may argue that treatment is beneficial for patients with mechanical symptoms of catching or locking or those with early disease, prespecified subgroup analyses also failed to show efficacy in this population of patients. However, patients suspected of having large meniscal ("bucket handle") tears, in whom arthroscopic surgery is considered an effective intervention, were excluded from this study.

Our study had limitations. Bias is possible be-

Table 3. Secondary Outcome Measures in the Patients.*

Measure	Baseline		Mo 3		Mo 6		Mo 12		Mo 18		Mo 24		P Value†
	Surgery (N=92)	Control (N=86)	Surgery (N=90)	Control (N=80)	Surgery (N=90)	Control (N=73)	Surgery (N=80)	Control (N=77)	Surgery (N=78)	Control (N=70)	Surgery (N=88)	Control (N=80)	
WOMAC‡													
Total score	1187±483	1043±542	743±481	824±520	780±518	757±510	811±576	741±514	850±609	775±532	874±624	897±583	0.22
Pain	239±105	214±122	141±109	172±124	143±113	155±118	155±125	147±116	179±140	158±115	168±134	185±132	0.14
Stiffness	117±50	103±48	80±54	84±53	86±53	82±47	85±56	81±51	94±59	80±49	93±60	88±51	1.00
Physical function	830±355	726±397	522±341	568±369	551±382	520±368	570±417	513±370	578±427	537±385	612±448	623±439	0.26
SF-36 Physical Component Summary§	33.8±7.6	33.9±8.6	38.7±9.0	37.7±10.2	38.7±9.3	38.1±10.2	38.3±10.7	37.7±10.0	37.7±11.9	38.4±10.4	37.0±11.4	37.2±10.6	0.93
ASES¶													
Pain	69.0±15.6	65.4±17.0	73.9±15.8	68.6±17.0	71.5±16.9	67.9±17.0	70.5±20.0	69.5±16.8	69.8±18.9	66.6±19.0	68.8±18.5	63.8±19.8	0.23
Function	77.4±16.8	79.5±17.2	80.7±18.2	81.9±19.6	83.8±14.7	83.2±16.1	81.4±19.1	84.4±15.8	82.0±18.5	83.2±18.5	83.5±17.0	81.9±18.4	0.20
Other symptoms	72.1±17.2	71.6±17.0	77.4±17.0	74.7±15.8	78.3±15.8	74.3±15.2	78.4±18.4	76.1±16.5	76.3±16.3	73.1±18.8	78.8±16.3	73.4±18.2	0.05
MACTAR‖	348±85	320±99	257±108	249±109	234±118	246±115	232±128	225±117	251±141	221±115	238±146	244±133	0.58
Standard-gamble utility score**	0.79±0.23	0.81±0.20	0.81±0.21	0.80±0.22	0.84±0.20	0.81±0.22	0.82±0.21	0.86±0.16	—	—	0.87±0.18	0.86±0.16	0.87

* Plus–minus values are means ±SD. Not all patients attended each follow-up visit.

† P values for the difference between the surgery and control groups at 24 months were calculated by analysis of covariance, with adjustment for baseline score and disease severity.

‡ The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) comprises three subscales (pain, stiffness, and physical function) composed of 24 questions. Scores can range from 0 to 2400; higher scores indicate more severe disease.

§ The Short Form-36 (SF-36) is a self-administered questionnaire that assesses the concepts of physical functioning, role limitations due to physical problems, social function, bodily pain, general mental health, role limitations due to emotional problems, vitality, and general health perceptions. Scores can range from 0 to 100; higher scores indicate better quality of life. The SF-36 has become the most commonly used global health-status tool in orthopedic trials.

¶ The Arthritis Self-Efficacy Scale (ASES), a validated measure of perceived ability to cope with the consequences of arthritis, consists of 20 questions that evaluate pain, functional status, and other arthritis-related symptoms. Scores on the three subscales can range from 10 to 100; higher scores indicate greater self-efficacy.

‖ The McMaster–Toronto Arthritis Patient Preference Disability Questionnaire (MACTAR) is a validated patient-preference questionnaire that measures changes in impaired activities due to arthritis and their relative importance to the patient. The patient identifies the five most significant activities and quantifies the degree of trouble or distress due to the arthritis on a visual-analogue scale. Scores can range from 0 to 500; higher scores indicate greater disability.

** Utility scores were generated by the standard-gamble method. Scores can range from 0.0 (death) to 1.0 (perfect health).

The New England Journal of Medicine

cause of the lack of a sham-surgery control. However, such bias would be expected to favor surgery and would not be expected to explain the present results. The observation that the two study groups had very similar exposures to drug treatment also suggests that bias, at least in the form of differential prescription of cointerventions, did not influence the results.

A second limitation is that only 68% of the patients who were evaluated for participation were deemed eligible and ultimately assigned to treatment. However, the majority of the excluded patients had substantial malalignment (38%) or declined participation (35%). The stringent inclusion and exclusion criteria ensured appropriate patient selection and optimized the chance of identifying a benefit of surgery. Thus, we do not believe that the participants in this trial were less likely to have a response to arthroscopic therapy than those treated in the community.

In summary, the results of this randomized, controlled trial show that arthroscopic surgery provides no additional benefit to optimized physical and medical therapy for the treatment of osteoarthritis of the knee.

Supported by the Canadian Institutes of Health Research.

No potential conflict of interest relevant to this article was reported.

This article is dedicated to the memory of Dr. Alexandra Kirkley.

We thank Beverley Jasevicius for assistance in the preparation of the manuscript.

