Tibiofemoral Osteoarthritis After Surgical or Nonsurgical Treatment of Anterior Cruciate Ligament Rupture: A Systematic Review

Kyle P. Harris  
*Bucks County Community College*

Jeffrey B. Driban  
*Tufts Medical Center*

Michael R. Sitler  
*Bucks County Community College*

Nicole M. Cattano  
*West Chester University of Pennsylvania, ncattano@wcupa.edu*

Easwaran Balasubramanian  
*Temple University*

Follow this and additional works at: [http://digitalcommons.wcupa.edu/spomed_facpub](http://digitalcommons.wcupa.edu/spomed_facpub)  
Part of the [Sports Sciences Commons](http://digitalcommons.wcupa.edu/spomed_facpub)

**Recommended Citation**

[http://dx.doi.org/10.4085/1062-6050-49.3.89](http://dx.doi.org/10.4085/1062-6050-49.3.89)
Tibiofemoral Osteoarthritis After Surgical or Nonsurgical Treatment of Anterior Cruciate Ligament Rupture: A Systematic Review

Kyle P. Harris, MS, ATC*; Jeffrey B. Driban, PhD, ATC, CSCS†; Michael R. Sitler, EdD, ATC, FNATA*; Nicole M. Cattano, MPH, ATC‡; Easwaran Balasubramanian, MD§

*Department of Health, Physical Education and Nursing, Bucks County Community College, Newtown, PA; †Division of Rheumatology, Tufts Medical Center, Boston, MA; ‡Department of Sports Medicine, West Chester University, PA; §Department of Clinical Orthopedic Surgery, Temple University, Philadelphia, PA

Objective: To determine if surgical or nonsurgical treatment of anterior cruciate ligament rupture affects the prevalence of posttraumatic tibiofemoral osteoarthritis (OA).

Data Sources: Studies published between 1983 and April 2012 were identified via EBSCOhost and OVID. Reference lists were then screened in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

Study Selection: Studies were included if (a) treatment outcomes focused on a direct comparison of surgical versus nonsurgical treatment of anterior cruciate ligament rupture, (b) the prevalence of tibiofemoral OA was reported, and (c) they were written in English. Studies were excluded if (a) the included patients were treated with cast immobilization after surgery, (b) the mean follow-up was less than 10 years, or (c) the patients underwent anterior cruciate ligament revision surgery.

Data Extraction: Two independent investigators reviewed the included articles using the Newcastle-Ottawa Scale. Frequency of OA, surgical procedure, nonsurgical treatments, and participant characteristics were extracted and summarized. We calculated prevalence (%) and 95% confidence intervals for treatment groups for each individual study and overall. We developed 2 × 2 contingency tables to assess the association between treatment groups (exposed had surgery, referent was nonsurgical treatment) and the prevalence of OA.

Data Synthesis: Four retrospective studies were identified (140 surgical patients, 240 nonsurgical patients). The mean Newcastle-Ottawa Scale score was 5 (range = 4–6 [of 10] points). Average length of follow-up was 11.8 years (range = 10–14 years). The prevalence of OA for surgically treated patients ranged from 32.6% to 51.2% (overall = 41.4%, 95% confidence interval = 35.0%, 48.1%) and for nonsurgical patients ranged from 24.5% to 42.3% (overall = 30.9%, 95% confidence interval = 24.4%, 38.3%).

Conclusions: Although OA prevalence was higher in the surgical treatment group at a mean follow-up of 11.8 years, no definitive evidence supports surgical or nonsurgical treatment after anterior cruciate ligament injury to prevent posttraumatic OA. Current studies have been limited by small sample sizes, low methodologic quality, and a lack of data regarding confounding factors.

Key Words: lower extremity, knee injuries, prevalence

Key Points
- This is the first systematic review to directly compare surgical and nonsurgical treatment of anterior cruciate ligament ruptures.
- No definitive evidence supports surgical or nonsurgical treatment after anterior cruciate ligament injury to prevent posttraumatic osteoarthritis.
- Research to date has been limited by small samples, low methodologic quality, and a lack of data on confounding factors.

