[RESEARCH REPORT]

BARBARA A.M. SNOEKER, MSc[1] • **ROBERT LINDEBOOM,** PhD[1] • **AEILKO H. ZWINDERMAN,** PhD[1]
PATRICE W.J. VINCKEN, MD, PhD[2] • **JORIS A. JANSEN,** MD[3] • **CEES LUCAS,** PhD[1]

Detecting Meniscal Tears in Primary Care: Reproducibility and Accuracy of 2 Weight-Bearing Tests and 1 Non–Weight-Bearing Test

STUDY DESIGN: Diagnostic accuracy study using a cross-sectional design.

OBJECTIVES: To determine the interexaminer reliability and the diagnostic accuracy in primary care of 1 existing weight-bearing meniscal test, the Thessaly test, 1 new weight-bearing test, the deep squat test, and 1 non–weight-bearing test, the joint-line tenderness test.

BACKGROUND: Meniscal tears are difficult to detect in primary care. Although valuable in secondary care, weight-bearing physical examination tests require validation in primary care in unselected patients.

METHODS: Between October 2009 and December 2013, 121 patients (age range, 18-65 years) seen in primary care and suspected of having internal derangement of the knee of less than 6 months in duration were included in the study. Diagnostic accuracy of the 3 meniscal tests was determined based on assessment with magnetic resonance imaging. The meniscal tests were performed by 3 trained physical therapists, who were not informed about the patient history and magnetic resonance imaging results. Each test was performed independently by 2 of the 3 trained physical therapists in alternating pairs.

RESULTS: The Thessaly test and the deep squat test had a moderate level of interexaminer reliability, with kappas of 0.54 and 0.46, respectively. The joint-line tenderness test had poor interexaminer reliability and was therefore not assessed for diagnostic accuracy. The following results are reported separately for both examiners. The Thessaly test had a sensitivity of 66.7% (95% confidence interval [CI]: 53.0%, 78.0%) and 51.2% (95% CI: 36.8%, 65.4%), a specificity of 37.9% (95% CI: 27.2%, 50.0%) and 43.5% (95% CI: 30.2%, 57.8%), a positive likelihood ratio of 1.07 (95% CI: 0.82, 1.41) and 0.91 (95% CI: 0.62, 1.33), and a negative likelihood ratio of 0.88 (95% CI: 0.54, 1.45) and 1.12 (95% CI: 0.72, 1.76). Similarly, the deep squat test had a sensitivity of 74.5% (95% CI: 61.1%, 84.5%) and 76.7% (95% CI: 62.3%, 86.9%), a specificity of 42.4% (95% CI: 31.2%, 54.4%) and 36.2% (95% CI: 24.0%, 50.5%), a positive likelihood ratio of 1.29 (95% CI: 0.97, 1.68) and 1.20 (95% CI: 0.92, 1.58), and a negative likelihood ratio of 0.60 (95% CI: 0.35, 1.04) and 0.64 (95% CI: 0.33, 1.25).

CONCLUSION: Although the Thessaly and deep squat tests have a moderate level of reliability, neither test is sufficiently accurate to help in the diagnosis of meniscal tears in primary care. Future research should focus on other relevant patient variables instead of on physical examination tests in the detection of meniscal tears.

LEVEL OF EVIDENCE: Diagnosis, level 3b.
J Orthop Sports Phys Ther 2015;45(9):693-702. Epub 10 Jul 2015. doi:10.2519/jospt.2015.5712

KEY WORDS: *diagnosis, knee, tenderness, Thessaly*

In the UK, 15% of general practitioner consultations are related to musculoskeletal disorders.[8] In the Netherlands, knee injuries are the second most common musculoskeletal condition that requires a general practitioner consultation (48 out of 1000 consultations per year). The incidence of consultations related to meniscal tears is 2 per 1000 annually.[10] In the short term, meniscal tears lead to disability in daily functioning, absenteeism from work, and inability to perform sports.[2] In the long term, meniscal tears may lead to knee osteoarthritis.[11] Meniscal tears are difficult to detect in primary care.[14,15,30] More accurate diagnosis of knee disorders in primary care would enable general practitioners to more accurately refer patients with a high index of suspicion of

[1]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands. [2]Department of Radiology, Rijnland Ziekenhuis, Leiderdorp, the Netherlands. [3]Department of Orthopaedics, Rijnland Ziekenhuis, Leiderdorp, the Netherlands. The protocol for the study was approved by the Centrale Commissie Mensgebonden Onderzoek, the Netherlands. The authors certify that they have no affiliations with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the article. Address correspondence to Barbara A.M. Snoeker, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam Zuidoost, the Netherlands. E-mail: b.a.snoeker@amc.uva.nl

