Clinical examination of anterior cruciate ligament rupture: a systematic review and meta-analysis

Wenhan HUANG1,2, Yu ZHANG1,2, Zilong YAO1,2, Linmin MA2

1Graduate School of Southern Medical University, Guangzhou, China
2Department of Orthopedics, Guangzhou General Hospital of Guangzhou Military Command, Guangzhou, China

Objective: The purpose of this study was to evaluate the sensitivity and specificity of 3 tests for assessing anterior cruciate ligament (ACL) ruptures.

Methods: MEDLINE, EMBASE, Cochrane Library, and CBM (Chinese Biomedical Literature Database) searches were performed. Studies selected for data extraction were those that addressed the accuracy of at least 1 physical diagnostic test for ACL rupture in comparison with a clinical reference standard such as arthroscopy, arthrotomy, or magnetic resonance imaging (MRI). The references of the included studies were also reviewed. Searches were limited to English and Chinese languages.

Results: Sixteen studies that assessed the accuracy of the 3 tests for diagnosing ACL ruptures met the inclusion criteria. Study results were, however, heterogeneous. The Lachman test is the most sensitive test to determine ACL tears, showing a pooled sensitivity of 87.1% (95% confidence interval [CI] 0.84–0.90). The pivot shift test is the most specific test, showing a pooled specificity of 97.5% (95% CI 0.95–0.99); additionally, it has the highest positive likelihood ratios (LR+) of 16.00 (95% CI 7.34–34.87). The Lachman test has the lowest negative likelihood ratios (LR−) of 0.17 (95% CI 0.11–0.25).

Conclusion: In cases of suspected ACL injury, it is recommended to perform the pivot shift test, as it is highly specific and has greater likelihood and discrimination of accurately diagnosing ACL rupture. The Lachman test has great efficacy in ruling out a diagnosis of ACL rupture because of the lowest negative likelihood ratios.

Keywords: Anterior cruciate ligament; examination; ruptures.

Level of Evidence: Level I, Diagnostic study.

The anterior cruciate ligament (ACL) is one of the most commonly injured ligaments of the knee. In America, estimates of ACL injury cases range from 80,000 to 250,000 per year, with approximately 100,000 of these patients undergoing ACL reconstruction surgery.[1] With an estimated cost for these injuries of almost one billion dollars per year, identifying risk factors and developing prevention strategies as well as optimal treatment methods have widespread health and economic implications. Physical examinations are commonly used, especially in remote areas and small clinics where magnetic resonance imaging (MRI) or arthroscopy is not commonly used to make a diagnosis. The 3 primary diagnostic assessments of these manual tests are the anterior drawer test,
the Lachman test, and the pivot shift test, which are the most commonly used in practice and provide great efficacy. Some reports\(^2,^3\) suggest that the pivot shift test is very specific yet has poor sensitivity. The Lachman test is the most valid test to determine ACL tears, showing an excellent pooled sensitivity.\(^4\) Previous studies\(^5^–^7\) confirmed these findings. Furthermore, some studies\(^8,^9\) found that carefully performed clinical examination can give equal or better diagnosis of meniscal and ACL injuries in comparison to MRI scan. Our current practice of routinely requesting MRI scans to confirm the diagnosis may be altered, as unnecessary MRI scanning increases financial burden and delays patient treatment.\(^10\) MRI may be used to rule out such injuries rather than to di-

<table>
<thead>
<tr>
<th>Table 1. Operationalization of the search strategy.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Steps</strong></td>
</tr>
<tr>
<td>#1</td>
</tr>
<tr>
<td>#2</td>
</tr>
<tr>
<td>#3</td>
</tr>
<tr>
<td>#4</td>
</tr>
<tr>
<td>#5</td>
</tr>
<tr>
<td>#6</td>
</tr>
<tr>
<td>#7</td>
</tr>
<tr>
<td>#8</td>
</tr>
<tr>
<td>#9</td>
</tr>
<tr>
<td>#10</td>
</tr>
</tbody>
</table>
agnose them. Despite these findings, previous reports about the 3 tests had various results and small numbers of subjects; these differences created concern regarding the validity of clinical diagnostic tests for ACL ruptures. Furthermore, few databases were searched among for the majority of reviews, leading to a lack of statistical power from the results. Thus, there is a need for further study with meta-analysis to have quantitative data which can help clinicians to make the best diagnosis.

