Systematic Review/Meta-Analysis

Diagnostic Accuracy of Physical Examination Tests for Suspected Acute Anterior Cruciate Ligament Injury: A Systematic Review and Meta-Analysis

Shigeharu Tanaka, a Yu Inoue, Yuki Masuda, Hui Tian, Hungu Jung, Ryo Tanaka

Keywords: Anterior cruciate ligament injury, diagnostic test accuracy, physical examination test, systematic review

https://doi.org/10.26603/001c.36434

International Journal of Sports Physical Therapy
Vol. 17, Issue 5, 2022

Background
Many tests are used to examine the knee when anterior cruciate ligament (ACL) injury is suspected. However, evidence of diagnostic accuracy in the Lachman, anterior drawer, pivot shift, and lever sign tests is limited.

Purpose
The purpose of this study was to perform a systematic review and meta-analysis of original research studies that assessed the diagnostic accuracy of four physical examination tests for ACL injury acutely after an injury.

Study design
Systematic review and meta-analysis.

Methods
A literature search was conducted in the PubMed, MEDLINE, CINAHL, Web of Science, and Ichushi databases. Original articles with prospective cohort and cross-sectional studies in English and Japanese were included. The searched words were "anterior cruciate ligament", "injury", "rupture", "tear", "lachman test", "pivot shift test", "anterior drawer test", "lever sign test". The methodological quality of the diagnostic studies was evaluated using QUADAS-2. Summary sensitivity, specificity, likelihood ratio (LR)+, and LR− with 95% confidence intervals were calculated.

Results
Eight studies were included in this review. The methodological quality of the included studies was mostly favorable. For the domain of flow and timing in the QUADAS-2, three studies did not assess the timing between the reference and index tests. The pooled sensitivities were 0.79, 0.78, 0.55, and 0.82 in the Lachman, anterior drawer, pivot shift, and lever sign tests, respectively, and the pooled specificities were 0.91, 0.91, 0.96, and 0.88, respectively. The lever sign test had the lowest LR− (0.21) and the pivot shift test had the highest LR+ (11.60). The area under the curve for the four physical examinations was > 0.70.

Conclusion
The lever sign and pivot shift tests are useful for diagnosing ACL injuries in an acute setting.

Corresponding author:
Dr. Shigeharu Tanaka
Division of Physical Therapy, School of Rehabilitation, Faculty of Health and Social Services, Kanagawa University of Human Services, 1-10-1 Heisei-cho, Yokosuka, Kanagawa 238-8522, Japan
Phone: +81-46-828-2792, Fax: +81-46-828-2793
E-mail: tanaka-d4k@kuhs.ac.jp
INTRODUCTION

Anterior cruciate ligament (ACL) injury is one of the most common injuries in the knee joint occurring amongst adolescents and young adults during sports that involve sudden stops or changes in direction, jumping, and landing. Diagnosis of ACL injury in the initial medical consultation is difficult because of the common acute signs and symptoms of an ACL injury, such as immediate pain, swelling, and hemarthrosis. According to two previous studies in the U.K, correct ACL injury diagnosis was made in 28.2% and 14.4% of cases in the acute phase, respectively, which suggests that diagnostic accuracy is lower and difficult in the acute phase, as the median time to diagnosis was six weeks after ACL injury. Poor diagnostic accuracy within the first six weeks after injury is often related to an unnecessary delay in the diagnosis and subsequent treatment, and increases the risk of secondary injury to the knee joint. Accurate diagnosis in the acute stage is needed to reduce these secondary risks and appropriately manage ACL injuries. Diagnostic imaging modalities and physical examinations are used to diagnose ACL injuries. Magnetic resonance imaging (MRI) is the gold standard owing to its high diagnostic accuracy. A previous systematic review indicated that the positive likelihood ratio (LR+) and negative likelihood ratio (LR−) were 14.4 (95% confidence interval [CI] = 9.2–22.5) and 0.1 (95% CI = 0.1–0.3), respectively, indicating that MRI is a useful independent tool for diagnosing ACL injury. Conversely, physical examinations performed by athletic trainers, physical therapists, and medical doctors are inexpensive and can be performed without the use of highly specialized machines. Four tests are commonly used to diagnose or determine ACL injury: the Lachman test, the anterior drawer test, the pivot shift test, and the lever sign test. The diagnostic accuracy of these tests has been examined in systematic reviews and meta-analyses. The Lachman test had the lowest LR− of 0.17 (95% CI = 0.11–0.25), and the pivot shift test had the highest LR+ of 16.00. Additionally, the latest systematic review, conducted in 2019 regarding the lever sign test, suggested that it may be used in addition to other tests to rule out the presence of an ACL injury.