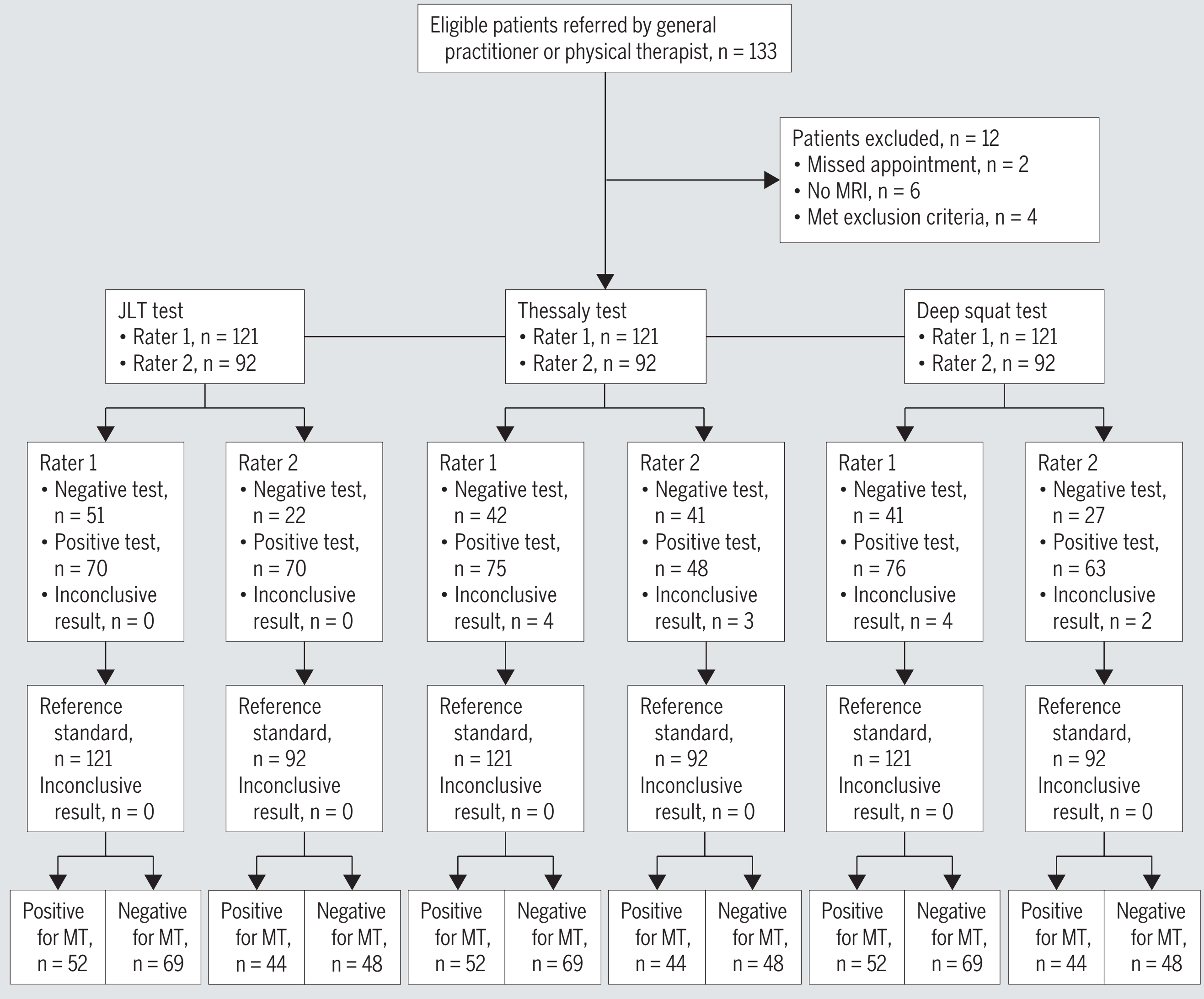

FIGURE 1. Flow chart of the recruitment process. Abbreviations: JLT, joint-line tenderness test; MRI, magnetic resonance imaging; MT, meniscus tear.

internal derangement to secondary care, or to provide conservative treatment for patients with a low suspicion of internal derangement.[7,8] Both strategies aim to get the patients back to their normal level of functioning as quickly as possible.[11,26]

An accurate early diagnosis of meniscal tears is important. The joint-line tenderness test, McMurray test, and Apley test are the clinical tests most often used to detect meniscal tears in primary care.[21] However, several studies have concluded that these non–weight-bearing tests performed by the examiner have limited diagnostic accuracy.[1,21,23,27,28,31] A recent meta-analysis by Ockert et al[23] found that the joint-line tenderness test, McMurray test, and Apley test had a sensitivity of 64%, 51%, and 38%, respectively, and a specificity of 61%, 78%, and 84%, respectively. Radiological confirmation with magnetic resonance imaging (MRI) scan is typically required for accurate diagnosis.[6,25,32] Magnetic resonance imaging has been reported to have a sensitivity of 91.4% and a specificity of 81.1% for the diagnosis of medial meniscal tears, and a sensitivity of 88.8% and specificity of 76% for lateral meniscal tears.[6] While the use of MRI may assist in the diagnosis and management of patients with knee pathologies seen in primary care,[8] MRI is too expensive to be used as the first-option diagnostic procedure. Therefore, clinical tests to help determine the need for imaging are warranted.[7]

The use of more sensitive physical examination tests in primary care could improve the management of patients with knee pathologies by general practitioners.[7,23,30,32] Weight-bearing tests have been proposed to counter the shortcomings of non–weight-bearing tests, in particular their limited sensitivity. With weight-bearing tests, such as the Thessaly test, the menisci are compressed by the patient's body weight while performing rotation of the knee, a movement suggested to mimic a common mechanism of injury for the menisci.[15,29]

Previous studies have shown that the Thessaly test is more reproducible than non–weight-bearing tests, with kappa coefficients exceeding 0.85.[14,15] The same studies reported accuracy measures with sensitivity and specificity greater than 90%.[14,15] However, these studies were on selected patients awaiting arthroscopy, with healthy individuals as controls, and the results are therefore subject to spectrum bias.[18] In primary care, the accuracy of such tests may be lower due to less contrast among study participants, all individuals having knee symptoms/pathologies.[14,18,28]

Therefore, the purpose of this study was to estimate the interexaminer reliability and diagnostic accuracy of 2 weight-bearing tests (the Thessaly test and the deep squat test[14,15,23,28,30]) and 1 non–weight-bearing test (joint-line tenderness test) for the detection of meniscal tears in primary care.

