HEALTH TECHNOLOGY ASSESSMENT

VOLUME 19 ISSUE 62 AUGUST 2015
ISSN 1366-5278

Diagnostic accuracy of the Thessaly test, standardised clinical history and other clinical examination tests (Apley's, McMurray's and joint line tenderness) for meniscal tears in comparison with magnetic resonance imaging diagnosis

Mark Blyth, Iain Anthony, Bernard Francq, Katriona Brooksbank, Paul Downie, Andrew Powell, Bryn Jones, Angus MacLean, Alex McConnachie and John Norrie

DOI 10.3310/hta19620

NHS
National Institute for Health Research

Diagnostic accuracy of the Thessaly test, standardised clinical history and other clinical examination tests (Apley's, McMurray's and joint line tenderness) for meniscal tears in comparison with magnetic resonance imaging diagnosis

Mark Blyth,[1*] Iain Anthony,[1] Bernard Francq,[2] Katriona Brooksbank,[1] Paul Downie,[3] Andrew Powell,[1] Bryn Jones,[1] Angus MacLean,[1] Alex McConnachie[2] and John Norrie[4]

[1]Orthopaedic Research Unit, Glasgow Royal Infirmary, Glasgow, UK
[2]Robertson Centre for Biostatistics, Glasgow University, Glasgow, UK
[3]Muirside Medical Practice, Baillieston Health Centre, Glasgow, UK
[4]Centre for Healthcare Randomised Trials (CHaRT), University of Aberdeen, Aberdeen, UK

*Corresponding author

Declared competing interests of authors: John Norrie is a member of the National Institute for Health Research Health Technology Assessment (HTA) and Efficacy and Mechanism Evaluation Editorial Board and HTA Commissioning Board.

Published August 2015
DOI: 10.3310/hta19620

This report should be referenced as follows:

Blyth M, Anthony I, Francq B, Brooksbank K, Downie P, Powell A, *et al*. Diagnostic accuracy of the Thessaly test, standardised clinical history and other clinical examination tests (Apley's, McMurray's and joint line tenderness) for meniscal tears in comparison with magnetic resonance imaging diagnosis. *Health Technol Assess* 2015;**19**(62).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents*®/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/163/02. The contractual start date was in September 2012. The draft report began editorial review in July 2014 and was accepted for publication in March 2015. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Diagnostic accuracy of the Thessaly test, standardised clinical history and other clinical examination tests (Apley's, McMurray's and joint line tenderness) for meniscal tears in comparison with magnetic resonance imaging diagnosis

Mark Blyth,[1*] Iain Anthony,[1] Bernard Francq,[2] Katriona Brooksbank,[1] Paul Downie,[3] Andrew Powell,[1] Bryn Jones,[1] Angus MacLean,[1] Alex McConnachie[2] and John Norrie[4]

[1]Orthopaedic Research Unit, Glasgow Royal Infirmary, Glasgow, UK
[2]Robertson Centre for Biostatistics, Glasgow University, Glasgow, UK
[3]Muirside Medical Practice, Baillieston Health Centre, Glasgow, UK
[4]Centre for Healthcare Randomised Trials (CHaRT), University of Aberdeen, Aberdeen, UK

*Corresponding author Mark.Blyth@ggc.scot.nhs.uk

Background: Reliable non-invasive diagnosis of meniscal tears is difficult. Magnetic resonance imaging (MRI) is often used but is expensive and incidental findings are problematic. There are a number of physical examination tests for the diagnosis of meniscal tears that are simple, cheap and non-invasive.

Objectives: To determine the diagnostic accuracy of the Thessaly test and to determine if the Thessaly test (alone or in combination with other physical tests) can obviate the need for further investigation by MRI or arthroscopy for patients with a suspected meniscal tear.

Design: Single-centre prospective diagnostic accuracy study.

Setting: Although the study was performed in a secondary care setting, it was designed to replicate the results that would have been achieved in a primary care setting.

Participants: Two cohorts of patients were recruited: patients with knee pathology ($n = 292$) and a control cohort with no knee pathology ($n = 75$).

Main outcome measures: Sensitivity, specificity and diagnostic accuracy of the Thessaly test in determining the presence of meniscal tears.

Methods: Participants were assessed by both a primary care clinician and a musculoskeletal clinician. Both clinicians performed the Thessaly test, McMurray's test, Apley's test, joint line tenderness test and took a standardised clinical history from the patient.

Results: The Thessaly test had a sensitivity of 0.66, a specificity of 0.39 and a diagnostic accuracy of 54% when utilised by primary care clinicians. This compared with a sensitivity of 0.62, a specificity of 0.55 and diagnostic accuracy of 59% when used by musculoskeletal clinicians. The diagnostics accuracy of the other tests when used by primary care clinicians was 54% for McMurray's test, 53% for Apley's test, 54% for the joint line tenderness test and 55% for clinical history. For primary care clinicians, age and past history of osteoarthritis were both significant predictors of MRI diagnosis of meniscal tears. For musculoskeletal clinicians age and a positive diagnosis of meniscal tears on clinical history taking were significant predictors of MRI diagnosis. No physical tests were significant predictors of MRI diagnosis in our multivariate models.

The specificity of MRI diagnosis was tested in subgroup of patients who went on to have a knee arthroscopy and was found to be low [0.53 (95% confidence interval 0.28 to 0.77)], although the sensitivity was 1.0.

Conclusions: The Thessaly test was no better at diagnosing meniscal tears than other established physical tests. The sensitivity, specificity and diagnostic accuracy of all physical tests was too low to be of routine clinical value as an alternative to MRI. Caution needs to be exercised in the indiscriminate use of MRI scanning in the identification of meniscal tears in the diagnosis of the painful knee, due to the low specificity seen in the presence of concomitant knee pathology. Further research is required to determine the true diagnostic accuracy and cost-effectiveness of MRI for the detection of meniscal tears.

Trial registration: Current Controlled Trial ISRCTN43527822.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of tables

List of figures

Glossary

Likelihood ratio for negative test How many times more likely you are to not have a meniscal tear if you have a negative test compared with someone who has a positive test.

Likelihood ratio for positive test How many times more likely you are to have a meniscal tear if you have a positive test compared with someone who has a negative test.

Musculoskeletal clinician Orthopaedic specialist knee surgeon or extended scope physiotherapist with specialist knee interest (operating in a secondary care orthopaedic setting).

Negative predictive value The proportion of patients with a negative test who do not have a meniscal tear.

Odds ratio Diagnostic odds ratio of a test is the ratio of the odds of positivity in subjects with a tear relative to the odds in subjects without a tear. It is calculated according to the formula: odds ratio = (true positive/false negative)/(false positive/true negative).

Positive predictive value The proportion of patients with a positive test who actually have a meniscal tear.

Primary care clinician General practitioner- or community-based physiotherapist.

Sensitivity The proportion of those patients with a meniscal tear who are correctly identified as positive by the test. Low sensitivity indicates a high rate of false-negative tests.

Specificity The proportion of those patients without a meniscal tear who are correctly identified as negative by the test. Low specificity indicates a high rate of false-positive tests.

List of abbreviations

ACL	anterior cruciate ligament
AUC	area under the curve
BMI	body mass index
CI	confidence interval
GP	general practitioner
HTA	Health Technology Assessment
MRI	magnetic resonance imaging
NPV	negative predictive value
PPV	positive predictive value
STARD	STAndards for the Reporting of Diagnostic accuracy studies

Plain English summary

The menisci of the knee play an important role in absorbing forces transmitted through the knee during movements. Damage to the menisci can result in pain, may limit movement and can require surgery. Currently the best, non-invasive way to diagnose meniscal tears is using a magnetic resonance imaging (MRI) scan. However, MRI scanning is expensive and is a limited resource within the NHS.

This study was designed to determine if the a simple physical examination test, the Thessaly Test, was a suitable alternative for general practitioners (GPs) to use to determine if patients have a meniscal tear or not. We have compared the accuracy of the Thessaly Test, three other physical tests for meniscal tears (Apley's Test, McMurray's Test and joint line tenderness Test) and a standardised clinical history with the results achieved using MRI. All patients were examined by both a primary care clinician (GP or community physiotherapist) and an orthopaedic musculoskeletal clinician. In total, 282 patients underwent a MRI scan (239 with knee problems and 43 controls with no knee problems).

Our results show that neither the Thessaly Test nor any other simple physical examination tests are suitable alternatives for GPs to use alone to diagnose meniscal tears. The best non-invasive, non-imaged method of diagnosing meniscal tears was a clinical history taken by an experienced musculoskeletal clinician.

The findings of this study have implications for primary care clinicians who may wish to consider referral of patients with significant knee pain to a specialist clinician for assessment rather than automatic referral for MRI.

Scientific summary

Background

The menisci are two semilunar, fibrocartilaginous discs located between the medial and lateral articular surfaces of the femur and tibia in each knee. The menisci play an important role in the knee providing multiple complex functions, including load bearing, stress distribution and shock absorption. Loads transmitted through the joint to the cartilage are partially borne by the menisci; hence they play an important role in both the protection of the cartilage and the subsequent development of degenerative osteoarthritis of the knee should the menisci become damaged.

Damage to the menisci commonly occurs in two different scenarios: in young active individuals during sporting activity and in older individuals as a result of long-term degeneration of the menisci.