Anterior cruciate ligament (ACL) reconstruction is often the treatment recommended after ACL rupture in a physically active individual. The intended outcome of the surgery is to restore knee anatomy and biomechanics to a functional level, thereby reducing shear and torsional stresses on the menisci and articular surfaces and permitting a return to previous physical activities. After ACL reconstruction surgery, the short-term functional results appear favorable, however, over the long term, at least 28% to 87% of these patients develop posttraumatic tibiofemoral osteoarthritis (OA).

Some patients opt for nonsurgical treatment of their ACL injuries. Patients recommended for this type of treatment typically have sufficient dynamic knee stability for their desired level of function and no secondary joint injury (eg, meniscal tear, collateral ligament sprain). For patients treated nonsurgically, the reported rates of OA range from 11% to 73%. Although these patients are believed to have less disruption of lower
extremity biomechanics after ACL rupture, it remains unclear if the likelihood of OA differs with surgical or nonsurgical treatment of the knee.15,35,40,44,48,49

No systematic review has been reported to date on the prevalence of OA in patients with ACL ruptures treated surgically versus nonsurgically. Our purpose was to conduct a systematic review to determine if OA prevalence differed between the treatments. Studies used in this systematic review focused on a direct head-to-head comparison of surgical reconstruction and nonsurgical treatment. The advantage of a head-to-head approach was that the studies used the same criteria, such as the radiographic threshold for determining OA and the patient’s previous level of function. Our intent was to inform evidence-based clinical care in the treatment of patients with ACL ruptures.

METHODS

Inclusion and Exclusion Criteria

Studies were included if (a) treatment outcomes focused on a direct comparison of surgical versus nonsurgical treatment of ACL ruptures, (b) the prevalence of tibiofemoral OA was reported, and (c) they were written in English. Studies were excluded if (a) patients were placed in casts after surgery, (b) the mean follow-up was less than 10 years, or (c) patients underwent ACL revision surgery. Before 1983, the standard postsurgical care was straight-leg casting for 8 to 12 weeks. More contemporary postsurgical treatment calls for mobilization and rehabilitation beginning immediately after surgery. Because the long-term outcomes may differ between these postsurgical treatment protocols, we excluded articles published before 1983. Randomized control trials were initially considered for inclusion; however, none met the eligibility criteria for inclusion. If the reviewer (K.P.H.) was unsure that a study met all the necessary criteria, the study was reviewed by the other authors, and a consensus was reached.

Search Strategy

We conducted a comprehensive literature search from 1983 through April 2012 with the assistance of an experienced reference librarian. Databases searched with EBSCOhost were Academic Search Premier, CAB abstracts, CINAHL, Education Research Complete, Education Resources Information Center, MEDLINE, SPORTDiscus with full text, and Research Starters-Education. Databases searched with OVID were Cochrane Database of Systematic Reviews, ACP Journal Club (1991 to April 2012), Database of Abstracts of Reviews of Effects (second quarter of 2012), Cochrane Library Central Register of Controlled Trials (second quarter of 2012), Cochrane Library Methodology Register, Cochrane Library Health Technology Assessment, Cochrane Library Economic Evaluation Database, Journals@Ovid, Global Health, and Ovid MEDLINE. Key words used in the database searches were ACL or anterior cruciate ligament and osteoarthritis or osteoarthrosis or degenerative joint disease or arthritis or coxarthrosis or gonarthrosis and reconstruct* or repair or surgery or replacement and meniscus or menisci or tear or torn or injury or injuries or injured. Although we focused on tibiofemoral OA, we did not specify the type of OA for the literature search to avoid eliminating studies of both tibiofemoral and patellofemoral OA.

Study Selection

The primary search yielded 799 studies, and the lead author (K.P.H.) screened the titles, key words, dates of publication, and abstracts. A total of 759 articles were eliminated because their titles or abstracts indicated that the studies either did not meet the inclusion criteria or they met the exclusion criteria. We then obtained the full text of the 40 remaining articles and further screened them for all inclusion and exclusion criteria. The reference lists of all 40 full-text articles were searched manually to identify any additional articles not located through the electronic database search process; no additional articles were cited. A total of 11 articles were provisionally identified as meeting the inclusion criteria. For 2 studies, it was unclear if patients were placed in casts after surgery. Therefore, we contacted the first author of each and confirmed that patients were treated with casts, so the studies were excluded. Three additional studies were later excluded during the data-extraction process because prevalence data were not reported. Another study was excluded due to a mixed study design (cohort and cross-sectional study design) and the inclusion of patients who did not improve with nonsurgical treatment. One additional study was later eliminated for not meeting the inclusion criteria for an adequate length of follow-up. The study-elimination process is shown in the Figure. The final 4 studies are presented in Table 1.