The aim of this study was to perform a meta-analysis of original research studies that involves the diagnostic sensitivity and specificity of the 3 assessments to evaluate the diagnostic accuracy of the anterior drawer, Lachman, and pivot shift tests.

**Materials and methods**

Relevant studies were searched for according to the search strategy of the Cochrane Collaboration. Medical Subject Headings (MeSH) and free word were restricted to English and Chinese. The following electronic databases were utilized to identify studies: MEDLINE, EMBASE (Excerpta Medica Database), the Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library, and CBM (the Chinese Biomedical Literature Database). The search strategy flows are as shown in Table 1.

The title and the abstract of all identified articles were reviewed for relevance. The study quality and extracted data were assessed by 2 authors. In cases of disagreement between the 2 authors, consensus was reached by discussion with a third author. Corresponding authors of the reviewed articles were contacted to obtain missing data if necessary. All available abstracts were examined, and the full-text articles were examined and evaluated for the included studies. The description of evidence of the included studies is based on the evidence pyramid.

Articles included were written in English or Chinese, and addressed the accuracy of at least 1 physical diagnostic test for ACL rupture studies using arthroscopy, arthrotomy, or MRI as reference standards. Arthroscopy and arthrotomy are commonly accepted as the clinical gold standards for ACL rupture. MRI is now also considered a gold standard, as recent literature[11] has shown an excellent correlation between MRI and arthroscopic and arthrotomy findings. Although accompanying lesions such as meniscus or medial/lateral collateral ligament deficiency appear in knee joints with ACL rupture, for the purposes of this study, we chose patients with isolated ACL-deficiency whose diagnoses were confirmed by the gold standard.

The methodological quality of the selected studies was assessed and rated by 2 reviewers independently.
Table 2.  Regarding the classification of acute and chronic ACL ruptures, decisions adopted by the authors of the studies included in this meta-analysis were followed. If no mention is made of acute or chronic, the authors did not specify.

<table>
<thead>
<tr>
<th>Source (year)</th>
<th>Level of evidence</th>
<th>NO. of subjects</th>
<th>Patient population standard</th>
<th>Reference</th>
<th>Examination maneuver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperman, 1990</td>
<td>2</td>
<td>32</td>
<td>Consecutive patients with knee pain; chronic</td>
<td>Arthroscopy</td>
<td>Lachman</td>
</tr>
<tr>
<td>Lee, 1988</td>
<td>4</td>
<td>41</td>
<td>Nonconsecutive patients who underwent MRI and then arthroscopy</td>
<td>Arthroscopy</td>
<td>ADT</td>
</tr>
<tr>
<td>Lachman Harilainen, 1987</td>
<td>2</td>
<td>350</td>
<td>Consecutive patients with knee injuries; acute</td>
<td>Arthroscopy</td>
<td>ADT</td>
</tr>
<tr>
<td>Wagemakers, 2010</td>
<td>2</td>
<td>64</td>
<td>Consecutive patients with knee pain; chronic</td>
<td>MRI</td>
<td>ADT</td>
</tr>
<tr>
<td>Kim, 1995</td>
<td>2</td>
<td>147</td>
<td>Consecutive patients with knee pain; chronic</td>
<td>Arthroscopy</td>
<td>ADT</td>
</tr>
<tr>
<td>Nickinson, 2010</td>
<td>5</td>
<td>698</td>
<td>Nonconsecutive patients who underwent arthroscopy</td>
<td>Arthroscopy</td>
<td>Lachman PST</td>
</tr>
<tr>
<td>Katz, 1986</td>
<td>5</td>
<td>85</td>
<td>Nonconsecutive patients who underwent arthroscopy</td>
<td>Arthroscopy</td>
<td>Lachman PST</td>
</tr>
<tr>
<td>Sandberg, 1986</td>
<td>2</td>
<td>182</td>
<td>Consecutive patients with knee injuries; acute</td>
<td>Arthroscopy</td>
<td>ADT</td>
</tr>
<tr>
<td>Steinbrück, 1988</td>
<td>5</td>
<td>300</td>
<td>Nonconsecutive patients who underwent arthroscopy</td>
<td>Arthroscopy</td>
<td>ADT</td>
</tr>
<tr>
<td>Rubinstein, 1994</td>
<td>2</td>
<td>39</td>
<td>Consecutive patients with knee injuries; chronic</td>
<td>Arthroscopy</td>
<td>Lachman PST</td>
</tr>
<tr>
<td>Richter, 1996</td>
<td>2</td>
<td>74</td>
<td>Consecutive patients with knee injuries; chronic and acute</td>
<td>Arthroscopy</td>
<td>ADT</td>
</tr>
<tr>
<td>Hardaker, 1990</td>
<td>2</td>
<td>132</td>
<td>Consecutive patients with knee injuries and hemarthrosis presenting to sports medicine clinic; acute</td>
<td>Arthroscopy</td>
<td>Lachman PST</td>
</tr>
<tr>
<td>Schwarz, 1997</td>
<td>2</td>
<td>58</td>
<td>Consecutive patients with knee injuries</td>
<td>Arthroscopy</td>
<td>Lachman PST</td>
</tr>
<tr>
<td>Dejour, 2013</td>
<td>2</td>
<td>300</td>
<td>Consecutive patients with knee injuries; acute and chronic</td>
<td>Arthroscopy; MRI</td>
<td>PST</td>
</tr>
</tbody>
</table>