METHODS

Procedure

Using a cross-sectional design, the results of 3 meniscal tests (index tests) were compared to those of MRI (reference test). Twenty-four general practitioners from 12 clinics and 28 physical therapists from 3 clinics from the cities of Leiden, Leiderdorp, and Voorschoten, the Netherlands were asked to refer patients fulfilling inclusion and exclusion criteria to our research centers (**FIGURE 1**). Magnetic resonance imaging generally took place within 2 weeks of performing the meniscal tests, with the maximum duration between clinical tests and MRI being 5 weeks. To avoid partial verification, MRI was performed for every participant, irrespective of the index test results. Depending on planning and scheduling of the medical center performing the imaging, MRI could be performed before or after the meniscal tests.

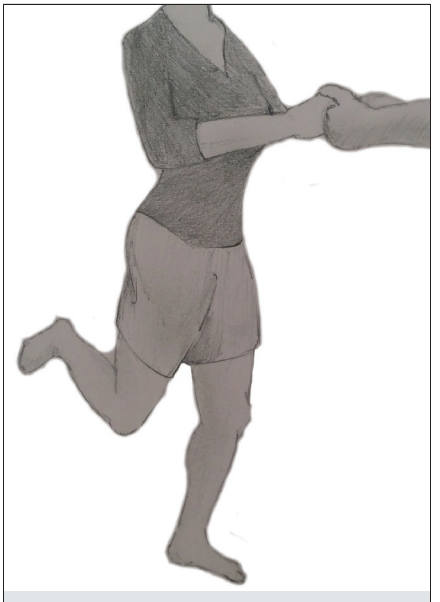

FIGURE 2. Thessaly test.

Participants were screened for inclusion and exclusion criteria and enrolled in the study after giving their written informed consent. Subsequently, baseline demographic characteristics were recorded and the meniscal tests were conducted by 3 trained physical therapists in 1 of the research centers, who were blinded to the participants' baseline characteristics and history of knee symptoms. From the 30th participant onward, each test was performed independently by 2 of the 3 physical therapists in alternating pairs to evaluate reproducibility. At each research center, the tests were performed by the same selected physical therapists. The meniscal tests were performed independently and without knowledge of the test results from the other examiners or knowledge of the MRI results. Test results were recorded on identical but separate knee report forms. The participants were not told whether the test results were positive or negative. For the MRI scan, the radiologist recorded the results on a knee report form that was sent to the primary investigator (B.S.). The radiologist was blinded to patient history and meniscal clinical test results. The Dutch medical ethics authority Centrale Commissie Mensgebonden Onderzoek approved the study protocol and provided ethical approval.

FIGURE 3. Joint-line tenderness test.

Participants

The inclusion criteria were that participants had to be between 18 and 65 years of age, to have knee symptoms for less than 6 months, and to be suspected of having internal derangement of the knee, based on the Dutch general practitioner guidelines for nontraumatic and traumatic knee problems.[3,4] Potential participants could not have had a radiographic evaluation of their knee within the previous 6 months and were excluded if they underwent knee arthroplasty or other surgical interventions for the knee joint in the past. The exclusion criteria also included all MRI contraindications (eg, pacemaker, claustrophobia, intracranial aneurysm clips, or orbital metallic foreign body).

Test Procedure

We evaluated 2 existing tests for meniscal tears (the Thessaly test and joint-line tenderness test) and 1 newly introduced test (the deep squat test), which was specifically designed to have high sensitivity to help in the diagnosis of meniscal tears.

Thessaly Test The Thessaly test is performed with the patient standing flat-footed on the affected limb, with the knee in 20° of flexion (**FIGURE 2**).[15] With the examiner holding the patient's hands, the patient rotates the knee and body 3 times in each direction, while keeping the knee in 20° of flexion. The test is always per-

FIGURE 4. Deep squat test.

formed initially on the unaffected knee, enabling the patient to recognize a possible positive result in the symptomatic knee when questioned about symptoms during and following the test. The test is positive if the patient experiences medial or lateral joint-line discomfort or has a sense of locking or catching.

Joint-Line Tenderness Test The examiner passively positions the knee joint in 90° of flexion (**FIGURE 3**) before systematically palpating the tibiofemoral joint line for the presence of tenderness.[12] For the medial meniscus, palpation is performed starting near the medial border of the patellar tendon, following the tibiofemoral joint line toward the posterior aspect of the knee. The same technique is used on the lateral aspect of the knee for the lateral meniscus. A positive test result is defined as a local area of tenderness that exceeds normal discomfort when compared to the unaffected knee.

Deep Squat Test The deep squat test is based on the biomechanical theory that a deep squat increases the compression of the posterior horn of the meniscus, causing knee pain in the presence of a torn meniscus (**FIGURE 4**).[20,28] The deep squat test is performed with the patient standing flatfooted on the floor. The patient squats as deeply as possible while the examiner holds the patient's hands for balance. The test is considered positive if the patient experiences internal knee pain during flexion or a sense of locking.

TABLE 1 PARTICIPANT BASELINE CHARACTERISTICS*

Characteristic	Overall (n = 121)	Meniscal Tear Based on MRI (n = 52)	No Meniscal Tear Based on MRI (n = 69)	*P* Value
Age, y	43.2 ± 12.2	46.4 ± 11.4	40.8 ± 12.2	.01
Sex (male), n (%)	68 (56)	39 (75)	29 (42)	<.001
Mass, kg	81.8 ± 15.2	85.0 ± 11.9	79.2 ± 17.0	.03
Height, m	1.76 ± 0.10	1.80 ± 0.08	1.74 ± 0.10	.003
Body mass index, kg/m^2	26.1 ± 4.0	26.3 ± 3.0	25.9 ± 4.6	.69
Pain severity VAS†	49.4 ± 25.9	42.5 ± 24.3	54.6 ± 26.1	.01
Duration of symptoms, wk‡	7.0 (0.5-27.0)	7.5 (0.5-27.0)	5.0 (1.0-27.0)	.90
Symptom side (right), n (%)	62 (51)	27 (52)	35 (51)	.78
Performing sports (yes), n (%)	93 (77)	44 (85)	49 (71)	.08
Weight bearing during trauma (yes), n (%)	41 (34)	25 (48)	16 (23)	.001
Work-related kneeling or squatting more than 1 h/d (yes), n (%)	34 (28)	17 (33)	17 (25)	.33
Pain during passive extension (yes), n (%)	79 (65)	36 (69)	43 (62)	.431
Mechanism of injury, n (%)				.775
Unknown	42 (35)	14 (27)	28 (41)	
(Rotation) trauma	48 (40)	24 (46)	24 (35)	
After sports	31 (26)	14 (27)	17 (25)	
Effusion, n (%)§				.068
No effusion	44 (36)	11 (21)	33 (48)	
Effusion+	66 (55)	36 (69)	30 (44)	
Effusion++	11 (9)	5 (10)	6 (9)	
Locking (yes), n (%)	30 (25)	8 (15)	22 (32)	.04
Giving way (yes), n (%)	70 (58)	32 (62)	38 (55)	.47