Assessment of Study Quality

Two independent reviewers (K.P.H., N.M.C.) assessed the quality of all included studies using the Newcastle-Ottawa Scale (NOS), which we modified for use in this systematic review (Appendix). Although not developed specifically for OA research, the NOS has been recommended for the qualitative evaluation of observational studies because it is easy to use and includes specific items based on study design (eg, case control, cohort). A recent analysis of several quality-assessment tools also deemed the NOS appropriate for its intended use in this study. We transferred the NOS to an electronic form for more efficient assessment and operationally defined the NOS criteria to make them applicable to the study population. Specifically, questions that assessed the representativeness of the cohorts (items 1 and 2) were modified to meet the objective of our systematic review. The question regarding the representativeness of the ACL population (item 1) was changed from “representativeness of the community of selection” to “representativeness of the general ACL-deficient population.” The question assessing the selection of the nonexposed cohort (item 2) was changed from “drawn from the same community as the exposed cohort” to “drawn from the general ACL-deficient population.” The question assessing the ascertainment of exposure of the exposed cohort (item 3) was defined as the “surgical cohort.” To assess the exposure of the nonsurgical cohort, a fifth item was added. Items used to assess the comparability of cohorts were defined as follows: (a) Did the study control for secondary injuries (ie, meniscal, ligamentous injuries other than ACL injury)? and b) Did the
study control for the body mass index (BMI) of participants at either baseline or follow-up? This indicated that the authors attempted to control for important, known influences on the development of OA. Of the 3 items designated to assess the outcomes of each study, 2 items were further defined. The adequate follow-up time for the outcome (prevalence of OA) to occur (item 2) was defined as 10 or more years, and the adequacy of follow-up cohorts (item 3) was defined as greater than 80%. The modified total possible score was 10 points, as opposed to 9 in the original instrument.

To standardize the way in which items were scored by the reviewers, we established criteria for specific questions on the NOS. To score the representativeness of the surgical group (item 1), the following criteria were assessed: (a) Were patients randomly or consecutively chosen (reduced risk of selection bias)? (b) Were patients between 15 and 55 years of age at the time of injury? (c) Was the cohort mean BMI less than 30? and (d) Was surgery performed within 8 weeks of injury? Studies that fulfilled all 4 criteria were considered representative of the general population with ACL ruptures. Studies adhering to 3 of the 4 criteria were considered to be somewhat representative. Investigations adhering to 2 or fewer of the 4 criteria were not considered representative of the general population with ACL ruptures. To score the representativeness of the nonsurgical group (item 2), we used 3 criteria: (a) Were patients randomly or consecutively chosen? (b) Were patients between 15 and 55 years of age at the time of injury? and (c) Was the cohort mean BMI less than 30? Studies that fulfilled all 3 criteria were considered representative of the general population with ACL ruptures.
Those adhering to 2 or fewer of the 3 criteria were not considered representative of the general population with ACL ruptures.

Before rating the quality of the 4 included studies, we pilot tested the rating procedures using 3 sample articles that were not included in the analysis. Studies were randomized by a blinded, independent investigator (M.R.S.). Each article was then read and scored independently by 2 raters. Consensus scores were determined for each article. If the scores of rater 1 (K.P.H.) and rater 2 (N.M.C.) agreed, then that score was used as the consensus score. If the scores of raters 1 and 2 differed by 1 point, the raters discussed and agreed on a consensus score. If the scores of raters 1 and 2 differed by 2 or more points, a third rater (J.B.D.) reviewed the article, and a final score was used as the consensus score. If the scores of raters 1 and 2 differed by 1 point, the raters discussed and agreed on a consensus score. If the scores of raters 1 and 2 differed by 2 or more points, a third rater (J.B.D.) reviewed the article, and a final score was used as the consensus score.