Based on these findings of systematic reviews, physical examination is useful in clinical practice for diagnosing ACL injury. However, previous meta-analyses have not analyzed acute and chronic conditions individually, and no systematic review has examined whether these physical examination tests are useful in acute settings for diagnosing ACL injuries. There is a possibility that the diagnostic accuracy for ACL injury differs between acute and chronic conditions because of the effects of signs and symptoms in acute conditions, such as pain, swelling, and hemarthrosis. Additionally, there has been no systematic review examining the diagnostic accuracy of these four physical examinations, including the lever sign test, in the same analysis, and the evidence for the efficacy of these tests for the diagnosis of acute ACL injury is limited. Clarifying the diagnostic accuracy of physical examination tests can help healthcare providers with the decision-making process in clinical practice.

Thus, this study aimed to perform a systematic review and meta-analysis of original research studies that involved the diagnostic accuracy of four physical examination tests for ACL injuries in acutely after an injury in adolescents and young adults.

MATERIALS AND METHODS

This study was conducted according to the Preferred Reporting Items for Systematic Review and Meta-analysis of Diagnostic Test Accuracy (PRISMA-DTA). The study protocol of this systematic review was registered in the University Hospital Medical Information Network (UMIN), the biggest registration system in Japan. The study approval number was UMIN000044766.

PHYSICAL EXAMINATION TESTS

This systematic review included studies evaluating the accuracy of four physical examination tests (Lachman, anterior drawer, pivot shift, and lever sign tests) in diagnosing ACL injuries. All tests are performed in the supine position. The Lachman test is performed with the knee flexed at 15°. The examiner stabilizes the distal femur with one hand and grasps the proximal tibia with the other hand. Then, the examiner applies an anterior tibial force to the proximal tibia. The anterior drawer test is performed with the knee flexed at 90°. The examiner sits on the patient’s foot and grasps behind the proximal tibia with the thumbs palpating the tibial plateau and index fingers palpating the tendons of the hamstring muscle group medially and laterally. The anterior tibial force is then applied by the examiner. In these two tests, a positive test for a torn ACL is indicated by greater anterior tibial displacement on the affected side when compared to the unaffected side. In the pivot shift test, the examiner stands on the side of the patient’s knee under examination. The examiner wraps one arm around the patient’s leg, pinning it firmly and flexing the knee to 90° while using the palm of the other hand to rotate the tibia medially, effectively subluxing the lateral tibial plateau. The examiner slowly extends the knee while maintaining the rotation of the tibia. As the patient’s knee reaches full extension, the tibial plateau relocates. A positive test traditionally is indicated by an audible or palpable click. In the lever sign test, the patient lies in the supine position with both legs fully extended. One fist of the examiner is placed under the distal third of the patient’s quadriiceps of same leg. A positive test is indicated by a heel rise (Figure 1).

Level of Evidence
Level 2
Studies were included in this review if they met the following inclusion criteria: (1) original articles in English or Japanese; (2) single-gate studies, such as prospective cohort and cross-sectional studies including acute settings; (3) reported findings that enabled the calculation of the number of true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) values for the test accuracy of index tests (the present study included articles that reported sensitivity, specificity, and prevalence to manually calculate TP, TN, FP, and FN if the study did not show specific numbers); (4) reported data on adolescents and young adults; (5) reported findings of partial or complete ACL injury; and (6) evaluated MRI and/or arthroscopy as the reference standard. Based on previous studies, the acute phase was defined as six weeks in this study. No limits regarding the date of publication were established, although the following types of studies were excluded: (1) participants who had undergone surgery in knee joints affected by ACL injury, (2) participants with other associated diseases such as osteoarthritis and meniscus injury, (3) diagnosis determined using questionnaires, (4) diagnosis determined using devices, and (5) participants who were diagnosed with laxity.