Abbreviations: MRI, magnetic resonance imaging; VAS, visual analog scale.
**Values are mean ± SD unless otherwise indicated.*
†Zero means no pain and 100 means worst pain possible for pain in the previous week.
‡Values are median (minimum-maximum).
§Effusion+, minimal displacement of effusion with ballottement test; effusion++, significant displacement of effusion with ballottement test.

MRI as Reference Standard

The MRI scans were performed in 2 different hospitals, the Rijnland Ziekenhuis in Leiderdorp and the Rode Kruis Ziekenhuis in Beverwijk, the Netherlands. The scanning protocol at the Rijnland Ziekenhuis, using a 1.5-T Achieva or Ingenia system (Koninklijke Philips NV, Eindhoven, the Netherlands), consisted of a sagittal T2-weighted sequence, a sagittal proton density (PD)–weighted sequence, a coronal PD-weighted sequence, and a coronal PD-weighted sequence with fat suppression. The scanning protocol at the Rode Kruis Zieken-

TABLE 2 Results of Magnetic Resonance Imaging (n = 121)

Characteristic	Value*
Median (minimum-maximum) duration of symptoms, wk	7 (0.5-27)
Median (minimum-maximum) duration between meniscal clinical tests and MRI, d	4 (–25 to 38)
Meniscal tears†	52 (43.0)
Medial meniscus	41 (78.8)
Lateral meniscus	11 (21.2)
Anterior horn	5 (9.6)
Posterior horn	41 (78.8)
Combination of CL and meniscal tear	9 (17.3)
Combination of chondropathy and meniscal tear	20 (38.4)
CL tears‡	14 (11.6)
Isolated ACL tears	3 (21.4)
Isolated PCL tears	2 (14.3)
Chondropathy in medial and lateral compartment	31 (25.6)
Participants with chondropathy for all those without meniscal tears	16 (23.2)
Patellofemoral chondropathy	7 (5.8)
No knee pathology	42 (34.7)

Abbreviations: ACL, anterior cruciate ligament; CL, cruciate ligament; MRI, magnetic resonance imaging; PCL, posterior cruciate ligament.
**Values are n (%) unless otherwise indicated.*
†Participants could have more than 1 condition.
‡Nine of the 14 participants had concurrent injuries.

huis, using a 1.5-T Philips Intera system (Koninklijke Philips NV), consisted of a sagittal PD-weighted sequence and an axial and coronal PD-weighted sequence with fat suppression. Two radiologists with more than 8 years of experience with musculoskeletal MRI interpreted all MRI films.

A positive diagnosis for meniscal pathology was based on the appearance of any meniscal tear or detachment of the meniscus from its supportive structures. If no abnormalities on MRI were found, this was classified as a normal, intact meniscus. The radiologist was not informed of the findings from the physical examination.

Statistical Analysis

Absolute and relative frequencies were used to summarize categorical patient characteristics. Normally distributed numerical data were summarized using means and standard deviations. Data not normally distributed were summarized using medians and interquartile scores.

To evaluate interexaminer reliability, kappa statistics were calculated separately for each test.[5] In addition, to determine the level of agreement between the 3 clinical tests, kappa statistics comparing positive and negative results of one test versus the results of another test were calculated based on the results of the first examiner.[5] The agreement between test procedures was performed for all test combinations.

Only the clinical tests with at least a moderate level of interexaminer agreement (κ>0.40) were analyzed further for diagnostic accuracy.[5,24] Diagnostic accuracy of the individual tests was evaluated against the MRI results and quantified using sensitivity and specificity rates,[9,16] which were calculated separately for the medial and lateral meniscus, as well as for both menisci combined. Positive and negative predictive values and likelihood ratios were used to compute posttest probability for a meniscal tear.[9,16] As test results were obtained from 2 examiners, the sensitivity, specificity, and predictive values for each examiner were calculated separately. Diagnostic odds ratios (DORs) for overall accuracy of a meniscal test were obtained using univariate logistic regression analysis.[9,16] All analyses were performed using the statistical program SPSS Version 20.0 (SPSS Inc, Chicago, IL).

RESULTS

Study Population

We recruited 121 participants between October 2009 and December 2013 (**FIGURE 1**). Baseline characteristics of those participants are summarized in TABLE 1. Magnetic resonance imaging results indicated that there were 69 participants without and 52 with a meniscal tear. Of the 93 participants who engaged in sporting activities, 31 (33%) were involved in contact sports. Thirty-four (28%) participants performed work activities that involved more than 1 hour per day of kneeling or squatting. Seventy-one participants (59%) had a body mass index greater than 25 kg/m^2. The data in **TABLE 1** show statistically significant differences between those with and without meniscal tears for age (46.4 versus 40.8 years), sex (75% versus 42% male), mass (85.0 versus 79.2 kg), height (1.80 versus 1.74 m), pain severity (42.5 versus 54.6 on a 100-mm visual analog scale), engaging in sports (84.6% versus 71.0%), weight bearing during trauma (48.1% versus 23.2%), and a history of knee locking (15.4% versus 31.9%). The median duration of symptoms before participants consulted a general practitioner or physical therapist was 7 weeks (range, 0.5-27 weeks).