We created a spreadsheet for data extraction. The following information was extracted from each of the studies by the primary investigator (K.P.H): (a) publication information: first author’s name, journal, and year of publication; (b) study methods: study design, report of meniscal injury, and NOS score; (c) OA outcomes: definition of OA and time of follow-up; (d) patient descriptors: source of exposed (surgical) and nonexposed (nonsurgical) cohorts, matching variables (eg, age, sex, Tegner activity score), sport, level of participation, percentage lost to follow-up, sample size, age, weight, height, BMI; and (e) outcomes measures: prevalence of OA, frequency of OA (#/100), and adjustment variables (eg, meniscal injury, age, BMI).

For each study, the number of patients with OA at follow-up, stratified by treatment group (surgical, nonsurgical), was populated in a 2 × 2 contingency table. Prevalence (%) of OA was calculated using the following formula: number with OA in a treatment group/total in the treatment group at baseline. Prevalence ratios (PRs) were calculated using the formula (# surgical and OA/all surgical)/(# nonsurgical and OA/all nonsurgical). The nonsurgical treatment group was considered the referent group. We calculated 95% confidence intervals (CIs) for proportions and PRs using standard methods and prevalence proportions and PRs for each individual study and for all studies combined (overall prevalence).

### RESULTS

All 4 reviewed studies used a retrospective cohort design. A total of 380 patients (260 men, 120 women), 140 (37%) of whom were treated surgically, were included in the 4 studies. The mean patient ages reported at follow-up and time of trauma were 37.9 and 24.9 years, respectively. In all 4 studies, bone–patellar tendon–bone autograft was the primary surgical method. The reported postsurgical follow-up for each study ranged from 10 to 14 years (mean = 11.8 years).

### Methodologic Quality

No study fulfilled all of the NOS criteria. The highest score recorded was 6 (of 10 possible points), and the lowest was 4. The mean score for all studies was 5 points. All studies included patients with meniscal injury; authors of only 1 study adjusted statistically for this using a logistic regression model. Two groups used matching variables (age, sex, and Tegner activity score) in allocating the exposed and nonexposed cohorts. No authors reported

---

**Table 1. Study Design, Participant Characteristics, Exposures, and Outcomes for Studies (n = 4) Included in the Systematic Review Extended on Next Page**

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Cohort Study Design</th>
<th>Participant Characteristics</th>
<th>Surgical (n)</th>
<th>Nonsurgical (n)</th>
<th>Exposure Definition: Treatment Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lohmander et al (2004)</td>
<td>Retrospective</td>
<td>Age range, 26–40 y, women only, Sweden, league soccer: identified through insurance company archive</td>
<td>Various reconstruction methods (62% autologous patellar tendon graft), n = 41 with radiographic data</td>
<td>No explanation of nonoperative treatment, n = 26 with radiographic data</td>
<td></td>
</tr>
<tr>
<td>Kessler et al (2008)</td>
<td>Retrospective</td>
<td>Age range = 12.5–54 y, males = 62%, females = 38%, participants matched age, sex, body mass index, Tegner score, follow-up</td>
<td>Bone–patellar tendon–bone reconstruction, procedures performed by various practitioners, radiographic data: n = 60</td>
<td>Standardized, nonoperative treatment program, carried out by various physiotherapists, radiographic data: n = 49</td>
<td></td>
</tr>
<tr>
<td>von Porat et al (2004)</td>
<td>Retrospective</td>
<td>Age range = 30–56 y, men only, Swedish league soccer players identified through insurance company archive</td>
<td>Various surgical methods (most common: patellar tendon graft)</td>
<td>No explanation of nonoperative treatment</td>
<td></td>
</tr>
<tr>
<td>Meuffels et al (2009)</td>
<td>Retrospective</td>
<td>Mean age = 37.7 y, males = 76%, females = 24%, Netherlands, participants matched by age, sex, Tegner score</td>
<td>Bone–patellar tendon–bone reconstruction, performed by 1 of 2 surgeons, radiographic data: n = 50</td>
<td>Rehabilitation program: edema reduction, range of motion, hamstrings and quadriceps strengthening exercises</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: KL, Kellgren-Lawrence Scale.