Reliable non-invasive diagnosis of meniscal tears is difficult. Magnetic resonance imaging (MRI) is often referred to as the gold standard for non-invasive diagnosis of meniscal tears. However, incidental meniscal findings on MRI of the knee are common in the general population. These incidental findings increase with age and are often not associated with pain. The only completely accurately way to diagnose meniscal tears is to perform an arthroscopy of the knee in order to image the menisci directly. However, this is an invasive procedure and therefore carries risks for the patient.

As an alternative to imaging or surgery there are a number of physical examination tests described for the diagnosis of meniscal tears. The most commonly used physical tests include the joint line tenderness Test, McMurray's Test and Apley's Test. These tests have been in use for many years, but are widely acknowledged to have limited specificity and sensitivity, particularly in the presence of other knee pathologies such as anterior cruciate ligament rupture. More recently Karachalios *et al.* have described a new physical test to detect meniscal tears – the Thessaly Test [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**:955–62]. The Thessaly Test is reported to have a high sensitivity and specificity.

The aim of this prospective study was to determine the diagnostic accuracy of the Thessaly Test and to determine if the Thessaly Test (alone or in combination with other tests) can obviate the need for further investigation by MRI or arthroscopy for patients with suspected meniscal tear.

Although the primary purpose of the study was to evaluate whether or not the Thessaly Test is of use to primary care clinicians in the diagnosis of meniscal tears, the study was actually conducted in a secondary care setting. The rationale for this was that any one individual general practitioner (GP) is likely to see only a very small number of patients with a meniscal tear in the course of an average year. In order to avoid the need for a large network of primary care clinicians who would contribute only relatively small numbers of patients, the study was designed to be undertaken in secondary care, where large numbers of the target patient group are available. Primary care clinicians were brought in to the secondary care setting to evaluate patients. This pragmatic design enhanced the efficiency of the study and provided an opportunity to compare the diagnostic accuracy of the Thessaly Test when utilised by both primary care clinicians and secondary care specialist musculoskeletal clinicians.

Objectives

- To determine the diagnostic accuracy of the Thessaly Test for meniscal tears in the knee and whether or not this test can obviate the need for further investigation by arthroscopy or MRI.
- To determine how the Thessaly Test compares with clinical history and to other commonly used physical examinations (McMurray's Test, Apley's Test and the joint line tenderness Test) in diagnosing meniscal tears.
- To determine if the presence of arthritis or other knee pathologies influences the accuracy of the Thessaly Test.
- To determine if the use of combinations of physical tests (such as the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test) provides better specificity and sensitivity than a single test alone in the diagnosis of meniscal tears.
- To determine the ability of non-specialist GPs to use the Thessaly Test in comparison with specialist knee clinicians.

Methods

This was a single-centre study undertaken at Glasgow Royal Infirmary between 2013 and 2014, study registration number ISRCTN43527822. A total of 367 patients were recruited (covering all age ranges) – 292 patients with suspected knee pathology, 75 patients with no knee pathology.

Knee pathology group

Inclusion criteria

- Patients referred to the knee clinic at Glasgow Royal Infirmary.

Exclusion criteria

- Age < 18 years.
- Unable to give informed consent.
- Previous knee replacement surgery.

Control group

Inclusion criteria

- Patients attending the hand clinic at Glasgow Royal Infirmary or responding to advertisement posters placed within Glasgow Royal Infirmary.

Exclusion criteria

- Age < 18 years.
- Unable to give informed consent.
- Previous knee replacement surgery.
- A history of knee pain in the last 6 months.
- Osteoarthritis.
- Rheumatoid arthritis.

A subgroup of knee pathology patients were recruited directly from a single GP practice. The purpose of this was to allow comparison with the main study group recruited in secondary care to ensure that the demographics of the main group matched those that would be expected to present to a typical GP practice.

Independent assessments of patients knees were carried out by orthopaedic musculoskeletal specialists ($n = 11$) and primary care clinicians ($n = 12$). Each clinician undertook the following diagnostics tests for meniscal tears: the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test. In addition, each clinician made a diagnosis based on a standardised clinical history. In order to avoid bias the order in which each test was performed was randomised.

Magnetic resonance imaging scans were performed as a 'gold standard' for diagnosing meniscal tear in 282 patients (239 knee pathology and 43 controls). The accuracy of MRI diagnosis was further validated in 77 patients who required arthroscopy of the knee. Knee radiography were performed to determine presence of osteoarthritis in 264 patients (all knee pathology patients).

For each diagnostic test we calculated the sensitivity, specificity, negative and positive predictive value, likelihood ratio, odds ratio and diagnostic accuracy.

Multivariate logistic regression and stepwise logistic regression models were used to test whether or not combinations of physical tests were predictive of MRI diagnosis.

Results

Fifty-six per cent of patients in the knee pathology group had a meniscal tear diagnosed on a MRI scan, compared with just 12% in the control group. All patients in the control group were asymptomatic. The medial meniscus was more commonly affected (71.4%). In the knee pathology group 18.3% of patients had a concomitant anterior cruciate ligament injury and 27% of patients had osteoarthritis grade II or higher on the Kellgren and Lawrence scale.

When the Thessaly Test was utilised by primary care clinicians it had a sensitivity of 0.66, a specificity of 0.39 and a diagnostic accuracy of 54%. Similar diagnostic accuracy was achieved for the other tests: McMurray's Test 54%, Apley's Test 53%, joint line tenderness Test 54% and clinical history 55%.

When used by musculoskeletal clinicians the Thessaly Test had a sensitivity of 0.62, a specificity of 0.55 and diagnostic accuracy of 59%. The diagnostic accuracy of the other tests was McMurray's Test 63%, Apley's Test 58%, joint line tenderness Test 64% and clinical history 69%.

Sensitivity of the Thessaly Test was influenced by age and the presence of other knee pathologies when used by primary care clinicians and by the presence of other knee pathologies when used by musculoskeletal clinicians.

Multivariable logistic regression with all factors included and stepwise logistic regression was performed to identify factors predictive of MRI diagnosis. For primary care clinicians, age and past history of osteoarthritis were both significant predictors of MRI diagnosis of meniscal tear. For musculoskeletal clinicians, age and a positive diagnosis of meniscal tear on clinical history taking were significant predictors of MRI diagnosis. No physical tests were predictive in our models for either group of clinicians.

The specificity of MRI diagnosis was tested in subgroup of patients who went on to have a knee arthroscopy. The specificity was 0.53 [95% confidence interval (CI) 0.28 to 0.77] and 0.6 (95% CI 0.32 to 0.84) after patients who had previously had knee surgery were removed from the analysis. The sensitivity of MRI compared with arthroscopy was 1.0.

Conclusions

The data generated by this study suggests that the Thessaly Test is no better at diagnosing meniscal tears than other established physical tests. We found that neither the Thessaly Test alone or in combination with other physical tests could be reliably used by primary care clinicians as an alternative, or surrogate, for MRI scanning to diagnose meniscal tears in the knee.

We have also found that MRI diagnosis of meniscal tears in comparison with arthroscopic diagnosis was less accurate than is commonly reported in the literature. This is particularly true when MRI is used indiscriminately rather than in targeted populations of patients with specific 'suspected' knee pathologies. Further well-defined studies are required to determine the true accuracy of MRI for meniscal tear.

The findings of this study have implications for primary care clinicians who may wish to consider referral of patients with significant knee pain and injuries to a specialist clinician for assessment and treatment rather than automatic referral for MRI imaging of the knee. This approach is likely to maximise the benefit from MRI imaging and minimise the cost of unnecessary or inappropriate imaging.

Trial registration

This trial is registered as ISRCTN43527822.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Background

The menisci are two semilunar, fibrocartilaginous discs located between the medial and lateral articular surfaces of the femur and tibia in each knee.[1] Although both menisci are crescent shaped, the medial meniscus appears comma shaped, with larger separation between the anterior and posterior horns, and the lateral meniscus is 'C' shaped with less separation between the horns (see *Figures 1* and *2*).[1] Lateral menisci are reported to show greater variation in size, shape and mobility than medial mensci.[2] The anatomy of the menisci is unique, being made up of circumferentially and radially oriented collagen fibres that provide resistance to 'hoop' stress and shear forces respectively.[3] The meniscal microstructure of fibrochondrocytes situated in the inner portion and fibroblasts situated on the outer portion enable the key properties of these structures. The fibrochondrocytes synthesise extracellular matrix and the fibroblasts produce collagen and proteoglycans.[3]

The meniscus is a relatively avascular structure; at maturity 10–25% of the peripheral meniscal tissue is supplied by the popliteal artery.[4] Three regions can be described, dividing the meniscus by its vascular supply. The 'red zone' is located in the peripheral third of the meniscal body and is heavily vascularised, the 'red–white zone' comprises the central third of the meniscal body and the 'white zone' is located in the inner third of the body of the meniscus and has no vascular supply (*Figure 1*).[4,5] The healing capacity of each area of the meniscus is directly related to its vascularisation, with the white zone susceptible to permanent damage and degenerative lesions.[4] The peripheral red zone is a thick, convex structure attached to the joint capsule whereas the inner white zone tapers to a thin edge.[6] To allow effective articulation with the convex femoral condyles the superior surface of the menisci are concave. The inferior surfaces of the menisci are flat to match the tibial plateau.[6]

The 'red zone' is heavily vascularised, the 'red–white zone' constitutes the central third of the meniscal body and has limited vascularity, and the 'white zone' is located in the inner third of the body of the meniscus and has no vascular supply.