MRI Results

The median time between the clinical meniscal tests and the MRI was 4 days, with the MRI as many as 25 days before to 38 days after the clinical examination.

The MRI scan revealed that 41 participants had medial meniscal tears (78.8% of the tears) and 11 participants had lateral meniscal tears (21.2% of the tears). Of all participants with meniscal tears, 9 (17.3%) also had a cruciate ligament tear. Five of the 14 participants with a cruciate ligament tear had an isolated injury to the cruciate ligament (**TABLE 2**).

Reproducibility of Meniscal Tests

As indicated in the Methods section, the second examiner was added starting with the 30th participant. In addition, for the Thessaly and deep squat tests, the results for 4 participants could not be interpreted and were consequently excluded. Therefore, interexaminer reliability results for the joint-line tenderness test were from 92 participants and those for the Thessaly and deep squat tests were from 88 participants (**TABLES 3** through **6**). The Thessaly test had the highest interexaminer reliability (κ = 0.54), followed by the deep squat test (κ = 0.46). The joint-line tenderness test had poor reliability (κ = 0.17).

TABLE 3 Interexaminer Reliability of the Clinical Tests

Test	Kappa*
Thessaly test (n = 88)	0.54 (0.37, 0.72)
Deep squat test (n = 88)	0.46 (0.26, 0.65)
Joint-line tenderness test (n = 92)	0.17 (−0.02, 0.36)

**Values in parentheses are 95% confidence interval.*

TABLE 4 Interexaminer Agreement for the Thessaly Test (n = 88)

	Positive Thessaly: Examiner 1	Negative Thessaly: Examiner 1	Total
Positive Thessaly: examiner 2	38	11	49
Negative Thessaly: examiner 2	9	30	39
Total	47	41	88

TABLE 5 Interexaminer Agreement for the Deep Squat Test (n = 88)

	Positive Deep Squat: Examiner 1	Negative Deep Squat: Examiner 1	Total
Positive deep squat: examiner 2	49	9	58
Negative deep squat: examiner 2	12	18	30
Total	61	27	88

Agreement Between Clinical Tests

The level of agreement between clinical tests was based on the results from the first examiner, with a total of 114 and 117 participants. The level of agreement between the clinical tests was poor (**TABLE 7**).

Diagnostic Accuracy of Meniscal Tests

The levels of agreement between MRI test results and meniscal tear test results, with their associated test accuracy parameters (sensitivity, specificity, positive predictive values, negative predictive values, and DORs), are provided in **TABLES 8** and **9** for both examiners separately.

For the Thessaly test for both meniscal tears, sensitivity was 66.7% (95% confidence interval [CI]: 53.0%, 78.0%) and 51.2% (95% CI: 36.8%, 65.4%), specificity was 37.9% (95% CI: 27.2%, 50.0%) and 43.5% (95% CI: 30.2%, 57.8%), positive likelihood ratio was 1.07 (95% CI: 0.82, 1.41) and 0.91 (95% CI: 0.62, 1.33), and negative likelihood ratio was 0.88 (95% CI: 0.54, 1.45) and 1.12 (95% CI: 0.72, 1.76) for examiners 1 and 2, respectively. For a medial meniscal tear, sensitivity was 65.9% (95% CI: 51.1%, 78.1%) for examiner 1 and 50.0% (95% CI: 34.5%, 65.5%) for examiner 2; for a lateral meniscal tear, sensitivity was 72.7% (95% CI: 43.4%, 90.3%) for examiner 1 and 54.5% (95% CI: 28.0%, 78.7%) for examiner 2. Given a positive clinical test result, for both meniscal tears, the probability increased from 43.6% to 45.3% for examiner 1 and decreased from 48.3% to 45.8% for examiner 2. Given a negative clinical test result, the probability for a meniscal tear decreased to 40.5% for examiner 1 and increased to 51.2% for examiner 2.

For the deep squat test for both meniscal tears, sensitivity was 74.5% (95% CI: 61.1%, 84.5%) and 76.7% (95% CI: 62.3%, 86.9%), specificity was 42.4% (95% CI: 31.2%, 54.4%) and 36.2% (95% CI: 24.0%, 50.5%), positive likelihood ratio was 1.29 (95% CI: 0.97, 1.68) and 1.20 (95% CI: 0.92, 1.58), and negative likelihood ratio was 0.60 (95% CI: 0.35, 1.04) and 0.64 (95% CI: 0.33, 1.25) for examiners 1 and 2, respectively. For a medial meniscal tear, sensitivity was 72.7% (95% CI: 58.1%, 83.7%) for examiner 1 and 72.2% (95% CI: 56.0%, 84.2%) for examiner 2; for a lateral meniscal tear, sensitivity was 81.8% (95% CI: 52.3%, 94.9%) for examiner 1 and 90.9% (95% CI: 62.3%, 98.4%) for examiner 2. Given a positive clinical test result, for both meniscal tears, the probability increased from 43.5% to 50.0% for examiner 1 and from 47.7% to

TABLE 6 Interexaminer Agreement for the Joint-Line Tenderness Test (n = 92)

	Positive JLT: Examiner 1	Negative JLT: Examiner 1	Total
Positive JLT: examiner 2	44	9	53
Negative JLT: examiner 2	26	13	39
Total	70	22	92

Abbreviation: JLT, joint-line tenderness test.