*Maximum possible score for modified instrument = 10.*
inclusion and exclusion criteria for patients treated nonsurgically or fulfilled the requirement of "representative of the general ACL-deficient population" for either the exposed or nonexposed cohort. All studies\textsuperscript{23,25,32,41} ascertained exposure of the surgically treated cohort through surgical records or a structured interview. No group reported the absence of OA (eg, on radiographs) at the start of the study, mitigating the ability to determine the incidence of OA.

**Prevalence of Tibiofemoral OA**

Overall, OA prevalence for the 4 studies\textsuperscript{23,25,32,41} ranged from 24.5\% to 51.2\%. The OA prevalence for surgically treated patients ranged from 32.6\% to 51.2\% (overall = 41.4\%, 95\% CI = 35.0, 48.1) and for nonsurgical patients ranged from 24.5\% to 42.3\% (overall =30.9\%, 95\% CI = 24.4, 38.3). Because all studies included patients with compromised menisci, we did not report prevalence data comparing isolated ACL ruptures with nonsolated ACL ruptures.

**Prevalence Ratios of Tibiofemoral OA**

Across the 4 studies, the PRs of tibiofemoral OA among surgical patients compared with nonsurgical controls ranged from 1.01 to 1.84. The surgical group had a 34\% higher prevalence of tibiofemoral OA (PR = 1.34; 95\% CI = 1.01, 1.77) compared with the nonsurgical group (Table 2).

**Radiologic Classification System**

To quantify radiographic OA, authors of all 4 included studies\textsuperscript{23,25,32,41} used the Kellgren-Lawrence score.\textsuperscript{64} Two studies\textsuperscript{25,41} defined the presence of OA as greater than or equal to Kellgren-Lawrence grade 1. Combined, these studies demonstrated an OA prevalence of 45.9\% (95\% CI = 35.7, 56.4) for surgically treated patients and 25.7\% (95\% CI = 17.1, 36.6) for nonsurgically treated patients. The surgical group had a higher prevalence of tibiofemoral OA than did the nonsurgically treated patients (PR = 1.79, 95\% CI = 1.14, 2.81). Two studies\textsuperscript{23,32} defined OA as a Kellgren-Lawrence score of greater than or equal to grade 2. Use of this scale resulted in a reported OA prevalence of 38.5\% (95\% CI = 30.5, 47.0) for surgically treated patients and 35.1\% (95\% CI = 26.1, 45.3) for nonsurgically treated patients. Based on these 2 studies, the prevalence of tibiofemoral OA was similar among both treatment groups (PR = 1.09, 95\% CI = 0.77, 1.56). Furthermore, radiographic OA was determined using at least 2 patient positions. Two studies\textsuperscript{25,41} used full–weight-bearing anteroposterior radiographs with the knee at 0\% of extension, and the other 2 studies\textsuperscript{23,32} used full–weight-bearing anteroposterior radiographs with the knee in 15\% of flexion.

**DISCUSSION**

This is the first systematic review to report on the prevalence of OA in patients with ACL ruptures treated surgically or nonsurgically based on studies directly comparing these treatments.\textsuperscript{23,25,27,28,32,41,54–56} This approach allowed us to use the PR to compare OA prevalence between the treatments. Unfortunately, because the included studies did not rule out the presence of OA at baseline, the incidence of tibiofemoral OA could not be determined. This limited our systematic review to the evaluation of OA prevalence and the association (ie, the PR) between the treatments. Based on these results, the prevalence of tibiofemoral OA may be greater among surgically treated ACL-deficient patients than among nonsurgically treated patients.
ACL-deficient patients. Unfortunately, the lack of methodologic quality and insufficient data in these studies prohibit a conclusive statement. Regardless, the tibiofemoral OA prevalence rates for both treatments were higher than for the general population, which may suggest that an optimal long-term treatment strategy for preventing tibiofemoral OA after ACL injury is yet to be determined.