The menisci play an important role in the knee, providing multiple complex functions including load bearing, stress distribution and shock absorption. Loads transmitted through the joint to the cartilage are partially borne by the menisci, giving them an important role in both the protection of the cartilage and the subsequent development of degenerative osteoarthritis of the knee should the menisci become damaged.

Damage to the menisci commonly occurs in two different scenarios. In young active individuals the menisci are often damaged during sporting activity. Acute tears of the menisci occur secondary to a rotational and

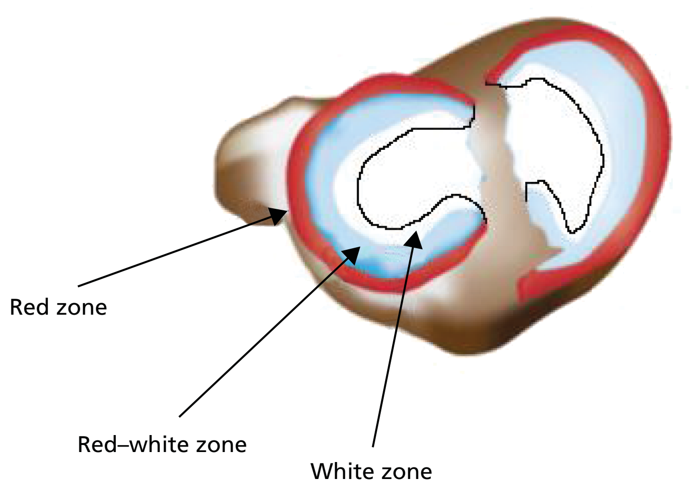


FIGURE 1 Blood supply to the menisci.

compressive load placed on the knee during movement from a flexed position (knee bent) to an extended position (knee straight). Sport participants often describe knee pain after twisting their leg while the foot is planted on the ground with the affected limb bearing their full weight. The combination of compression and rotation that occurs during this movement results in shear stresses that can tear the meniscal tissue. At the time of injury there is often pain and swelling of the knee joint, with twisting or pivoting movements often exacerbating the pain. The pain and swelling subside over a period of weeks after an acute injury, but patients can present weeks later with typical symptoms including intermittent localised pain, locking or catching, giving way, difficulty with deep knee bending and kinesiophobia or fear of movement.

During a 10-year epidemiological study in Switzerland and Germany 17,397 patients with 19,530 sports injuries were documented: 40% of injuries related to the knee joint; 10.8% had a medial meniscus lesion with a further 3.7% having a lateral meniscus lesion.[7] Degenerative meniscal tears are more common in older people. Tears in this age group usually result from long-term degeneration of the meniscus itself and cause joint swelling, joint line pain and mechanical locking.[8–10] The presence of meniscal tears has been identified as an important risk factor in the development of knee osteoarthritis.[10,11] Degenerative meniscal tears are associated with articular cartilage defects, loss of articular cartilage volume and prevalence of radiographical osteoarthritis.[12] Individuals with osteoarthritis of the knee are twice as likely to have a meniscal tear as age-matched individuals without osteoarthritis.[13]

A number of different patterns of meniscal tear have been reported including vertical or longitudinal (bucket-handle), flap or oblique, radial or transverse, torn horn and complex degenerative tears (*Figure 2*).

Reliable non-invasive diagnosis of meniscal tears is difficult. Magnetic resonance imaging (MRI) is often referred to as the gold standard for non-invasive diagnosis of meniscal tears. However, incidental meniscal findings by MRI of the knee are common in the general population. These incidental findings increase with

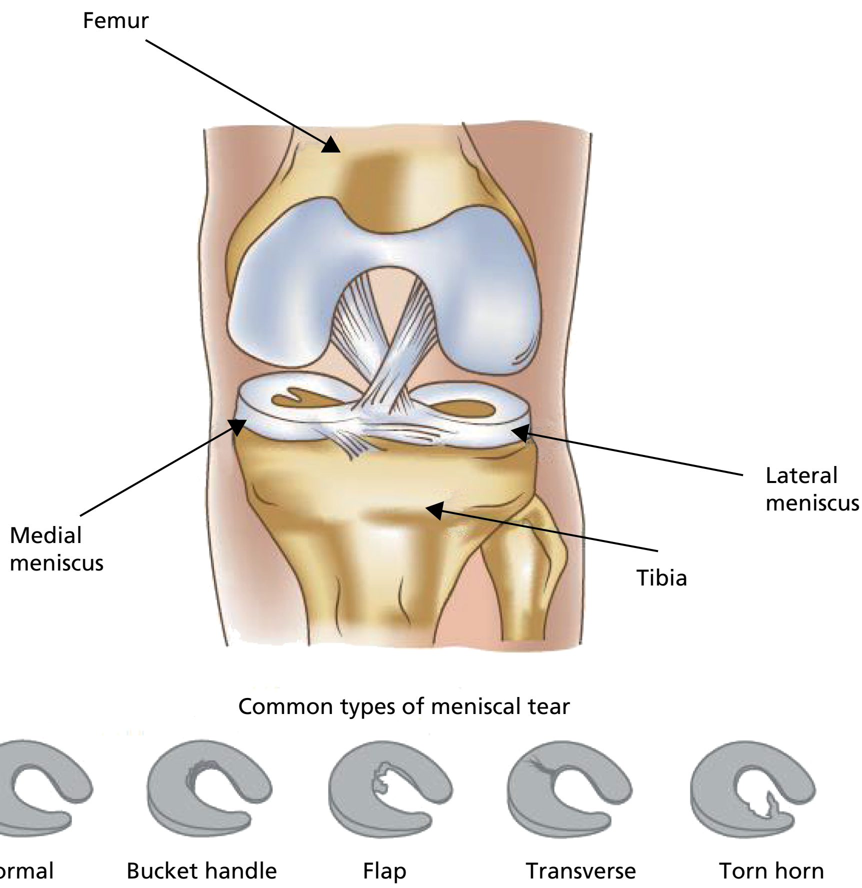


FIGURE 2 Menisci of the knee and common tear patterns.

age and are often not associated with pain.[13] The only completely accurate way to diagnose meniscal tears is to perform an arthroscopy of the knee in order to image the menisci directly. However, this is an invasive procedure and therefore carries risks for the patient. Infection is a rare complication of arthroscopy, but septic arthritis may have devastating consequences and therefore every procedure should be well justified.[14] Reported rates in the literature are low and vary from 0.08% to 0.42%.[15,16] Several systematic reviews taking arthroscopy findings as 'gold standard' compared with MRI results have shown that MRI is both sensitive and specific.[17,18] MRI is reported to be better at identifying patients with medial meniscal tears rather than lateral tears.[17] Sensitivity of 93.3% [95% confidence interval (CI) 91.7% to 95.0%] and specificity of 88.4% (95% CI 85.4% to 91.4%) have been reported for medial meniscal tears by MRI.[18] For lateral meniscal tears sensitivity and specificity are 79.3% (95% CI 74.3% to 84.2%) and 95.7% (95% CI 94.6% to 96.8%) respectively.[18] We believe that performing unnecessary arthroscopy procedures cannot be ethically justified in this patient group and that MRI is an acceptable surrogate for diagnosis of meniscal tears.

However, meniscal damage is also a frequent finding on MRI of the osteoarthritic knee, limiting the value of this diagnostic tool for meniscal tears in this section of the population.[19] Among patients with clinical and radiographic findings of osteoarthritis of the knee the prevalence of meniscal tears has been reported to be as high as 68–90%.[20,21] It is therefore essential that a detailed history, including mechanism of injury, onset of symptoms and a physical examination, is carried out prior to MRI scanning for meniscal tear.[22]

In addition to imaging tests, there are a number of physical examination tests described for the diagnosis of meniscal tears. However, a meniscal tear can be difficult to diagnose as symptoms are often non-specific and associated injuries can disguise a tear in the meniscus.[23] The most commonly used physical tests include the joint line tenderness Test, McMurray's Test and Apley's Test. These tests have been in use for many years, but are widely acknowledged to have limited specificity and sensitivity, particularly in the presence of other knee pathologies, such as anterior cruciate ligament rupture.[24]

Although limited in diagnostic accuracy compared with MRI scanning or arthroscopy, these physical tests have the advantage of being relatively easy to perform and also to incur no additional cost. More recently Karachalios *et al.*[25] have described a new physical test to detect meniscal tears – the Thessaly Test. When performed at 20° of knee flexion the Thessaly Test is reported to have a high diagnostic accuracy for detecting both lateral and medial meniscal tears. *Table 1* summarises the current literature with respect to the diagnostic capabilities of the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test.

Summary description of physical tests for meniscal tears

The joint line tenderness Test involves palpation of the joint line with the knee in 90° of flexion. The test is considered positive when there is pain along the joint line on palpation (*Figure 3*).

The McMurray's Test (*Figure 4*) was originally described in 1940 and designed to detect tears in the posterior portion of the meniscus. A test is considered positive when a click can be heard and/or felt on joint line palpation when the knee is bent beyond 90° flexion and the tibia is rotated on the femur into full internal rotation then full external rotation (to test the lateral and medial meniscus respectively).