TABLE 7 Level of Agreement Between Clinical Tests*

Tests	Kappa†	Agreement
Deep squat versus Thessaly test (n = 114)	0.10 (–0.09, 0.29)	Poor
Deep squat versus joint-line tenderness test (n = 117)	0.14 (–0.04, 0.32)	Poor
Joint-line tenderness test versus Thessaly test (n = 117)	0.18 (0.00, 0.36)	Poor

**Results based on the first examiner.*
†Values in parentheses are 95% confidence interval.

52.4% for examiner 2. Given a negative clinical test result, the probability of a meniscal tear decreased to 31.7% for examiner 1 and 37.0% for examiner 2.

In univariate analysis, DOR was measured for the association between the test result and the presence of a meniscal tear (**TABLES 8** and **9**). The deep squat test was not significantly associated with the presence of a meniscal tear (DOR = 2.15; 95% CI: 0.97, 4.78 for examiner 1 and DOR = 1.87; 95% CI: 0.74, 4.71 for examiner 2).

DISCUSSION

Our study suggests that 2 weight-bearing tests for meniscal tears, the Thessaly test and the deep squat test, have moderate interexaminer reliability. Because the joint-line tenderness test has poor interexaminer reliability, it was not further analyzed for accuracy.[24] The level of agreement between the Thessaly and deep squat tests was poor.

With low diagnostic accuracy, as reflected by positive likelihood ratios and negative likelihood ratios around 1, the Thessaly test could not predict a meniscal tear for patients seen in primary care. The likelihood ratios of the deep squat test indicated only a small and nonsignificant change in the probability that participants had or did not have a meniscal tear after a positive or negative test result. Therefore, the deep squat test alone is not sufficiently accurate to detect or rule out meniscal tears in primary care.[9]

Contrary to previous studies that have reported sensitivity and specificity rates exceeding 90%,[14,15] we observed a much lower sensitivity and specificity for the weight-bearing Thessaly test in our primary care sample (less than 68% and less than 44%, respectively). Differences may be explained by the spectrum of participants in these earlier studies. Differences in patient spectrum and referral filter may lead to overestimation of the sensitivity and specificity of a test.[18] In these previous studies, the knees of individuals already awaiting arthroscopy were compared to the knees of healthy individuals[15] or to the healthy, noninjured knee of the injured participants,[14] which might have resulted in overestimation of the test accuracy.[18] Later studies performed in secondary care reported lower sensitivity and specificity rates for the Thessaly test.[13,17,22] In our primary care setting with unselected patients, we observed substantially lower accuracy of the meniscal tests.

Compared to the Thessaly test, the deep squat test was more sensitive as a screening test for meniscal tears in our primary care sample. However, with small and nonsignificant likelihood ratios and a nonsignificant DOR, the deep squat test seemed insufficient to detect or rule out meniscal tears.[9] Various authors have suggested that a single test cannot predict meniscal tears and that a combination of tests should be used.[17,19,26-28] Unfortunately, the modest accuracy and reproducibility of the 2 tests observed in our study suggest that combining the Thessaly and deep squat tests would not improve prediction.

To improve prediction, variables that describe important differences between participants with and without meniscal tears from this study could be used for that purpose. For example, participants with and without meniscal tears differed in age, sex, body mass index, sports participation, and weight-bearing status during trauma. These findings are consistent with the results of an earlier systematic review on the risk factors for meniscal tears.[29] In contrast to the findings of the previously published review, the present study found that work-related kneeling and squatting were equally reported by participants with and without meniscal tears. This may be because no information on the duration in years of work-related kneeling and squatting was collected and, therefore, could not be adjusted for.

There are several limitations in this study. First, MRI was used as a reference standard, not arthroscopy, as it was unethical to perform arthroscopy in unselected individuals in primary care. Although arthroscopy is more accurate, we considered MRI to be a valid alternative because of its reported accuracy.[6] However, the use of MRI could have led to more false-positive and false-negative

TABLE 8 Diagnostic Accuracy Statistics for the Thessaly Test*

	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Diagnostic OR	Positive LR	Negative LR
Rater 1 (n = 117)	66.7 (53.0, 78.0)	37.9 (27.2, 50.0)	45.3 (34.6, 56.6)	59.5 (44.5, 73.0)	1.22 (0.57, 2.62)	1.07 (0.82, 1.41)	0.88 (0.54, 1.45)
Rater 2 (n = 89)	51.2 (36.8, 65.4)	43.5 (30.2, 57.8)	45.8 (32.6, 59.7)	48.7 (34.3, 63.5)	0.81 (0.35, 1.86)	0.91 (0.62, 1.33)	1.12 (0.72, 1.76)

Abbreviations: LR, likelihood ratio; NPV, negative predictive value; OR, odds ratio; PPV, positive predictive value.
**Values in parentheses are 95% confidence interval.*

TABLE 9 Diagnostic Accuracy Statistics for the Deep Squat Test*

	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Diagnostic OR	Positive LR	Negative LR
Rater 1 (n = 117)	74.5 (61.1, 84.5)	42.4 (31.2, 54.4)	50.0 (39.0, 61.0)	68.3 (53.0, 80.4)	2.15 (0.97, 4.78)	1.29 (0.97, 1.68)	0.60 (0.35, 1.04)
Rater 2 (n = 90)	76.7 (62.3, 86.9)	36.2 (24.0, 50.5)	52.4 (40.3, 64.2)	63.0 (44.2, 78.5)	1.87 (0.74, 4.71)	1.20 (0.92, 1.58)	0.64 (0.33, 1.25)

Abbreviations: LR, likelihood ratio; NPV, negative predictive value; OR, odds ratio; PPV, positive predictive value.
**Values in parentheses are 95% confidence interval.*

results. To evaluate the influence of MRI on the interpretation of the results, we requested arthroscopy reports of all participants who underwent knee surgery. Of 41 identifiable participants who received arthroscopy after MRI (34%), only 4 had a meniscal tear identified with arthroscopy that was not identified with MRI. For 3 of these 4 participants, the deep squat test yielded positive test results, and for 1 participant the Thessaly test was positive. All participants who underwent arthroscopy had a meniscal tear.