Surgical ACL treatments focus on reconstructing the damaged ligament to restore normal knee biomechanics. In contrast, nonsurgical ACL treatment consists of joint mobility training to regain full range of motion, muscle strengthening, and neuromuscular training to promote the restoration of knee function.67 Shortly after injury, both surgically and nonsurgically treated patients appear encouraging, but only 10% to 14% of patients actually returned to their preinjury activity level without limitations.35,45 Furthermore, the rates of ACL-deficient patients who are unable to adequately cope with knee instability and later opt for surgery range from 12% to 39%.35,42,47,68

One explanation for the increased long-term prevalence of OA after ACL injury is the disruption of joint biochemistry. Immediately after injury, the joint undergoes a cascade of changes (eg, increase in inflammatory mediators and cartilage turnover markers) that disrupt the equilibrium between synthesis and catabolism of articular cartilage, influencing how articular cartilage and subchondral bone respond to new loading patterns.3,10,49,72 Elevated levels of C-telopeptide fragments in synovial fluid (a biomarker of cartilage degeneration) and matrix metalloproteinases (catabolic enzymes involved in the degradation of the extracellular matrix) occur within hours of ACL rupture; they gradually decrease over the next year but never return to preinjury levels.52,73 In the nonsurgically treated ACL-deficient knee, type II collagen cleavage begins to return to normal at 12 months after injury, whereas type II collagen synthesis remains elevated at 12 and 24 months postinjury.10 At the time of ACL reconstruction surgery, biochemical synthesis and degeneration of type II collagen were elevated in the injured knee compared with the normal knee.52 At 12 months postsurgery, although cleavage of type II collagen has returned to normal limits, synthesis of type II collagen is elevated and remains so through 24 months. At 24 months, aggrecan turnover begins to approach normal levels but is still elevated. Unfortunately, current treatment strategies do not address these biochemical changes to the joint and, therefore, both surgically and nonsurgically treated patients may be susceptible to tibiofemoral OA.

Although ACL reconstruction does not address the aforementioned biochemical concerns, it may correct the disrupted joint kinematics. However, the long-term outcomes are yet to be determined.13,74–76 Tashman et al13 used a 3-dimensional radiographic stereophotogrammetric motion-analysis system to determine joint kinematics in 6 ACL-reconstructed patients during downhill running. In all reconstructed knees, the femur was more externally rotated and adducted relative to the tibia than in the uninjured contralateral knee. Vertical loading during heel strike and loading rate directed in line with the tibia were less in patients with reconstructed ACLs than in healthy control participants.75 Collectively, these findings demonstrate

---

### Table 2. Frequency, Prevalence (%), and Unadjusted Prevalence Ratios of Tibiofemoral Osteoarthritis by Treatment (Surgical or Nonsurgical) Among Athletes with Anterior Cruciate Ligament Ruptures

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Frequency</th>
<th>Prevalence, % (95% Confidence Interval)</th>
<th>Crude Prevalence Ratios (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical</td>
<td>21</td>
<td>41</td>
<td>51.2 (36.4, 65.8)</td>
</tr>
<tr>
<td>Nonsurgical</td>
<td>11</td>
<td>26</td>
<td>42.3 (25.5, 61.0)</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>67</td>
<td>47.8 (36.3, 59.5)</td>
</tr>
<tr>
<td>Kessler et al41 (2008)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical</td>
<td>27</td>
<td>60</td>
<td>45.0 (33.1, 57.5)</td>
</tr>
<tr>
<td>Nonsurgical</td>
<td>12</td>
<td>49</td>
<td>24.5 (14.6, 38.1)</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>109</td>
<td>35.8 (27.4, 45.1)</td>
</tr>
<tr>
<td>Surgical</td>
<td>29</td>
<td>89</td>
<td>32.6 (23.7, 42.9)</td>
</tr>
<tr>
<td>Nonsurgical</td>
<td>21</td>
<td>65</td>
<td>32.3 (22.2, 44.4)</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>154</td>
<td>32.5 (25.6, 40.2)</td>
</tr>
<tr>
<td>Meuffels et al25 (2009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical</td>
<td>12</td>
<td>25</td>
<td>48.0 (30.0, 66.5)</td>
</tr>
<tr>
<td>Nonsurgical</td>
<td>7</td>
<td>50</td>
<td>28.0 (14.3, 47.6)</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>75</td>
<td>38.0 (25.9, 51.8)</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical</td>
<td>89</td>
<td>215</td>
<td>41.4 (35.0, 48.1)</td>
</tr>
<tr>
<td>Nonsurgical</td>
<td>51</td>
<td>165</td>
<td>30.9 (24.4, 38.3)</td>
</tr>
<tr>
<td>Total</td>
<td>140</td>
<td>380</td>
<td>36.8 (32.1, 41.8)</td>
</tr>
</tbody>
</table>

*a Prevalence percentages calculated using the following formula: (#surgical with osteoarthritis/total surgical)*100. Prevalence ratios calculated as (#surgical with osteoarthritis/all surgical)/(#nonsurgical with osteoarthritis/all nonsurgical).
persistent altered kinematics and biochemical changes not only postinjury but also postsurgery.