Apley's Test is carried out with the patient prone and the knee flexed to 90° (*Figure 5*). The tibia is then compressed onto the knee joint while being externally rotated. If this manoeuvre produces pain, this constitutes a positive test.[30,31]

TABLE 1 Literature summary of the sensitivity and specificity of the Thessaly Test, the joint line tenderness Test, McMurray's Test and Apley's Test

Test	Sensitivity (95% CI)	Specificity (95% CI)	Reference (*n*)	Diagnostic 'gold standard'
McMurray's Test	70.5% (67.4% to 73.4%)	71.1% (69.3% to 72.9%)	Hegedus *et al.*[26] (meta-analysis pooled data)	Mixed arthroscopy and MRI
Apley's Test	60.7% (55.7% to 65.5%)	70.2% (68.0% to 72.4%)	Hegedus *et al.*[26] (meta-analysis pooled data)	Mixed arthroscopy and MRI
Joint line tenderness Test	63.3% (60.9% to 65.7%)	77.4% (75.6% to 79.1%)	Hegedus *et al.*[26] (meta-analysis pooled data)	Mixed arthroscopy and MRI
Thessaly Test (original paper)	89%[a] [medial]	97%[a] [medial]	Karachalios *et al.*[25] ($n = 213$ symptomatic and 197 asymptomatic)	Arthroscopy
Thessaly Test (original paper)	92%[a] [lateral]	96%[a] [lateral]	Karachalios *et al.*[25] ($n = 213$ symptomatic and 197 asymptomatic)	Arthroscopy
Thessaly Test	90%[a]	98%[a]	Harrison *et al.*[27] ($n = 116$)	Arthroscopy
Thessaly Test	59% (47% to 71%) [medial]	67% (45% to 83%) [medial]	Konan *et al.*[28] ($n = 109$)	Arthroscopy
Thessaly Test	31% (15% to 54%) [lateral]	95% (87% to 98%) [lateral]	Konan *et al.*[28] ($n = 109$)	Arthroscopy
Thessaly Test in ACL-deficient patients	79%[a]	40%[a]	Mirzatolooei *et al.*[29] ($n = 80$)	Arthroscopy

ACL, anterior cruciate ligament.
a No 95% CI reported.
Note
Where no side (medial or lateral) is given, the patient cohort was mixed.

FIGURE 3 Joint line tenderness Test.

FIGURE 4 McMurray's Test.

FIGURE 5 Apley's Test.

The Thessaly Test was first described in 2005 by Karachalios *et al.*[25] as a novel clinical examination used to detect meniscal tears in the knee. When performed at 20° of knee flexion it is reported to have a high diagnostic accuracy rate for detecting both lateral and medial meniscal tears.[28] The Thessaly Test is a dynamic reproduction of joint loading in the knee. The examiner supports the patient by holding the patient's outstretched hands while he/she stands flatfooted on the floor. The patient then rotates his or her knee and body, internally and externally, three times, keeping the knee in slight flexion (20°) (*Figure 6*). Patients with a suspected meniscal tear will experience medial or lateral joint line pain if the test is positive.

The aim of this prospective study was to determine the diagnostic accuracy of the Thessaly Test and to determine if the Thessaly Test (alone or in combination with other tests) can obviate the need for further investigation by MRI or arthroscopy for patients with suspected meniscal tears.

FIGURE 6 Thessaly Test.

Chapter 2 Study design/methods

This study was a single-centre diagnostic accuracy study comparing two cohorts of patients with and without knee pathology, assessed by two groups of clinicians – primary care clinicians and musculoskeletal clinicians. The primary goal was to assess the diagnostic accuracy of the Thessaly Test for meniscal tears; however, the diagnostic accuracy of three other physical tests and a standardised clinical history were also analysed.

Musculoskeletal clinicians included both orthopaedic specialist knee surgeons and extended scope physiotherapists with specialist knee interest (operating in a secondary care orthopaedic setting). Primary care clinicians were either general practitioner- (GP-) or community-based physiotherapists.

The study protocol was developed in response to a call for proposals from the National Institute for Health Research Heath Technology Assessment (HTA) programme. Funding was provided by HTA.

Ethical approval was obtained from the West of Scotland Research Ethics Committee (reference 12/WS/0225) prior to initiation of the study. The study was registered on the International Randomised Controlled Trial Number registry (ISRCTN 43527822).

The study was performed as a collaboration between the Orthopaedic Research Unit at Glasgow Royal Infirmary and the Robertson Centre for Biostatistics at the University of Glasgow.

Study objectives

- To determine the diagnostic accuracy of the Thessaly Test for meniscal tears in the knee and whether or not this test can obviate the need for further investigation by arthroscopy or MRI.
- To determine how the Thessaly Test compares with clinical history and to other commonly used physical examinations (McMurray's Test, Apley's Test and the joint line tenderness Test) in diagnosing meniscal tears.
- To determine if the presence of arthritis or other knee pathologies influences the accuracy of the Thessaly Test.
- To determine if the use of combinations of physical tests (such as the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test) provides better specificity and sensitivity than a single test alone in the diagnosis of meniscal tears.
- To determine the ability of non-specialist GPs to use the Thessaly Test in comparison with specialist knee clinicians.

Patient selection

Two cohorts of patients were recruited to the study between October 2012 and March 2014 (*Figure 7*). The first group had unspecified knee pathology and were typical of the target patient group that a primary care clinician may wish to assess using the Thessaly Test. These patients were primarily drawn from patients referred to a specialist knee clinic at Glasgow Royal Infirmary by GPs and other primary care clinicians. A proportion of this group were recruited directly from a single GP practice ($n = 11$). This group were used as an internal control to verify that subjects recruited at Glasgow Royal Infirmary were representative of the target population of patients who attend general practice with knee pain. The second control group consisted of patients with no current or recent knee pathology. These subjects were included in order to test the specificity of the Thessaly Test. Control subjects were recruited from two sources – from the hand injury clinic at Glasgow Royal Infirmary and from advertisement posters placed within Glasgow Royal Infirmary and local universities.

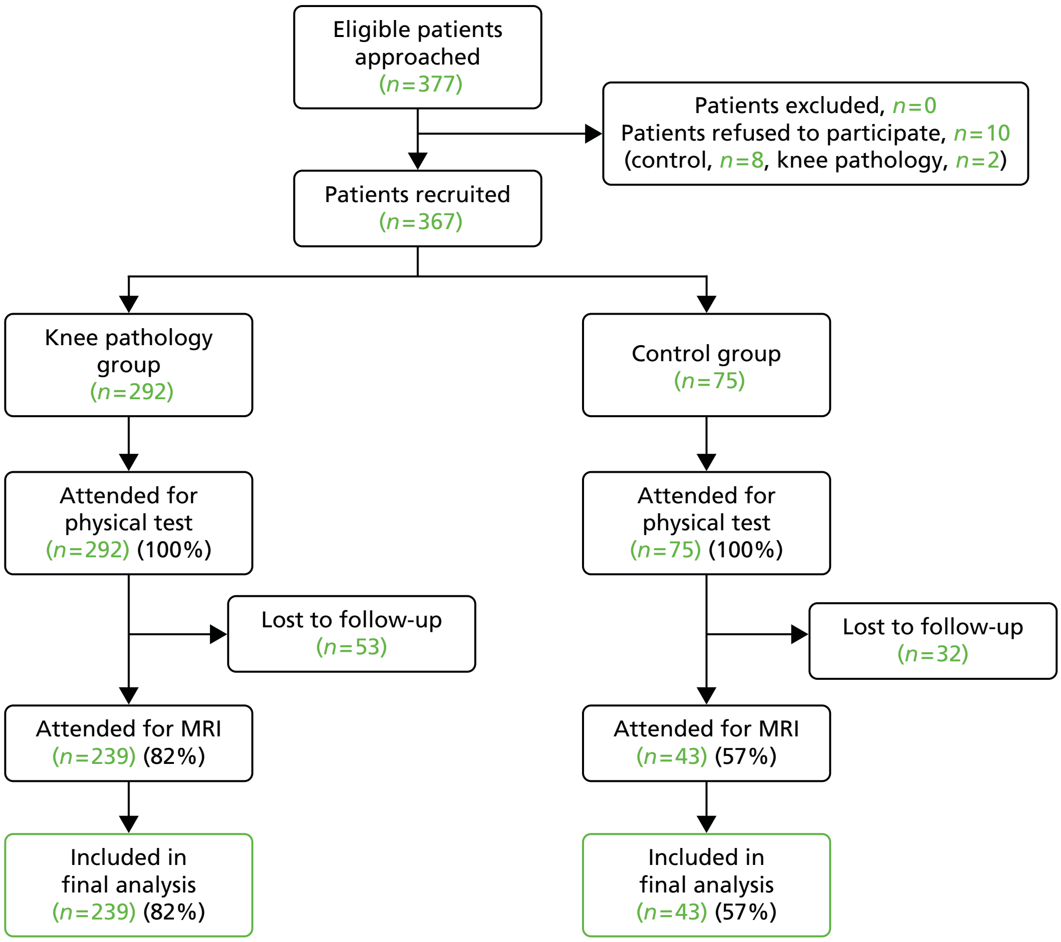


FIGURE 7 Study flow diagram. Note: patients who did not attend for a MRI scan were classed as lost to follow-up.

Only one knee from any one individual was included in the study.

Written informed consent was obtained from each patient.

As all patients attending the knee clinic were essentially eligible for the study, there were no screening failures from the knee clinic. Similarly, patients in the control group responding to our advert were effectively self-screened and therefore there were no screen failures from this source either.

Knee pathology group

Inclusion criteria

- Patients referred to the knee clinic at Glasgow Royal Infirmary.