SEARCH STRATEGY

The languages used in the search were English and Japanese. The electronic search was conducted using the following databases: PubMed, Medical Literature Analysis and Table 1. Search strategy

<table>
<thead>
<tr>
<th>#</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>“Anterior cruciate ligament”</td>
</tr>
<tr>
<td>#2</td>
<td>Injury</td>
</tr>
<tr>
<td>#3</td>
<td>Rupture</td>
</tr>
<tr>
<td>#4</td>
<td>Tear</td>
</tr>
<tr>
<td>#5</td>
<td>#2 OR #3 OR #4</td>
</tr>
<tr>
<td>#6</td>
<td>“Lachman test”</td>
</tr>
<tr>
<td>#7</td>
<td>“Pivot shift test”</td>
</tr>
<tr>
<td>#8</td>
<td>“Anterior drawer test”</td>
</tr>
<tr>
<td>#9</td>
<td>“Lever sign test”</td>
</tr>
<tr>
<td>#10</td>
<td>#6 OR #7 OR #8 OR #9</td>
</tr>
<tr>
<td>#11</td>
<td>#1 AND #5 AND #10</td>
</tr>
</tbody>
</table>

Retrieval System Online (MEDLINE), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Web of Science, and Ichushi (a Japanese medical database). The most recent search was conducted on April 23, 2021. The details of the search strategy are presented in Table 1. The same approach was used for all searches and was adopted as necessary, according to the specifics of each database.

SELECTION OF THE STUDIES

First, duplicate studies were removed after searching databases. One evaluator (ST) read the titles and abstracts of the identified articles and excluded irrelevant studies. The full texts of the selected studies were evaluated, and the suit-
ability for inclusion was determined by two independent evaluators (ST and YI). Disagreements between evaluators were resolved by consensus. In cases in which no consensus was reached, a third evaluator (RT) was consulted to determine eligibility.

RISK OF BIAS

The methodological quality of the diagnostic studies was evaluated using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 tool.\(^9\) This tool comprises four domains: patient selection, index test, reference standard, and flow and timing. Each domain of the QUADAS-2 was assessed in terms of the risk of bias and classified accordingly as low, high, or unclear by two independent evaluators, ST and RT. Applicability concerns of the three domains of the QUADAS-2, such as patient selection, index test, and reference standard, were assessed by ST and RT.

STATISTICAL ANALYSIS

To analyze the data, free software Meta-Disc version 1.4 was used. The statistical methodology was based on a previous study described by Devillé et al.\(^9\) Summary sensitivity, specificity, LR+, and LR− with 95% CI were calculated using random-effects models with the Der-Simonian and Laird method or fixed-effects models with the Mantel-Haenszel method, depending on the level of heterogeneity of the study. Data are presented as forest plots and summary receiver operating characteristic (SROC) curve plots. The percentage of variability was crossed using the chi-squared test (p < 0.10) and I\(^2\) statistics. A random-effects model was used if the heterogeneity test showed statistical significance (I\(^2\) > 50%, p < 0.01). Otherwise, a fixed effects model was used. In addition, the area under the curve (AUC) was calculated using 95% CI. A mean AUC-ROC value of > 0.70, was defined as an effective risk predictor.

RESULTS

A systematic electronic search of PubMed, MEDLINE, CINAHL, Web of Science, and Ichushi resulted in 3,836 studies. After duplicates were removed, 1,536 studies were included. The titles and abstracts of these studies were screened, and 86 full-text articles were evaluated for eligibility. Of the 86 articles, 40 were not applicable according to the study design, 14 did not describe test accuracy, 10 were review articles, eight articles did not use physical examination tests, and six articles did not include the acute phase. This, eight studies were included in this review. A summary of the literature search is shown in Figure 2. The characteristics of the included studies are summarized in Table 2.

A total of 620 participants were evaluated in the included studies. Six of the articles involved assessment using an index test by medical doctors, and two of the articles reported assessment using index test by physical therapists. The Lachman test was used in eight studies, the anterior drawer test in six studies, the pivot shift test in five studies, and the lever sign test in four studies. For the reference test, seven studies used MRI, one study used arthroscopy, and two studies used both MRI and arthroscopy. Index tests, which included the four physical examination tests of interest in this review, were performed while awake in four studies, under anesthesia in one study, and patient status was unclear in four studies (Table 2). In included studies, a test was considered positive if it met the following criteria: (1) +1 in the International Knee Documentation Committee parameters; (2) perceivable anterior subluxation of the tibia; (3) no positive endpoint to the subluxation of the knee in 10°–20° of flexion; and (4) the foot did not raise passively in the lever sign test.