The findings of all studies included in this systematic review were limited by the fact that the type and incidence of meniscal injury were not controlled. Patients who have undergone a meniscal repair or meniscectomy have a higher prevalence of OA than those with no meniscal injury.2,19,22,24,26,31,33,34,45,76–81 In an ACL-deficient knee, which is already experiencing biomechanical and biochemical joint changes, meniscal injuries compound the risk of developing OA by further altering the mechanical loading and contact points on the articular cartilage.1,12,14,45,79

Meniscal status is important to the long-term OA outcomes of patients with ACL ruptures.2,15,19–22,24,26,31,33,34,41,77,81–83 In addition, although partial meniscectomies resulted in a greater risk of radiographic changes, the risk was still lower than with total meniscectomies, and substantial function may remain in the residual meniscus.84 However, the potential effect of this remaining tissue on OA risk remains unknown.

The large age range in this systematic review indicates that these results are generalizable to the population with ACL ruptures, but it limits our ability to make inferences about specific groups of patients (eg, high school or college athletes). Although the authors of 2 studies25,41 matched patients according to their Tegner activity scores, authors of the other studies included patients ranging from high-level European soccer players to more sedentary patients injured in motor vehicle accidents or falls. Patients with various activity levels may react differently to treatment options and may have different risks of OA based on their activity levels (eg, sedentary lifestyle, high-level competition).84,85

Another limitation introduced by the large age range is that some patients may have already had joint degeneration at the time of surgery. Therefore, we could not assess whether incidence rates were different between nonsurgical and surgical treatments for ACL injuries.

It is also important to note that we identified only studies of patients with bone–patellar tendon–bone grafts. Other surgical techniques or grafts, such as the hamstrings tendon graft, were not investigated. Future researchers should evaluate the long-term effectiveness of new surgical approaches and graft selections because evidence suggests that patients reconstructed with hamstrings tendon grafts have lower rates of knee OA than those receiving bone–patellar tendon–bone grafts.86–88 Unfortunately, a critical challenge to performing long-term follow-up studies is that new treatment strategies may be adopted as the standard of care before high-quality, long-term follow-up studies can assess their long-term efficacy. This was evident during the study-selection process: a number of groups compared surgical and nonsurgical treatments yet included patients who were immobilized in casts after surgery. This practice is no longer the standard of care, but it was discarded recently enough that an insufficient amount of time has passed to allow degenerative changes to become detectable.

Both the number and quality of studies identified as eligible for inclusion in the systematic review were low. The small number of articles may reflect a publication bias (eg, papers without significant findings were not published). Among the 4 studies, the mean NOS score was 5 of 10 possible points, which was attributed to methodological weakness and inadequate reporting. Although we redefined certain NOS criteria in an effort to standardize the scoring, this could have resulted in lower study scores because some groups reported details concerning the item being assessed but did not provide enough information to meet the necessary number of criteria to qualify for credit on the NOS.

Although the Kellgren-Lawrence score is the most common method for detecting OA, disagreement exists as to the threshold for determining OA, which has been shown to affect the overall classification of OA.89–91 One strength of this systematic review was that we included only articles with direct comparisons: both treatment groups were evaluated using the same diagnostic criteria.23,25,32,41 A limitation to the radiographic OA classification system is the possibility of false negatives, especially in mild cases, when the diseased compartment or joint is compared with internal controls (ie, opposite compartment [medial versus lateral], contralateral knee). No authors of studies included in this systematic review used diagnostic magnetic resonance imaging, which is more costly than radiographic analysis but also more sensitive to degenerative changes.92

Furthermore, if we do not know whether tibiofemoral OA was present at baseline, we cannot truly determine the risk of tibiofemoral OA. This is also problematic with respect to previous or subsequent knee injuries and is especially limiting when assessing OA in a strictly athletic population. Future researchers must not only include patients with no signs of radiographic OA at baseline but also be diligent in collecting and reporting both this information and a comprehensive history of previously sustained injuries.