Exclusion criteria

- Age < 18 years.
- Unable to give informed consent.
- Previous knee replacement surgery.

Control group

Inclusion criteria

- Patients attending the hand clinic at Glasgow Royal Infirmary or responding to advertisement posters placed within Glasgow Royal Infirmary.

Exclusion criteria

- Age < 18 years.
- Unable to give informed consent.
- Previous knee replacement surgery.
- A history of knee pain in the last 6 months.
- Osteoarthritis.
- Rheumatoid arthritis.

Sample size

Sample size calculation was based on the following assumptions. Assuming the sensitivity of the Thessaly Test is around 75%, the study would need around 300 subjects to estimate the sensitivity to within ± 5%. A similar calculation for the width of the CI for a binomial proportion is appropriate for the specificity (e.g. if the specificity was around 90%, the required sample size to estimate the specificity to within ± 8% would be $n = 50$ participants). The power for the pairwise comparison of tests, or combinations of tests, will depend on the degree of disagreement between the tests [e.g. with around 220 pairs of measurements the study would have 90% power to detect a difference in proportions of 0.10 when the proportion of discordant pairs is expected to be 0.15 (using McNemar's Test)].

Patient assessment

Patients were examined by two types of independent practitioners: a trained orthopaedic musculoskeletal specialist ($n = 11$) and a primary care clinician ($n = 12$). The orthopaedic musculoskeletal specialist was either an orthopaedic consultant knee surgeon ($n = 3$) or an orthopaedic extended scope physiotherapist with a specialist interest in the knee ($n = 8$). Primary care clinicians were either a GP ($n = 4$) or a community physiotherapist ($n = 8$). There was a broad range of experience within the primary care clinician group ranging from newly qualified practitioners to a practitioner with more than 30 years of experience. We believe that the diverse mix of experience reflects reality in the primary care setting.

Each clinician performed the following physical tests: the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test. In addition, each clinician took a clinical history from the patient. The order in which physical tests were carried out was randomised to avoid bias in knowledge gained from performing any previous test. This was achieved using a random sequence generator to determine the order that each test should be performed in. The sequences were generated in advance and printed on study data collection forms. The forms were handed out in the order they were printed by the study co-ordinator (KB) to sequential examiners.

All patients in both groups were asked to attend for a MRI scan of their knee, any patient not attending for MRI was excluded from the final analysis. MRI diagnosis of meniscal tears was used as the gold standard reference for each of the physical tests examined in this study. All MRI scans were reported by radiologists who were not linked to the study and who were blind to the patients' status with regard to clinical history and examination findings. MRI scans for the knee pathology patients were carried out within 6 weeks of assessment.

Magnetic resonance imaging scan sequences are detailed below:

- T1 spin-echo Sagittal
- T2 fl2d Sagittal (flash-gradient ECHO)
- proton-density turbo-spin echo fat saturation transverse.
- proton-density turbo-spin echo fat saturation coronal.

In order to determine if the presence of osteoarthritis of the knee influenced the outcome of physical tests for meniscal tears, knee radiography was performed on all patients with knee pathology. No radiography was performed on control patients as it was deemed unethical to expose control patients to unnecessary X-rays. Degree of osteoarthritis of the knee was graded from knee radiography using the Kellgren and Lawrence system.[32] Kellgren and Lawrence grading was undertaken by a single-blinded examiner (AP).

A proportion of patients with a knee pathology subsequently had a routine arthroscopy as part of their treatment/diagnosis of their knee condition. In these patients we recorded the presence of any meniscal tears at the time of arthroscopy. Arthroscopy was normally carried out within 6 weeks of the MRI scan. These data were then correlated with the diagnosis of meniscal tears obtained from MRI images of the patient, in order to verify the accuracy of MRI imaging for meniscal tears by computing specificity, sensitivity, negative and positive predictive values (NPVs and PPVs) for MRI, with arthroscopy as the gold standard.

Physical tests

The Thessaly Test was carried out as described by Karachalios *et al.*[25] The Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test were all carried out as described in *Chapter 1*.

Standardised clinical history

A simple standardised clinical history was developed for this study (see *Standardised minimal clinical history data set*). Our concept was that any test that proved successful in this study required to be simple enough for national roll-out without requiring an extensive training programme. All clinicians were therefore given basic instruction in the use of the standardised clinical history in *Standardised minimal clinical history data set* and asked to use this as a minimum in their diagnosis.

Standardised minimal clinical history data set

- History of knee injury?
    - Was this a sports injury?
- Past history of anterior cruciate ligament (ACL) injury?
- Past history of arthritis?
- Previous surgery to the knee?
- Past history of medial meniscal tear?
- Past history of lateral meniscal tear?
- Presence of associated patella–femoral symptoms?

Primary outcome measures

- Sensitivity and specificity of the Thessaly Test in determining the presence of meniscal tears when employed by primary care clinicians.

Secondary outcomes measures

- Sensitivity and specificity of the Thessaly Test in determining the presence of meniscal tears when employed by specialist musculoskeletal clinicians.
- Sensitivity and specificity of McMurray's Test, Apley's Test, the joint line tenderness Test and clinical history in determining the presence of meniscal tears.
- Determination of the influence of osteoarthritis on the sensitivity and specificity of the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test.
- Determination of the influence of other knee pathologies such as ACL damage or patellofemoral instability on the sensitivity and specificity of the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test.
- Determination of the optimal combination of tests for most accurate diagnosis of meniscal tears.

Statistical analysis

The full statistical analysis plan is described in detail in *Appendix 1*. In brief, we calculated the specificity, sensitivity, PPV and NPV for each of the following individual tests: Thessaly Test, McMurray's Test, Apley's Test, the joint line tenderness Test and clinical history. We have used evidence of meniscal tear on MRI as our 'gold standard' reference for diagnosis. We have compared the relative rates between each individual test. These results are reported separately for specialist musculoskeletal clinicians and for primary care clinicians.

The following subanalysis was also performed:

- Patients with and without osteoarthritis of the knee were compared to determine if osteoarthritis influences the outcome of the Thessaly Test or any of the other physical tests. Other patient factors were also considered [body mass index (BMI), sex, age and previous surgery].
- A comparison of results of all four physical tests between specialist musculoskeletal clinicians and primary care clinicians was performed to determine if specialised training influences the outcome of test results.
- We have examined whether or not a combination of physical tests can produce greater diagnostic accuracy than a single test alone.
- In a subset of patients, we have compared the accuracy of MRI scan diagnosis of meniscal tears to that achieved using arthroscopy.

Statistical tests applied to data

A chi-squared test was used for comparisons between musculoskeletal clinicians and primary care clinicians for each physical test and clinical history, as well as to compare the sensitivity or specificity of the four physical tests along the different bins of the categorical (or categorised) covariates.

McNemar's Test was used to for comparisons between musculoskeletal clinicians and primary care clinicians for each physical test and clinical history. McNemar's Test assesses the significance of the difference between two correlated proportions, where the two proportions are based on the same sample of subjects. In this study we wished to determine the difference between the proportion of successful diagnoses obtained by the two groups of observers; hence, the choice of McNemar's Test over intraclass correlation.

Logistic regression was used to assess the influence of different covariates on the sensitivity or specificity of the four physical tests for musculoskeletal clinicians and primary care clinicians (all patients or knee patients).

Logistic regression using a stepwise procedure from the covariates and the four physical tests to build a model to predict MRI.

Guide to likelihood ratio graphs

The diagnostic abilities of two or more diagnostic tests are traditionally compared by their respective sensitivities and specificities. Comparison is relatively simple if both tests have similar sensitivity, but one has a superior specificity. However, comparison is more complicated when the specificities and sensitivities are both different and a trade-off of one over the other is required.

Likelihood ratio graphs are a simple graphic that readily facilitates comparison between two or more diagnostic tests, allowing an assessment of whether or not a decrease in specificity may be offset by a sufficient gain in sensitivity to yield a test that, nevertheless, has superior diagnostic ability.

For all likelihood ratio graphs in this report, the Thessaly Test is plotted as the reference test (solid black lines). Any comparator test plotted in zone I is superior in all aspects to the Thessaly Test. Comparators plotted in zone II are superior at detecting the absence of a meniscal tear, but inferior at detecting the presence of a tear. Comparators plotted in zone III are superior at detecting the presence of a tear, but inferior at detecting the absence of a tear. Any comparator plotted in zone IV is inferior to the Thessaly Test (*Figure 8*).

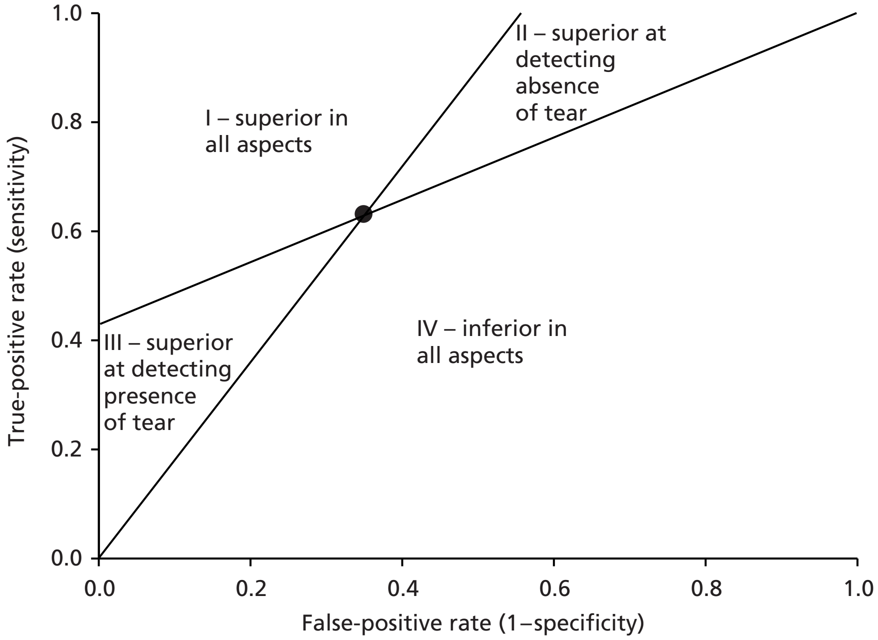


FIGURE 8 Likelihood ratio graph.