ADT: Anterior drawer test; PST: Pivot shift test; MRI: Magnetic resonance imaging.
The criteria for the evaluation of methodological quality was appraised using the QUADAS-2 tool. This tool comprises 4 domains: patient selection, index test, reference standard, and flow and timing. Each domain was assessed in terms of risk of bias and classified accordingly as low, high, or unclear. The QUADAS-2 tool allows for an objective and transparent rating of bias and applicability of primary diagnostic accuracy studies.

RevMan version 5.0 and Meta-Disc version 1.4 were used to analyze the data. Summary sensitivities, specificities, and positive and negative likelihood ratios (LR+/LR−) with 95% confidence interval (CI) were calculated using random effect models with Der-Simonian Laird method or fixed effect models with Mantel-Haenszel method, depending on the level of heterogeneity of the study. The data were presented with forest plots and receiver operating characteristic (ROC) curve plots. The percentage of variability was crossed by the chi-square test (p<0.10) and I2 statistics. The random effects model was used if the heterogeneity test showed statistical significance (I2>50%, p<0.10). If not, the fixed effects model was adopted. In addition, area under curve (AUC) was combined with 95% CI. A mean value of AUC-ROC>0.70 defined an effective risk predictor.

Results

A flow chart of the study selection process is presented in Figure 1. The EMBASE search provided 413 hits and the PubMed search provided 2617 hits. In addition, 56 hits and 40 hits were found in CENTRAL and CBM, respectively. From these 3126 total hits, 72 studies were selected based on their title and abstract. Fifty studies were excluded due to the selection criteria. Eight studies were excluded because the full text was not available. Consequently, 14 studies were included based on their full text. A total of 2502 patients were involved in these studies. Details of the included studies are shown in Table 2.

The methodological quality of the included studies is presented in Figures 2 and 3. Overall, the methodological quality of the included studies is favorable. Firstly, for domain 1, only 2 studies had high risk of bias in patient selection, as patients were selected without any specific description. Four studies did not clearly denote consecutive selection but had a clear presence of ACL deficiency. A consecutive or random sample of patients enrolled...
was found in other studies, and case-control design was avoided. Secondly, all of the included tests had low risk of bias in domain 2 index test. Thirdly, for domain 3, reference standard, 3 studies did not clarify the reference standard clearly because they did not state whether the reference standard results were interpreted without knowledge of the results of the index test. Lastly, for domain 4, results of the index test and reference standard were ideally collected on the same patients at the same time among 9 studies. In 2 studies with high risk, some patients were addressed with more than 1 reference standard to assess the accuracy of physical diagnostic tests for ACL rupture. In 1 study, a delay occurred between the index test and reference standard with an unclear risk, as it could be problematic for patients with acute infectious diseases such as ACL deficiency. The other 2 studies with unclear risk did not state clearly the interval between the index test and reference standard, but all patients in these studies accepted the same reference standard.