The risk of bias assessment of the included studies is described under the headings of the QUADAS-2 assessment tool (Figure 3). For the patient selection domain, only one study had an unclear risk of bias because there was no information regarding consecutive sampling. For the index test domain, one study had a high risk of bias. This study did not state whether the index test was interpreted without the knowledge of the results of the reference test, and it was unclear how this study judged a positive result. For the domain of flow and timing, three studies had a high risk of bias because these studies did not assess the timing between reference and index tests. All applicability concerns in the three domains were judged to have a low risk of bias.

The sensitivity and specificity with 95% CI and forest plots for the four physical examinations for all included studies are shown in Figure 4. Pooled sensitivities for the Lachman, anterior drawer, pivot shift, and lever sign tests were 0.79 (95% CI = 0.74–0.83), 0.78 (95% CI = 0.73–0.83), 0.55 (95% CI = 0.49–0.60), and 0.82 (95% CI = 0.76–0.87), respectively. Pooled specificities were 0.91 (95% CI = 0.87–0.94), 0.91 (95% CI = 0.86–0.95), 0.96 (95% CI = 0.91–0.98), and 0.88 (95% CI = 0.82–0.95), respectively. Forest plots for LR+ and LR− are shown in Figure 5. The lever
Table 2. Summary of characteristics of the included studies

<table>
<thead>
<tr>
<th>Source (year)</th>
<th>No. of subjects</th>
<th>Male</th>
<th>Female</th>
<th>Age years mean (range)</th>
<th>Assessor</th>
<th>Anesthesia</th>
<th>Reference test</th>
<th>Index test</th>
<th>TP</th>
<th>TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jarbo, 2017</td>
<td>102</td>
<td>58</td>
<td>44</td>
<td>23 (15-66)</td>
<td>Medical doctor</td>
<td>29 patients under anesthesia 15 patients while awake</td>
<td>MRI</td>
<td>Lachman test</td>
<td>40</td>
<td>56</td>
</tr>
<tr>
<td>Lee, 1988</td>
<td>79</td>
<td>NA</td>
<td>NA</td>
<td>NA (14-59)</td>
<td>Medical doctor</td>
<td>NA</td>
<td>MRI</td>
<td>Lachman test</td>
<td>21</td>
<td>56</td>
</tr>
<tr>
<td>Massey, 2017</td>
<td>91</td>
<td>61</td>
<td>30</td>
<td>28 (NA)</td>
<td>Medical doctor</td>
<td>while awake</td>
<td>MRI</td>
<td>Lachman test</td>
<td>63</td>
<td>17</td>
</tr>
<tr>
<td>Gürpınar, 2019</td>
<td>78</td>
<td>69</td>
<td>9</td>
<td>26 (17-44)</td>
<td>Medical doctor</td>
<td>while awake</td>
<td>MRI</td>
<td>Lachman test</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>Kıyak, 2018</td>
<td>62</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Medical doctor</td>
<td>while awake</td>
<td>MRI</td>
<td>Lachman test</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>Mulligan, 2015</td>
<td>45</td>
<td>21</td>
<td>24</td>
<td>40 (20-64)</td>
<td>Physical therapist</td>
<td>NA</td>
<td>MRI</td>
<td>Lachman test</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>Mulligan, 2017</td>
<td>60</td>
<td>38</td>
<td>22</td>
<td>42 (18-65)</td>
<td>Physical therapist</td>
<td>NA</td>
<td>MRI</td>
<td>Lachman test</td>
<td>9</td>
<td>26</td>
</tr>
<tr>
<td>Kostov, 2014</td>
<td>103</td>
<td>81</td>
<td>22</td>
<td>29 (16-58)</td>
<td>Medical doctor</td>
<td>NA</td>
<td>MRI</td>
<td>Lachman test</td>
<td>67</td>
<td>30</td>
</tr>
</tbody>
</table>

TP, true positive; TN, true negative; MRI, magnetic resonance imaging.
The lever sign test had the lowest LR− (0.21) and the pivot shift test had the highest LR+ (11.60). The SR OC curves for the four physical examinations are shown in Figure 6. The AUC of the four physical examinations were > 0.70.