Chapter 3 Study cohort demographics and description

Demographics

The study cohort demographics are displayed in *Tables 2* and *3*.

The control group were on average 4.7 years younger than the knee pathology group with the same age range for both groups (see *Table 2*). Forty-nine per cent of the control group were male compared with a significantly higher proportion of 64% in the knee pathology group (see *Table 3*). Although the mean ages and sex distribution were statistically significantly different we do not believe that this is important in the context of this particular study. Furthermore, our logistic regression models have adjusted for variables such age and sex.

The control group were included to allow us to test the specificity of the Thessaly Test in a group with no knee pathologies and not to provide a direct comparison with the knee pathology group. The specificity of a test indicates how often the test correctly predicts negative results. As increasing age is associated with meniscal tears a younger control cohort are less likely to have meniscal tears and therefore should allow more accurate assessment of the specificity of the test.

TABLE 2 Age and BMI descriptive statistics

		Age (years)[a]	BMI (kg/m^2)[a]
All patients	*n* (NA)	367 (0 or 0%)	345 (22 or 6%)
	Mean (SD)	41.6 (13.7)	27.5 (4.7)
	Min.–max.	18–71	18–47
Knee pathology group	*n* (NA)	292 (0 or 0%)	276 (16 or 5.5%)
	Mean (SD)	42.6 (13.4)	27.8 (4.8)
	Min.–max.	18–71	18–47
Control group	*n* (NA)	75 (0 or 0%)	69 (6 or 8%)
	Mean (SD)	37.9 (14.3)	26.2 (4.2)
	Min.–max.	18–71	20.4–41.7
p-value	*t*-test	0.011	0.008

Max., maximum; min., minimum; NA, not accounted; SD, standard deviation.
a The first value is the number of missing values and the second number is the percentage.

TABLE 3 Sex distribution

	Knee pathology group	Control group
Female, *n*	104	38
Male, *n*	188	37
% male	64	49

Fisher's exact test p-value = 0.023.

Approximately half of the study cohort had a meniscal tear identified by MRI (*Table 4*), these tears were predominantly medial (*Table 5*). In the knee pathology group 56% of patients ($n = 130$) had a meniscal tear compared with just 12% in the control group ($n = 5$). All patients in the control cohort who had meniscal tears diagnosed by MRI were asymptomatic ($n = 5$) and four of these five patients had osteoarthritis of the knee. In the knee pathology group 71.4% of those with meniscal tears had medial tears, 28.6% lateral tears and 11% had meniscal cysts (note some patients had more than one pathology). A total of 18.3% of the knee pathology cohort had an ACL injury and 27% had grade II or worse osteoarthritis of the knee on X-ray (*Table 6*). Other major knee pathologies observed in both groups are detailed in *Table 7*.

TABLE 4 Magnetic resonance imaging reported observations on study cohort

MRI observation		Knee pathology		Control	
		Yes	No	Yes	No
Meniscal tear	*n*	130	100	5	38
	%	56.5	43.5	11.6	88.4
Arthritis	*n*	81	148	9	34
	%	35.4	64.6	20.9	79.1
ACL injury	*n*	42	187	1	42
	%	18.3	81.7	2.3	97.7

TABLE 5 Meniscal tear location for knee pathology group ($n = 130$ with meniscal tear on MRI)

	Medial	Lateral
%	71.4	28.6

TABLE 6 Presence of osteoarthritis of the knee as defined by Kellgren and Lawrence[32] grade for knee pathology group

Grade	0	I	II	III	IV
n	109	83	35	29	8
%	41.3	31.4	13.3	11	3

TABLE 7 Other knee pathologies assessed for the study

Other knee pathologies	Knee pathology	Control group
Previous knee surgery, %	15	0
History of knee pain (last 6 months), %	100	0
Rheumatoid arthritis, %	0	0
Ligament instabilities, %	21	2
Tendonitis, %	3	0
Patellofemoral disorders, %	16	0

Chapter 4 Diagnostic accuracy of the Thessaly Test and other tests for diagnosis of meniscal tear

STAndards for the Reporting of Diagnostic accuracy studies (STARD) diagrams are provided in *Appendix 2* for each of the following tests: the Thessaly Test, the joint line tenderness Test, McMurray's Test, Apley's Test and clinical history. These figures give the exact number of patients assessed using each test and their respective test outcomes.

Accuracy of the Thessaly Test when used by primary care clinicians

When the study tests were performed by primary care clinicians, significant differences were found between the tests with respect to sensitivity and specificity of the tests, but not with NPVs and PPVs (*Table 8*).

The most sensitive test was the joint line tenderness Test [0.77 (95% CI 0.68 to 0.84)]; however, this test also had the lowest specificity in the hands of primary care clinicians [0.26 (95% CI 0.18 to 0.36)]. *Figure 9* shows a likelihood ratio graph for each test with the Thessaly Test set as the reference line (solid black). The graph shows slight superiority for McMurray's Test over the Thessaly Test when used by primary care clinicians. No single test was identified as being an appropriate surrogate for MRI diagnosis of meniscal tears by primary care clinicians (see *Table 9*). All tests including clinical history had a diagnostic accuracy of 55% or lower (*Table 9*).

For guidance on the interpretation of likelihood ratio graphs, see *Chapter 2, Guide to likelihood ratio graphs*.

TABLE 8 Accuracy of physical test and recorded clinical history vs. MRI for diagnosis of meniscal tears by primary care clinicians for subjects with knee pathology (95% CI)

Statistical test	Thessaly Test	McMurray's Test	Apley's Test	Joint line tenderness Test	Recorded clinical history	*p*-value
Sensitivity	0.66 (0.57 to 0.74)	0.58 (0.49 to 0.67)	0.53 (0.44 to 0.62)	0.77 (0.68 to 0.84)	0.65 (0.56 to 0.74)	0.001
Specificity	0.39 (0.29 to 0.50)	0.56 (0.45 to 0.66)	0.53 (0.42 to 0.63)	0.26 (0.18 to 0.36)	0.43 (0.33 to 0.54)	< 0.001
LR+	1.08 (0.88 to 1.33)	1.33 (1.01 to 1.75)	1.12 (0.85 to 1.46)	1.04 (0.89 to 1.21)	1.16 (0.93 to 1.43)	–
LR–	0.87 (0.61 to 1.25)	0.74 (0.56 to 0.98)	0.90 (0.69 to 1.17)	0.89 (0.57 to 1.41)	0.80 (0.57 to 1.11)	–
OR	1.24 (0.71 to 2.18)	1.79 (1.04 to 3.09)	1.24 (0.73 to 2.12)	1.16 (0.63 to 2.13)	1.45 (0.84 to 2.49)	–
PPV	0.47 (0.40 to 0.55)	0.53 (0.45 to 0.62)	0.50 (0.41 to 0.58)	0.45 (0.38 to 0.51)	0.48 (0.40 to 0.55)	0.598
NPV	0.47 (0.35 to 0.58)	0.50 (0.40 to 0.60)	0.45 (0.35 to 0.55)	0.46 (0.33 to 0.60)	0.49 (0.39 to 0.60)	0.96

LR–, likelihood ratio for negative test; LR+, likelihood ratio for positive test; OR, odds ratio.
p-values are based on a chi-squared distribution to assess whether or not the sensitivities, specificities, PPV or NPV are equal along the four physical tests and clinical history.