APPENDIX

The following persons participated in the study: surgeons — A. Amendola, P.J. Fowler, J.R. Giffin, A. Kirkley, R.B. Litchfield, R. McCalden, K.R. Willits; members of the Data Coordinating Center, Robarts Clinical Trials, Robarts Research Institute — B. Sarazin, B. Bergman, H. Sun, G.Y. Zou, L. Liddiard, B. Annunziello, L. Smith, T. Clayton, M. Brine, N. Ng; data monitoring committee — D. Whalen, A. Donner.

REFERENCES

1. Felson DT. Osteoarthritis of the knee. N Engl J Med 2006;354:841-8. [Erratum, N Engl J Med 2006;354:2520.]
2. American College of Rheumatology Subcommittee on Osteoarthritis Guidelines. Recommendations for the medical management of osteoarthritis of the hip and knee: 2000 update. Arthritis Rheum 2000;43:1905-15.
3. Hochberg MC, Altman RD, Brandt KD, et al. Guidelines for the medical management of osteoarthritis. II. Osteoarthritis of the knee. Arthritis Rheum 1995;38:1541-6.
4. Felson DT, Buckwalter J. Débridement and lavage for osteoarthritis of the knee. N Engl J Med 2002;347:132-3.
5. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002;347:81-8.
6. Gillespie WJ. Arthroscopic surgery was not effective for relieving pain or improving function in osteoarthritis of the knee. ACP J Club 2003;138:49.
7. Jackson RW. Arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002;347:1717.
8. Morse LJ. Arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002;347:1718.
9. Chambers KG, Schulzer M. Arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002;347:1718.
10. Ellis TJ, Crawford D. Arthroscopic surgery for arthritis of the knee. Curr Womens Health Rep 2003;3:63-4.
11. Ewing W, Ewing JW. Arthroscopic surgery for osteoarthritis of the knee. N Engl J Med 2002;347:1717.
12. Kelly MA. Role of arthroscopic debridement in the arthritic knee. J Arthroplasty 2006;21:Suppl 1:9-10.
13. Poehling GG. Degenerative arthritis arthroscopy and research. Arthroscopy 2002;18:683-7.
14. Siparsky P, Ryzewicz M, Peterson B, Bartz R. Arthroscopic treatment of osteoarthritis of the knee: are there any evidence-based indications? Clin Orthop Relat Res 2007;455:107-12.
15. Hawker G, Guan J, Judge A, Dieppe P. Knee arthroscopy in England and Ontario: patterns of use, changes over time, and relationship to total knee joint replacement. J Bone Joint Surg Am (in press).
16. Altman R, Asch E, Bloch D, et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 1986;29:1039-49.
17. Altman RD, Fries JF, Bloch DA, et al. Radiographic assessment of progression in osteoarthritis. Arthritis Rheum 1987;30:1214-25.
18. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494-502.
19. Blackburn WD Jr, Bernreuter WK, Rominger M, Loose LL. Arthroscopic evaluation of knee articular cartilage: a comparison with plain radiographs and magnetic resonance imaging. J Rheumatol 1994;21:675-9.
20. The epidemiology of chronic rheumatism. Vol. 2. Atlas of standard radiographs of arthritis. Oxford, England: Blackwell Scientific, 1963.
21. Bhattacharyya T, Gale D, Dewire P, et al. The clinical importance of meniscal tears demonstrated by magnetic resonance imaging in ostoearthritis of the knee. J Bone Joint Surg Am 2003;85-A:4-9.
22. Ryzewicz M, Peterson B, Siparsky PN, Bartz RL. The diagnosis of meniscus tears: the role of MRI and clinical examination. Clin Orthop Relat Res 2007;455:123-33.
23. Bellamy N. WOMAC osteoarthritis index: a user's guide. London, ON, Canada: University of Western Ontario, 1995.
24. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833-40.
25. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
26. Verhoeven AC, Boers M, van der Liden S. Validity of the MACTAR questionnaire as a functional index in a rheumatoid arthritis clinical trial. J Rheumatol 2000;27:2801-9.
27. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire — an individualized functional priority approach for assessing improvement in physical disability in

clinical trials in rheumatoid arthritis. J Rheumatol 1987;14:446-51.
28. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum 1989;32:37-44.
29. Torrance GW. Utility approach to measuring health-related quality of life. J Chronic Dis 1987;40:593-600.
30. Outerbridge RE. The etiology of chondromalacia patellae. J Bone Joint Surg Br 1961;43-B:752-7.
31. van Baar ME, Assendelft WJ, Dekker J, Oostendorp RA, Bijlsma JW. Effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review of randomized clinical trials. Arthritis Rheum 1999;42:1361-9.
32. Lorig K, Fries JF. The arthritis helpbook: a tested self-management program for coping with arthritis and fibromyalgia. 5th ed. New York: Perseus Books, 2000.
33. Kirkley A, Webster-Bogaert S, Litchfield R, et al. The effect of bracing on varus gonarthrosis. J Bone Joint Surg Am 1999;81:539-48.
34. Angst F, Aeschlimann A, Michel BA, Stucki G. Minimal clinically important rehabilitation effects in patients with osteoarthritis of the lower extremities. J Rheumatol 2002;29:131-8.
35. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities Osteoarthritis Index questionnnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000;27:2635-41.
36. Goldsmith CH, Boers M, Bombardier C, Tugwell P. Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. J Rheumatol 1993;20:561-5.
37. Agresti A, Klingenberg B. Multivariate test comparing binomial probabilities, with application to safety studies for drugs. Appl Stat 2005;54:691-706.
38. SAS system software, version 8.2, TS2 M0. Cary, NC: SAS Institute, 2001.
39. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979;35:549-56.