Sensitivity and specificity with 95% CI and forest plots for the anterior drawer test, Lachman test, and pivot shift test are shown in Figure 4 for all the included studies. The overall sensitivity of the anterior drawer test was 0.725, and specificity was 0.927. For the Lachman test, the overall sensitivity and specificity were 0.871

<table>
<thead>
<tr>
<th>Study</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee. 1988</td>
<td>0.78 (0.52 – 0.94)</td>
<td>1.00 (0.85 – 1.00)</td>
</tr>
<tr>
<td>Harlainen A. 1987</td>
<td>0.41 (0.33 – 0.50)</td>
<td>0.86 (0.80 – 0.90)</td>
</tr>
<tr>
<td>Wagemakers. 2010</td>
<td>0.82 (0.63 – 0.94)</td>
<td>0.58 (0.41 – 0.74)</td>
</tr>
<tr>
<td>Nickinson. 2010</td>
<td>0.86 (0.76 – 0.93)</td>
<td>0.98 (0.97 – 0.99)</td>
</tr>
<tr>
<td>Katz. 1986</td>
<td>0.41 (0.21 – 0.64)</td>
<td>0.95 (0.87 – 0.99)</td>
</tr>
<tr>
<td>Sandberg. 1986</td>
<td>0.78 (0.70 – 0.85)</td>
<td>0.80 (0.64 – 0.91)</td>
</tr>
<tr>
<td>Steinbrück. 1988</td>
<td>0.92 (0.86 – 0.95)</td>
<td>0.91 (0.85 – 0.95)</td>
</tr>
<tr>
<td>Rubinstein. 1994</td>
<td>0.76 (0.53 – 0.92)</td>
<td>0.83 (0.59 – 0.96)</td>
</tr>
<tr>
<td>Richter. 1996</td>
<td>0.67 (0.54 – 0.79)</td>
<td>0.88 (0.62 – 0.98)</td>
</tr>
</tbody>
</table>

Pooled Sensitivity = 0.73 (0.69 to 0.76)
Chi-square = 116.66; df = 8 (p=0.0000)
Inconsistency (I-square) = 93.1%

Pooled Specificity = 0.93 (0.91 to 0.94)
Chi-square = 10.72; df = 7 (p=0.1513)
Inconsistency (I-square) = 34.7%

Pooled Sensitivity = 0.87 (0.84 to 0.90)
Chi-square = 10.72; df = 7 (p=0.1513)
Inconsistency (I-square) = 34.7%

Pooled Specificity = 0.91 (0.89 to 0.93)
Chi-square = 52.26; df = 7 (p=0.0000)
Inconsistency (I-square) = 86.6%

Pooled Sensitivity = 0.49 (0.43 to 0.55)
Chi-square = 2 (p=0.0000)
Inconsistency (I-square) = 97.6%

Pooled Specificity = 0.98 (0.95 to 1.00)
Chi-square = 8 (p=0.0000)
Inconsistency (I-square) = 37.8%

Fig. 4. (a-c) Forest plots of sensitivity and specificity. [Color figure can be viewed in the online issue, which is available at www.aott.org.tr]
and 0.97, respectively. For the pivot shift test, the overall specificity was 0.975; however, this test had a low sensitivity of 0.490. LR+ and LR− are shown in Figure 5. The Lachman test had the lowest LR− (0.17), and the pivot shift test had the highest LR+ (16.00). The ROC curves for the anterior drawer test, Lachman test, and pivot shift test are displayed in Figure 6. The high ROC value is similar between the 3 tests (0.887, 0.928, 0.974, respectively).