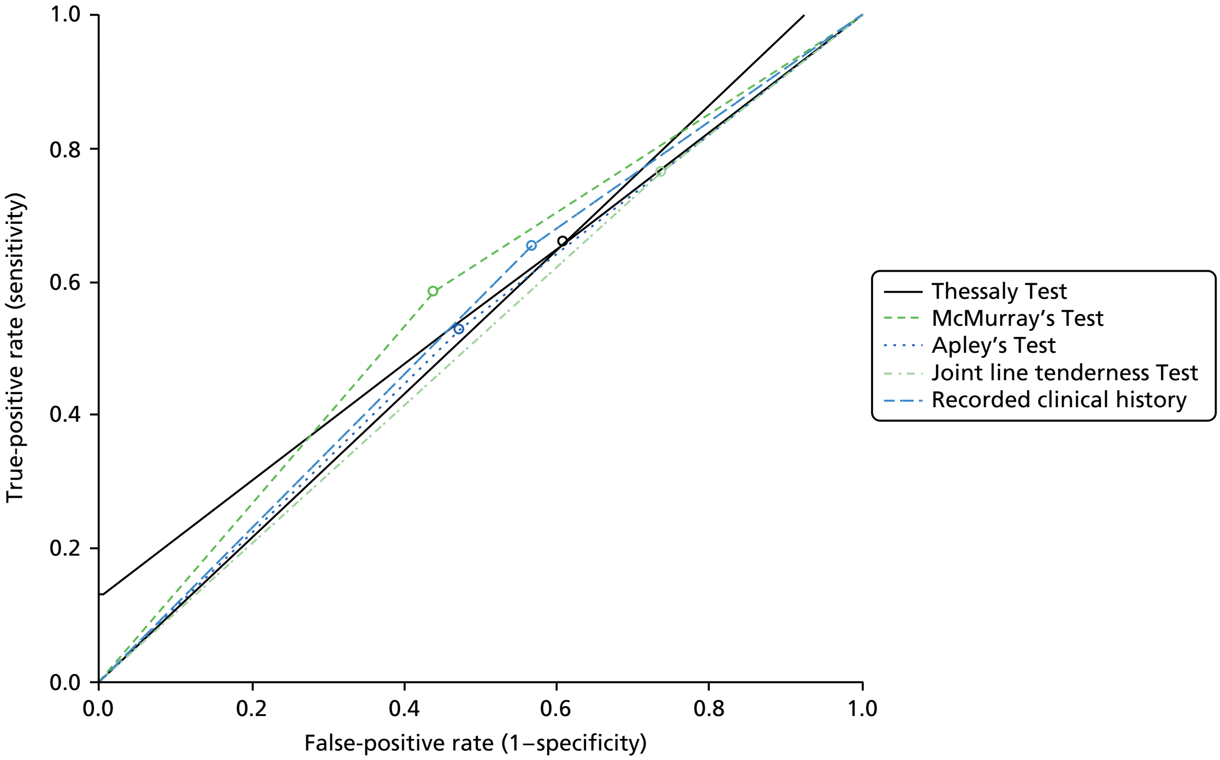


FIGURE 9 Likelihood ratio graph, primary care clinicians vs. MRI (knee pathology patients only).

TABLE 9 Implications of primary care clinicians using the Thessaly Test in routine practice as an assessment screening tool for meniscal tears to differentiate those patients with knee pathology who need a MRI scan/referral to secondary care for meniscal tears

Diagnostic test	In every 100 patients how many will have a MRI-positive meniscal tear?	In every 100 patients how many patients with a tear would be correctly sent for a MRI scan?	In every 100 patients how many patients with a tear would be missed?	In every 100 patients how many patients without a tear would be wrongly referred for a MRI scan?	Diagnostic accuracy[a]
Thessaly Test	56/100	37	19	27	54%
McMurray's Test	56/100	33	23	19	54%
Apley's Test	56/100	30	26	21	53%
Joint line tenderness Test	56/100	43	13	33	54%
Clinical history	56/100	36	20	25	55%

a Diagnostic accuracy = prevalence × sensitivity + [1 – (prevalence × specificity)]. Prevalence is based on frequency of MRI-positive tests identified in this study in patients with knee pathology.

We examined the influence of subject characteristics (age, sex and BMI), arthritis, presence of knee injury, ACL injury on the sensitivity and specificity of the Thessaly Test as performed by primary care clinicians. When examining the entire cohort, only the presence of a knee injury influenced sensitivity of the Thessaly Test (*Table 10*). Specificity was influenced by both the presence of an injury and the report of a sports injury. These results are intuitive and fit with the accepted pathogenesis of meniscal tear, at least in younger patients.

The STARD diagram for the Thessaly Test when carried out by primary care clinicians (see *Appendix 2*) shows that in 35 patients the clinician was unable to perform the Thessaly Test. This was due to pain in the knee at the time of the examination. In order to ensure that this effect has not biased the overall results we have compared the rate of meniscal tears diagnosed by MRI for the group of patients in whom the test was unable to be performed with the remaining study cohort. Eight of the 35 patients did not attend for a MRI scan and we therefore have no knowledge of their meniscal status. The results for the remaining 27 patients are detailed in *Table 11*. A statistical comparison of the two groups of patients, using Fisher's exact test, revealed no significant difference between the two groups ($p = 0.224$).

TABLE 10 Influence of subject characteristics on sensitivity and specificity of the Thessaly Test as performed by primary care clinicians: ORs (with 95% CI)

Subject characteristic		Sensitivity	Specificity
Sex	Female		
	Male	0.807 (0.306 to 2.127); $p = 0.664$	1.18 (0.449 to 3.103); $p = 0.737$
Age (years)		0.986 (0.945 to 1.029); $p = 0.526$	1.001 (0.957 to 1.046); $p = 0.976$
BMI (kg/m^2)	Low (< 27.5 kg/m^2)		
	High (≥ 27.5 kg/m^2)	0.808 (0.336 to 1.940); $p = 0.633$	0.546 (0.214 to 1.393); $p = 0.205$
Arthritis	Yes		
	No	0.638 (0.217 to 1.877); $p = 0.414$	2.312 (0.639 to 8.363); $p = 0.201$
Knee injury	No injury		
	Other injury	0.139 (0.028 to 0.681); $\mathbf{p = 0.015}$	0.088 (0.021 to 0.368); $\mathbf{p < 0.001}$
	Sports injury	1.982 (0.749 to 5.242); $p = 0.168$	0.242 (0.08 to 0.729); $\mathbf{p = 0.012}$
ACL	Yes		
	No	0.844 (0.102 to 6.983); $p = 0.875$	Not estimable
Past arthritis	Yes		
	No	1.098 (0.164 to 7.350); $p = 0.923$	0.703 (0.098 to 5.052); $p = 0.726$
Previous surgery	Yes		
	No	1.128 (0.283 to 4.494); $p = 0.865$	6.836 (0.533 to 87.635); $p = 0.14$

NA, not accounted; OR, odds ratio.
Significant *p*-values are in bold.

TABLE 11 Comparison of rate of meniscal tears diagnosed by MRI between patients where the Thessaly Test was not performed and was performed

MRI diagnosis of meniscal tear	Thessaly not performed	Thessaly performed
Yes, *n*	10	125
No, *n*	17	121

Fisher's exact test $p = 0.224$.

Accuracy of the Thessaly Test when used by musculoskeletal specialists

When the study tests were performed by specialist musculoskeletal clinicians, significant differences were found between the tests with respect to sensitivity and specificity and NPV of the tests, but not PPVs (*Table 12*).

The most sensitive test was the joint line tenderness Test [0.83 (95% CI 0.75 to 0.89)]; however, this test also had the lowest specificity [0.39 (95% CI 0.29 to 0.49)]. The test with the highest sensitivity was Apley's Test [0.72 (95% CI 0.61 to 0.81)], but Apley's Test had low specificity [0.43 (95% CI 0.34 to 0.52)]. *Figure 10* shows a likelihood ratio graph for each test with the Thessaly Test set as the reference line (solid black). The graph shows slight superiority for both clinical history and McMurray's Test over the Thessaly Test when used by musculoskeletal clinicians.

No diagnostic physical test was better at diagnosing a meniscal tear than a well-trained musculoskeletal clinician taking a clinical history (*Figure 10* and *Table 13*). McMurray's Test was a better overall test than the Thessaly Test with a better compromise of sensitivity and specificity (see *Figure 10*).

There were only subtle differences in the ability of each test to diagnose medial and lateral tears (*Table 14*).

Interestingly, the specificity and NPV was greater for all tests in the control group when analysed separately (*Table 15*). This indicates that the tests actually work well in differentiating a normal knee from a knee with a meniscal tear. However, all of the physical tests have limited ability to differentiate between a painful knee with a meniscal tear and painful knee due to any other cause. *Table 16* lists the diagnosis and clinical findings in patients who were positive for the Thessaly Test, but did not have a meniscal tear on MRI scan. This demonstrates the range of common conditions that affect the knee which are poorly differentiated by both the Thessaly Test and the other physical tests assessed in this study.

For guidance on the interpretation of likelihood ratio graphs, see *Chapter 2*, *Guide to likelihood ratio graphs*.

TABLE 12 Accuracy of physical tests and recorded clinical history vs. MRI for diagnosis of meniscal tears by musculoskeletal clinicians for knee pathology subjects only (95% CI)

	Test					
Location	Thessaly Test	McMurray's Test	Apley's Test	Joint line tenderness Test	Recorded clinical history	*p*-value
Sensitivity	0.62 (0.52 to 0.71)	0.63 (0.53 to 0.72)	0.43 (0.34 to 0.52)	0.83 (0.75 to 0.89)	0.76 (0.67 to 0.83)	< 0.001
Specificity	0.55 (0.44 to 0.66)	0.63 (0.53 to 0.73)	0.72 (0.61 to 0.80)	0.39 (0.29 to 0.49)	0.60 (0.50 to 0.70)	< 0.001
LR+	1.38 (1.05 to 1.81)	1.72 (1.26 to 2.33)	1.52 (1.04 to 2.21)	1.36 (1.14 to 1.62)	1.89 (1.46 to 2.46)	–
LR–	0.69 (0.51 to 0.93)	0.50 (0.44 to 0.78)	0.80 (0.65 to 0.97)	0.44 (0.28 to 0.69)	0.40 (0.29 to 0.57)	–
OR	2.00 (1.14 to 3.50)	2.93 (1.66 to 5.19)	1.91 (1.08 to 3.38)	3.08 (1.67 to 5.67)	4.69 (2.66 to 8.29)	–
PPV	0.55 (0.46 to 0.64)	0.57 (0.48 to 0.66)	0.59 (0.48 to 0.69)	0.50 (0.43 to 0.57)	0.62 (0.54 to 0.69)	0.227
NPV	0.52 (0.41 to 0.62)	0.57 (0.47 to 0.67)	0.49 (0.41 to 0.58)	0.64 (0.51 to 0.76)	0.66 (0.55 to 0.76)	0.075

LR–, likelihood ratio for negative test; LR+, likelihood ratio for positive test; OR, odds ratio.
p-values are based on a chi-squared distribution to assess whether or not the sensitivities, specificities, PPV or NPV are equal along the four physical tests and clinical history.