Second, the assessment of pain could have limited the findings. A typical finding was that participants with meniscal tears reported less pain at baseline than those without meniscal tears. Those with meniscal tears had a longer duration of pain (median, 7.5 weeks versus 5 weeks). As participants were asked to rate their pain in the previous week on the visual analog scale, their pain might have already diminished at the time of inclusion. Therefore, most participants with meniscal tears who were included in this study had a prolonged duration of knee symptoms. Another typical finding was that relatively fewer participants with meniscal tears experienced locking of the knee than participants without meniscal tears, which is in contrast to other reports.[19,31] One possible explanation is that participants might not have interpreted the question correctly during history taking. Another possible explanation is that the finding was due to chance.

Third, we asked general practitioners and physical therapists to refer participants according to the inclusion and exclusion criteria but could not be sure if the referrals were consecutive. For example, some participants with more severe knee complaints might have been directly referred for MRI by their general practitioner, and were therefore not included in our study. This might have led to a lower contrast between participants with and without meniscal tears, and therefore to an underestimation of test accuracy.

Due to the planning and organization of the involved hospital, the MRI scan could have been performed before the participants underwent the meniscal tests. Although we asked participants to remain unaware of their MRI results before undergoing the meniscal tests, 2 of them were aware of their MRI results before the meniscal tests. The examiner who performed the meniscal tests and the investigator who recorded the clinical characteristics were blinded to the MRI results.

For some participants, the MRI could have been performed more than 2 weeks before or after the meniscal clinical tests. The decision to limit the time of the MRI to within 2 weeks before or after the meniscal test was to prevent biased test results due to disease progression. However, the median delay for receiving MRI was only 4 days, and the few participants who had longer delays could not have influenced the results significantly.

The results of one test could have been influenced by another test. If a test provoked symptoms on the affected knee, the examiner who performed the test had to wait until the symptoms resolved (eg, pain, locking) before the second examiner performed a test. However, we cannot rule out that the provocation of one test could have led to a positive result on another test. Nevertheless, this was equally likely for every test result, as tests were randomly performed.

Finally, for the reliability analysis, our sample size was reduced, as the second examiner for every test was added from the 30th participant. So, 88 participants had all tests performed effectively by 2 examiners. The decision to analyze results for the 2 examiners separately resulted in wider CIs than would have been observed with pooled results. Precision would have improved if analyses had been performed on the pooled data.

CONCLUSION

Our study suggests that 2 weight-bearing tests for meniscal tears, the Thessaly test and the deep squat test, have moderate interexaminer reliability. Agreement between these tests was poor. For the Thessaly test as well as the deep squat test, diagnostic accuracy values were poor, which indicated that both tests were not sufficiently accurate to detect or rule out meniscal tears in primary care. Future research should focus on other relevant patient variables instead of physical examination tests in the detection of meniscal tears. Better detection of meniscal tears could improve the management of patients with knee pathologies in primary care. Therefore, a prediction model will be developed and validated by our research group. ◉

KEY POINTS

FINDINGS: Both the Thessaly test and the deep squat test demonstrated insufficient diagnostic accuracy for use in primary care. The joint-line tenderness test had poor interexaminer reliability, and diagnostic accuracy of this test was therefore not evaluated. The level of agreement between the 3 tests was poor.

IMPLICATIONS: Current clinical tests for the diagnosis of meniscal tears have limited clinical utility.

CAUTION: In the present study, MRI, instead of arthroscopy, was used as the reference test to diagnose meniscal tears. This could have led to more false-positive and false-negative results. Also, we cannot be sure that referrals were consecutively sent to our research center. As individuals with more acute knee symptoms might have been directly referred for MRI without inclusion into this study, this could have led to an underestimation of test results.