Discussion

As the increasing prevalence of ACL injuries brings financial burden to the healthcare sector, investigations must be urgently conducted on how to diminish the risk of developing ACL rupture. Therefore, it is of great socioeconomic significance to make a sound diagnosis in cases of suspected ACL rupture. We found that the pivot shift test is the most specific and has favorable LR+ and diagnostic odds. With a 0.17 LR− value, the

![Fig. 5. (a-c) Forest plots of positive/negative likelihood ratio. [Color figure can be viewed in the online issue, which is available at www.aott.org.tr]]
Lachman test may be more useful for ruling out ACL deficiency than for predicting ACL injury. The anterior drawer test may be the least efficient tool in diagnosing ACL deficiency of these 3 methods which are most often used by practitioners, as it is of unproven diagnostic value in this setting.

In this study, we provided evidence that the anterior drawer test has moderate specificity of 0.93 and poor sensitivity of 0.73. Tore et al.\cite{10} put forward 3 potential causes for a false-negative anterior drawer test in acute condition, especially in isolated ACL tears. First, hemarthrosis and reactive synovitis may preclude knee flexion to 90°, hindering performance of the test. Second, protective muscle action of the hamstrings secondary to joint pain provides a vector force opposite to the anterior translation of the tibia. Third, the posterior horn of the medial meniscus becomes buttressed against the posterior-most margin of the medial femoral condyle and may preclude anterior translation of the tibia.

Our results indicate that the Lachman test is the most sensitive method in diagnosing ACL rupture. Furthermore, this method has favorable performance in specificity, as the position of the knee during this test (20–30° of flexion) is less painful than the position of the knee during the anterior drawer test; hence, it reduces possible muscle action to protect the knee during testing.\cite{6} Additionally, the diagnosis of anterior cruciate ligament rupture is often difficult to establish, especially in recent injuries with acute hemarthrosis. The diagnostic accuracy of the Lachman test in recent ruptures when the patient is examined without general anesthetic is superior to that of the anterior drawer test.\cite{15}

The pivot shift test evaluates the combined tibiofemoral internal rotation and anterior tibial translation that occurs when the ACL is injured or deficient.\cite{16} The pivot shift test is a complex multiplanar maneuver that incorporates 2 main components: translation (the anterior subluxation of the lateral tibial plateau followed by its reduction) and rotation (the rotation of the tibia relative to the femur).\cite{17} The pivot shift test reproduces the phenomenon of giving way of the knee. The specificity of the pivot shift test is very high, namely 98% (95% CI 95–99). We found, however, very poor sensitivities of 49% (95% CI 43–55). The reason for the very low sensitivity may be explained by the fact that a patient with a chronic ACL-deficient knee is familiar with this unpleasant phenomenon and will show protective muscle action.\cite{18}

Our results are compatible with the findings of Scholten\cite{4} and Benjaminse\cite{6} et al. In the meta-analysis by Scholten et al., the pivot shift test had high specificity of 98% (95% CI 96–99) but poor sensitivity of 24% (95% CI 21–27). The Lachman test was the most valid test to determine ACL tears, showing pooled sensitivity of 85% (95% CI 83–87) and pooled specificity of 94% (95% CI 92–95). In the study by Benjaminse et al., the pivot shift test showed specificity ranging from 0.97 to 0.99 and sensitivity from 0.18 to 0.48. Both of the stud-
ies demonstrated that the pivot shift test is the most specific but has poor sensitivity. For the Lachman test, pooled sensitivity of 86% (95% CI 76–92) and specificity of 91% (95% CI 79–96) was reported by Scholten et al. In accordance with the report by Scholten et al., pooled sensitivity of 85% (95% CI 83–87) and specificity of 94% (95% CI 92–95) was reported by Benjaminse et al. These results are consistent with our findings in this meta-analysis. For the anterior draw test, neither of the studies suggested that an acceptable value of sensitivity was obtained. Sensitivity of the anterior drawer test was 0.62 (95% CI 42–78) from Scholten and 0.49 (95% CI 43–55) from the Benjaminse et al. study. Compared with the sensitivity of the anterior drawer test from previous studies listed above, corresponding results were achieved in our study, namely 0.73 (95% CI 69–76).