**DISCUSSION**

This systematic review included eight studies. The risk of bias of the included studies was evaluated using the QUADAS-2 tool. For the domains of patient selection, the index test, and the reference standard, the risk of bias was low. For the domain of flow and timing, three studies were evaluated as having a high risk of bias because they did not assess the timing between the reference and index tests. The present study included two new studies that were not included in previous systematic reviews. Thus, this study provides additional evidence and emphasize the method healthcare providers should select for ACL injury diagnosis in an acute setting in adolescents and young adults.

The AUC values of the Lachman, anterior drawer, pivot shift, and lever sign tests were large enough (0.9270, 0.9193, 0.8910, and 0.9038, respectively) for the physical examination tests to be considered effective in diagnosing ACL injuries. Furthermore, the pooled sensitivity was highest in the lever sign test, and the pooled specificity was highest in the pivot shift test. The lever sign test had the lowest LR− (0.21) and the pivot shift test had the highest LR+ (11.60). In general, LR+ above 10 and LR− below 0.1 are considered provide strong evidence to rule in or rule out diagnoses, respectively, under most circumstances. Based on the results, the lever sign test is useful to rule out among the tests, and the pivot shift test is useful to confirm an ACL injury in an acute setting.

Previous systematic reviews have assessed the diagnostic accuracy of physical examination tests for ACL injuries. In a previous systematic review in 2016, it was elucidated that the pivot shift test was the most specific, showing a pooled specificity of 97.5% (95% CI = 0.95–0.99), and had the highest LR+ of 16.00 (95% CI = 7.34–34.87). Reiman et al. conducted a systematic review of diagnostic accuracy regarding the lever sign test and found that LR+ and LR− were 9.2 (95% CI = 0.70–46.1) and 0.58 (95% CI = 0.18–1.28), respectively. This study supports these findings, and the pivot shift and lever sign tests are proficient tools for the diagnosis of ACL injuries in acute settings. Although LR+ and LR− are sufficient to diagnose ACL injuries, it is difficult to make a complete diagnosis using a single physical examination test. Therefore, a combination of these four physical examination tests with high diagnostic accuracy may help reduce the number of incorrect diagnoses in clinical practice.

Among the four physical examinations, the lever sign test was introduced as a new physical examination test in the 2010s. This test is relatively easy to perform and requires less examiner strength than the other three commonly used tests. The lever sign test uses the common force of gravity instead of relying on the sensation of the examiner, thus increasing test accuracy. The results of the current study revealed that the diagnostic accuracy of the lever sign test is favorable. Furthermore, Lichtenberg et al. examined the reliability of four physical examinations and indicated that the kappa values for the lever sign test were the highest among the physical examinations. This is a notable point for considering the use of the lever sign test in clinical practice.

This systematic review and meta-analysis have some limitations. First, this study searched for both English and Japanese articles in various databases; however, EMBASE, a database frequently used in other systematic reviews, could not be used. Furthermore, this systematic review focused only on diagnosis in acute conditions and excluded chronic conditions; thus, only eight studies were included. There may be a difference in the included studies and results between other systematic reviews and those in the present study. Additionally, this study did not include subjects with meniscal injury and laxity. These factors may limit the generalizability of the results. Second, this systematic review did not consider the relationship between anesthesia and test result, and included patients with complete or partial
Figure 4. Forest plots of sensitivity and specificity
Lachman test

Anterior drawer test

Pivot shift test

Lever sign test

Figure 5. Forest plots of positive/negative likelihood ratio
ACL injuries. In particular, anesthesia might be related to test accuracy due to complete relaxation of the patient. Therefore, clinicians should be cautious about its clinical applications. Additionally, two types of reference standards, MRI and arthroscopy, were used, which may have led to mixing bias.

CONCLUSION

The AUC of the Lachman, anterior drawer, pivot shift, and lever sign tests suggests high accuracy for use of each in the diagnosis of ACL injury. The lever sign test had the lowest LR− and the pivot shift test had the highest LR+. Thus, the lever sign test is recommended to rule out among the tests, and the pivot shift test is recommended to confirm an ACL injury in an acute setting.

COI STATEMENT

The authors have no conflicts of interest. This study did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

Submitted: December 05, 2021 CDT, Accepted: May 02, 2022 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