REFERENCES

1. Akseki D, Özcan O, Boya H, Pinar H. A new weight-bearing meniscal test and a comparison with McMurray's test and joint line tenderness. *Arthroscopy*. 2004;20:951-958. http://dx.doi.org/10.1016/j.arthro.2004.08.020
2. Baker P, Coggon D, Reading I, Barrett D, McLaren M, Cooper C. Sports injury, occupational physical activity, joint laxity, and meniscal damage. *J Rheumatol*. 2002;29:557-563.
3. Belo JN, Berg HF, Klein Ikkink AJ, Wildervanck-Dekker CMJ, Smorenburg HAAJ, Draijer LW. NHG-Standaard Traumatische knieproblemen (Eerste herziening). *Huisarts Wet*. 2010;54:147-158.
4. Belo JN, Bierma-Zeinstra SMA, Raaijmakers AJ, Van der Wissel F, Opstelten W. NHG-Standaard Niet-traumatische knieproblemen bij volwassenen (Eerste herziening). *Huisarts Wet*. 2008;51:229-240.
5. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
6. Crawford R, Walley G, Bridgman S, Maffulli N. Magnetic resonance imaging versus arthroscopy in the diagnosis of knee pathology, concentrating on meniscal lesions and ACL tears: a systematic review. *Br Med Bull*. 2007;84:5-23. http://dx.doi.org/10.1093/bmb/ldm022
7. DAMASK Trial Team. Cost-effectiveness of magnetic resonance imaging of the knee for patients presenting in primary care. *Br J Gen Pract*. 2008;58:e10-e16. http://dx.doi.org/10.3399/bjgp08X342660
8. DAMASK Trial Team. Effectiveness of GP access to magnetic resonance imaging of the knee: a randomised trial. *Br J Gen Pract*. 2008;58:e1-e9. http://dx.doi.org/10.3399/bjgp08X342651
9. Davidson M. The interpretation of diagnostic test: a primer for physiotherapists. *Aust J Physiother*. 2002;48:227-232.
10. Draijer LW, Belo JN, Berg HF, Geijer RM, Goudswaard AN. [Summary of the practice guideline 'Traumatic knee problems' (first revision) from the Dutch College of General Practitioners]. *Ned Tijdschr Geneeskd*. 2010;154:A2225.
11. Englund M, Roemer FW, Hayashi D, Crema MD, Guermazi A. Meniscus pathology, osteoarthritis and the treatment controversy. *Nat Rev Rheumatol*. 2012;8:412-419. http://dx.doi.org/10.1038/nrrheum.2012.69
12. Eren OT. The accuracy of joint line tenderness by physical examination in the diagnosis of meniscal tears. *Arthroscopy*. 2003;19:850-854. http://dx.doi.org/10.1016/S0749-8063(03)00736-9
13. Goossens P, Keijsers E, van Geenen RJ, et al. Validity of the Thessaly test in evaluating meniscal tears compared with arthroscopy: a diagnostic accuracy study. *J Orthop Sports Phys Ther*. 2015;45:18-24. http://dx.doi.org/10.2519/jospt.2015.5215
14. Harrison BK, Abell BE, Gibson TW. The Thessaly test for detection of meniscal tears: validation of a new physical examination technique for primary care medicine. *Clin J Sport Med*. 2009;19:9-12. http://dx.doi.org/10.1097/JSM.0b013e31818f1689
15. Karachalios T, Hantes M, Zibis AH, Zachos V, Karantanas AH, Malizos KN. Diagnostic accuracy of a new clinical test (the Thessaly test) for early detection of meniscal tears. *J Bone Joint Surg Am*. 2005;87:955-962. http://dx.doi.org/10.2106/JBJS.D.02338
16. Knottnerus JA, Buntinx F. *The Evidence Base of Clinical Diagnosis: Theory and Methods of Diagnostic Research*. 2nd ed. Chichester, UK: Wiley-Blackwell; 2009.
17. Konan S, Rayan F, Haddad FS. Do physical diagnostic tests accurately detect meniscal tears? *Knee Surg Sports Traumatol Arthrosc*. 2009;17:806-811. http://dx.doi.org/10.1007/s00167-009-0803-3
18. Leeflang MM, Rutjes AW, Reitsma JB, Hooft L, Bossuyt PM. Variation of a test's sensitivity and specificity with disease prevalence. *CMAJ*. 2013;185:E537-E544. http://dx.doi.org/10.1503/cmaj.121286
19. Lowery DJ, Farley TD, Wing DW, Sterett WI, Steadman JR. A clinical composite score accurately detects meniscal pathology. *Arthroscopy*. 2006;22:1174-1179. http://dx.doi.org/10.1016/j.arthro.2006.06.014
20. Lustig S, Scholes CJ, Balestro JC, Parker DA. In vivo assessment of weight-bearing knee flexion reveals compartment-specific alterations in meniscal slope. *Arthroscopy*. 2013;29:1653-1660. http://dx.doi.org/10.1016/j.arthro.2013.07.260
21. Meserve BB, Cleland JA, Boucher TR. A meta-analysis examining clinical test utilities for assessing meniscal injury. *Clin Rehabil*. 2008;22:143-161. http://dx.doi.org/10.1177/0269215507080130
22. Mirzatolooei F, Yekta Z, Bayazidchi M, Ershadi S, Afshar A. Validation of the Thessaly test for detecting meniscal tears in anterior cruciate deficient knees. *Knee*. 2010;17:221-223. http://dx.doi.org/10.1016/j.knee.2009.08.007
23. Ockert B, Haasters F, Polzer H, et al. [Value of the clinical examination in suspected meniscal injuries. A meta-analysis]. *Unfallchirurg*. 2010;113:293-299. http://dx.doi.org/10.1007/s00113-009-1702-2
24. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice*. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall; 2008.
25. Rosas HG, De Smet AA. Magnetic resonance imaging of the meniscus. *Top Magn Reson Imaging*. 2009;20:151-173. http://dx.doi.org/10.1097/RMR.0b013e3181d657d1
26. 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;455:123-133.
27. Scholten RJ, Devillé WL, Opstelten W, Bijl D, van der Plas CG, Bouter LM. The accuracy of physical diagnostic tests for assessing meniscal lesions of the knee: a meta-analysis. *J Fam Pract*.

2001;50:938-944.
28. Shrier I, Boudier-Revéret M, Fahmy K. Understanding the different physical examination tests for suspected meniscal tears. *Curr Sports Med Rep*. 2010;9:284-289. http://dx.doi.org/10.1249/JSR.0b013e3181f2727e
29. Snoeker BA, Bakker EW, Kegel CA, Lucas C. Risk factors for meniscal tears: a systematic review including meta-analysis. *J Orthop Sports Phys Ther*. 2013;43:352-367. http://dx.doi.org/10.2519/jospt.2013.4295
30. Solomon DH, Simel DL, Bates DW, Katz JN, Schaffer JL. The rational clinical examination. Does this patient have a torn meniscus or ligament of the knee? Value of the physical examination. *JAMA*. 2001;286:1610-1620. http://dx.doi.org/10.1001/jama.286.13.1610
31. Wagemakers HP, Heintjes EM, Boks SS, et al. Diagnostic value of history-taking and physical examination for assessing meniscal tears of the knee in general practice. *Clin J Sport Med*. 2008;18:24-30. http://dx.doi.org/10.1097/JSM.0b013e31815887a7
32. Yan R, Wang H, Yang Z, Ji ZH, Guo YM. Predicted probability of meniscus tears: comparing history and physical examination with MRI. *Swiss Med Wkly*. 2011;141:w13314. http://dx.doi.org/10.4414/smw.2011.13314