Regarding the quality of the included studies, the AUC values of the anterior drawer, Lachman, and pivot shift tests were large enough (0.887, 0.928, 0.974, respectively), and favorable methodology quality was employed, even though statistical heterogeneity must not be ignored. First, generally a value of AUC from 0.7 to 0.9 is considered as low accuracy of diagnosis, while a value from 0.5 to 0.7 is considered as medium accuracy. A result will be considered appreciable if the value of AUC is ≥0.90.[19] The maximum AUC value of 1 means that the diagnostic test is ideal for the differentiation in diagnostic test evaluation. Secondly, with respect to likelihood ratios, the Lachman test has the lowest LR− (0.17), and the pivot shift test has the highest LR+ (16.00). LR+ and LR−, which describe the discriminatory properties of positive and negative test results, respectively, were calculated.[20] It is possible to infer from these findings that the pivot shift test has the highest efficacy in diagnosing a suspected ACL rupture, while the Lachman test has the highest efficacy in ruling out a suspected ACL rupture.

The overall methodological quality of the included studies is favorable, except in domain 1. Firstly, for domain 1, a consecutive or random sample of patients enrolled was found in 8 studies, and a case-control design was avoided. Two studies had high risk of bias in patient selection, as patients were selected without any specific description. Additionally, 4 studies did not clearly denote consecutive selection. Thus, there was possibility of bias in this meta-analysis, potentially resulting in the overestimation of diagnostic accuracy. Secondly, for the domain 2, all index results were interpreted without knowledge of the results of the reference standard. This measure confirmed that the included studies could possibly avoid the introduced bias of the index test when conducted or interpreted. For the second signaling question—“If a threshold was used, was it prespecified?”—we are of the opinion that a threshold will not be used in clinical examination and therefore ignored this question. Thirdly, most of the reference standards (11 of 14), their conduct, and their interpretation avoided the introduced bias of reference standards such as arthroscopy, arthrotomy, and MRI, which are commonly used in practice and have high reliability and validity. Fourthly, for domain 4, results of the index test and reference standard are ideally collected from patients among 9 studies; acute ACL rupture will accept reference standards such as arthroscopy, arthrotomy, or MRI immediately after clinical examinations in case of delay. Chronic ACL rupture would allow several days delay between the index test and reference standard. In 2 studies with high risk of bias, some patients confirmed the accuracy of physical diagnostic test for ACL rupture with more than 1 reference standard, such as arthroscopy, arthrotomy, or MRI. This raised verification bias. Two studies with unclear risk did not state clearly the interval between the index test and reference standard, but all patients in both studies accepted the same reference standard. Overall, the quality of the included studies is favorable, and thus the resultant meta-analysis is convincing.

This meta-analysis has limitations. Firstly, most of the included studies were not performed recently, and most of the studies we searched in the databases did not provide enough details to calculate pooled statistical data. Therefore, only 14 studies were included in this meta-analysis. Secondly, the included studies were not homogeneous. Different study designs, setting, number of subjects, spectrum of the disease, and reference standard created heterogeneity. These differences make direct comparison of various studies difficult. As the number of included studies is small, the decision was made to not perform a meta-regression to determine the cause of heterogeneity. Thirdly, 2 or 3 clinical tests were conducted with anesthesia. This may lead to bias because the results of the test performed first could influence the testers’ grading of subsequent tests. Considering these aspects, there is undoubtedly a growing need for studies based on sound research with a high level of evidence.

In cases of suspected ACL injury, it is recommended to perform the pivot shift test, as it is very specific and has greater likelihood ratios in diagnosing ACL rupture. The Lachman test has favorable efficacy in ruling out a diagnosis of ACL rupture, while the anterior drawer test is the least proven of the 3 approaches in diagnosing ACL rupture.
Conflicts of Interest: No conflicts declared.

References