Although the systematic review is designed to present the body of literature concerning a specific topic, our systematic review was limited by insufficient reporting on a number of factors, which, greatly limits the results from being generalized to the population of those with ACL ruptures. Authors of the included studies did not report on many factors associated with the development of OA (eg, osteochondral lesions, previous injury). Higher-quality study designs would aid our understanding of how OA develops after surgical or nonsurgical ACL treatment. A randomized control trial93 comparing these 2 treatment options is thus far limited to a 2-year follow-up. We need more randomized clinical trials with sufficient posttreatment follow-up and effective control of confounding factors to increase our understanding of surgical versus nonsurgical management of ACL and the incidence of knee OA.

It is also important for future investigators to consider as subcohorts patients who do not undergo reconstruction but instead modify their level of activity (copers, noncopers, and adapters). Assessing the difference between surgical and nonsurgical treatment of ACL ruptures in copers and noncopers can lead to an improved understanding of the true effectiveness of the 2 treatment options by helping us to identify which treatment effectively decreases the episodes of instability and the occurrence of OA. Although some research has been completed to date,94 lengthier follow-up is needed to better assess the development of OA.

CONCLUSIONS

To date, no definitive evidence supports surgical or nonsurgical treatment after ACL injury to prevent post-
traumatic OA. The prevalence of OA in the included studies was slightly higher in surgically treated than in nonsurgically treated ACL patients at follow-up of approximately 12 years. However, large, overlapping confidence intervals indicate that there is no clear difference. This finding may have clinical importance, but the available studies were methodologically weak. Therefore, a significant relationship cannot be determined between having or not having ACL reconstruction surgery and developing tibiofemoral OA. The current studies were limited by small numbers, low methodologic quality, and a lack of data on confounding factors. Future authors should account for the presence of OA at baseline and focus on directly comparing the surgical and nonsurgical treatment of ACL ruptures while controlling for confounding factors (eg, age, meniscal status, BMI, physical activity).

REFERENCES


32. von Porat A, Roos EM, Roos H. High prevalence of osteoarthritis 14 years after an anterior cruciate ligament tear in male soccer players: a


Address correspondence to Kyle P. Harris, MS, ATC, Department of Health, Physical Education and Nursing, Bucks County Community College, 275 Swamp Road, Newtown, PA 18940. Address e-mail to kpf85@gmail.com.
### Appendix. Newcastle-Ottawa Scale Scores

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection (maximum = 1 point)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Representativeness of the general ACL-deficient population</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(a) Truly representative of the general ACL-deficient population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Somewhat representative of the general ACL-deficient population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Selected group of users (eg, nurses, volunteers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) No description of the derivation of the cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Selection of the nonexposed cohort (nonsurgical) (maximum = 1 point)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(a) Drawn from the general ACL-deficient population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Drawn from a different source</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Ascertainment of exposure (surgical) (maximum = 1 point)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(a) Secure records (eg, surgical records)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Structured interview</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Written self-report</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Ascertainment of nonexposure (nonsurgical) (maximum = 1 point)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(a) Secure records (eg, surgical records)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Structured interview</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Written self-report</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Demonstration that the outcome of interest was not present at start of study</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(a) Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparability (maximum = 2 points)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1. Comparability of the cohorts on the basis of design or analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) Study controls for secondary injury (ie, meniscal, ligamentous other than ACL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Study controls for body mass index</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome (maximum = 1 point each for items 1–3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Assessment of outcome</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>(a) Independent blind assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Record linkage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Self-report</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) No description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Was follow-up long enough for outcomes to occur?</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>(a) Yes (10 y or more)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Adequacy of cohort follow-up</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(a) Complete follow-up: all participants accounted for</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) Participants lost to follow-up unlikely to bias (small number lost or &gt;80% follow-up or description provided of those lost)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) Follow-up rate &lt;80% and no description of those lost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) No statement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall consensus score (out of possible 10)</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Time to read and score each article, average min</td>
<td>9</td>
<td>11</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Abbreviation: ACL, anterior cruciate ligament.