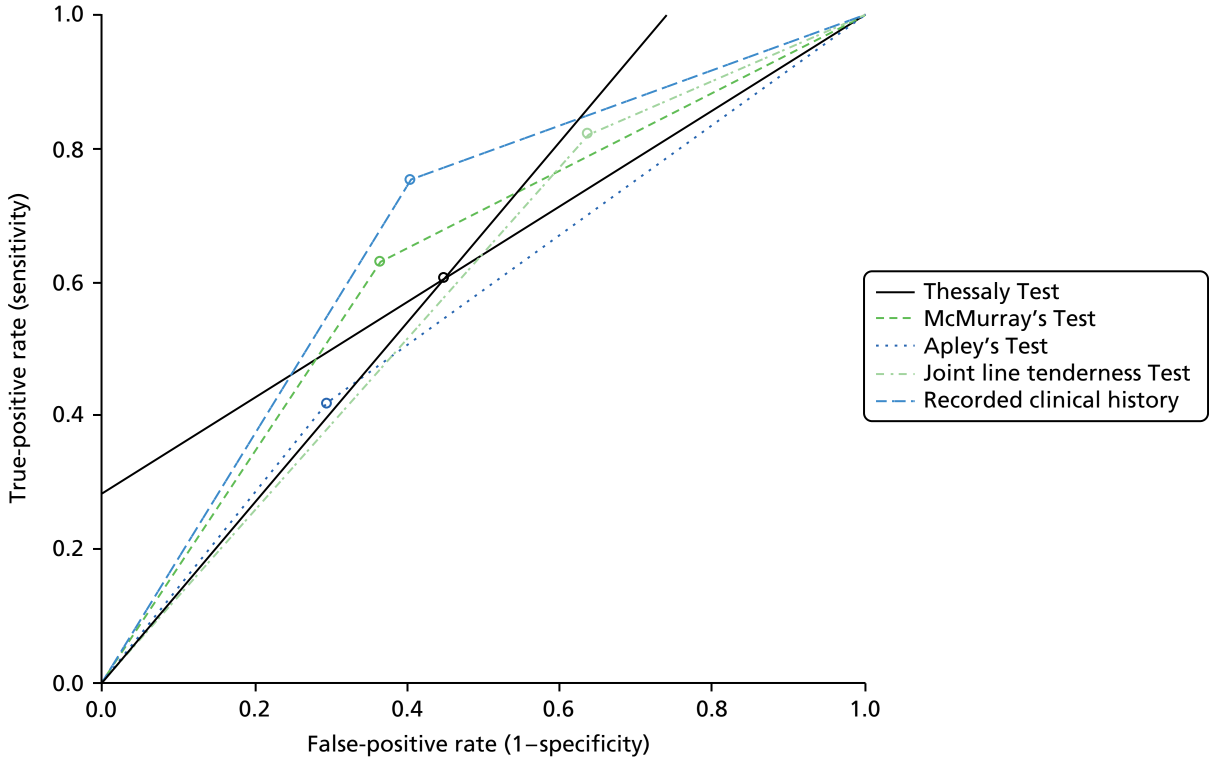


FIGURE 10 Likelihood ratio graph, musculoskeletal clinicians vs. MRI (knee pathology patients only).

TABLE 13 Implications of musculoskeletal clinicians using the Thessaly Test, McMurray's Test and clinical history in routine practice as an assessment screening tool for meniscal tears to differentiate those who need a MRI scan

Diagnostic test	In every 100 patients how many will have a MRI-positive meniscal tear?	In every 100 patients how many patients with a tear would be correctly sent for a MRI scan?	In every 100 patients how many patients with a tear would be missed?	In every 100 patients how many patients without a tear would be wrongly referred for a MRI scan?	Diagnostic accuracy[a]
Thessaly Test	56/100	35	21	20	59%
McMurray's Test	56/100	35	21	16	63%
Apley's Test	56/100	24	32	10	58%
Joint line tenderness Test	56/100	47	9	27	64%
Clinical history	56/100	43	13	18	69%

a Diagnostic accuracy = prevalence × sensitivity + [1 − (prevalence × specificity)]. Prevalence is based on frequency of MRI-positive tests identified in this study in patients with knee pathology.

TABLE 14 Comparison of diagnostic accuracy for medial and lateral meniscal tears (95% CI)

Test	Thessaly Test		McMurray's Test		Apley's Test		Joint line tenderness Test		Recorded clinical history	
Location	Medial	Lateral	Medial	Lateral	Medial	Lateral	Medial	Lateral	Medial	Lateral
Sensitivity	0.63 (0.52 to 0.74)	0.59 (0.41 to 0.75)	0.62 (0.50 to 0.73)	0.66 (0.48 to 0.81)	0.44 (0.33 to 0.55)	0.41 (0.25 to 0.58)	0.82 (0.72 to 0.89)	0.84 (0.68 to 0.94)	0.76 (0.66 to 0.85)	0.73 (0.56 to 0.86)
Specificity	0.51 (0.42 to 0.60)	0.46 (0.38 to 0.54)	0.55 (0.46 to 0.64)	0.51 (0.44 to 0.59)	0.68 (0.60 to 0.76)	0.64 (0.57 to 0.71)	0.33 (0.25 to 0.41)	0.29 (0.23 to 0.36)	0.51 (0.42 to 0.60)	0.43 (0.36 to 0.50)
LR+	1.3 (1.02 to 1.66)	1.09 (0.80 to 1.50)	1.38 (1.07 to 1.80)	1.35 (1.02 to 1.80)	1.38 (0.98 to 1.96)	1.13 (0.73 to 1.76)	1.22 (1.05 to 1.42)	1.18 (1.00 to 1.40)	1.56 (1.27 to 1.92)	1.28 (1.01 to 1.61)
LR–	0.71 (0.51 to 1.00)	0.89 (0.58 to 1.38)	0.69 (0.50 to 0.95)	0.67 (0.41 to 1.08)	0.82 (0.66 to 1.03)	0.93 (0.69 to 1.23)	0.55 (0.33 to 0.91)	0.56 (0.26 to 1.20)	0.46 (0.31 to 0.69)	0.63 (0.36 to 1.10)
OR	1.82 (0.31 to 0.97)	1.22 (0.39 to 1.72)	2.01 (0.28 to 0.88)	2.03 (0.23 to 1.05)	1.69 (0.34 to 1.04)	1.23 (0.4 to 1.68)	2.23 (1.17 to 4.27)	2.12 (0.84 to 5.37)	3.38 (1.87 to 6.12)	2.02 (0.93 to 4.42)
PPV	0.39 (0.31 to 0.48)	0.15 (0.09 to 0.22)	0.38 (0.30 to 0.47)	0.18 (0.12 to 0.26)	0.41 (0.31 to 0.52)	0.17 (0.10 to 0.26)	0.34 (0.28 to 0.41)	0.15 (0.10 to 0.20)	0.43 (0.35 to 0.51)	0.17 (0.12 to 0.24)
NPV	0.67 (0.57 to 0.77)	0.85 (0.76 to 0.91)	0.70 (0.60 to 0.79)	0.88 (0.80 to 0.94)	0.66 (0.57 to 0.74)	0.84 (0.77 to 0.90)	0.74 (0.61 to 0.84)	0.90 (0.80 to 0.96)	0.77 (0.67 to 0.85)	0.89 (0.81 to 0.95)

LR–, likelihood ratio for negative test; LR+, likelihood ratio for positive test; OR, odds ratio.

TABLE 15 Specificity and NPV in control patients assessed by musculoskeletal clinicians (95% CI)

Statistical test	Thessaly Test	McMurray's Test	Apley's Test	Joint line tenderness Test	Recorded clinical history	*p*-value
Specificity	1.00 (0.83 to 1.00)	1.00 (0.83 to 1.00)	1.00 (0.82 to 1.00)	0.95 (0.75 to 1.00)	1.00 (0.83 to 1.00)	0.407
NPV	0.80 (0.59 to 0.93)	0.83 (0.63 to 0.95)	0.83 (0.61 to 0.95)	0.83 (0.61 to 0.95)	0.80 (0.59 to 0.93)	0.997

TABLE 16 Diagnosis and clinical findings in patients who had a positive Thessaly Test (when performed by a musculoskeletal clinician), but were negative for a meniscal tear by MRI

$n = 39$	Yes	No
Primary osteoarthritis, %	43.6	56.4
Patellofemoral disorders, %	17.9	82.1
Ligament instabilities, %	7.7	92.3
Osteochondritis dissecans, %	2.6	97.4
Tendonitis, %	2.6	97.4
Seronegative arthritis, %	2.6	97.4

Comparison between primary care clinicians and specialist musculoskeletal clinicians using physical tests and clinical history to diagnose meniscal tears

Comparison between the two clinician groups (specialist musculoskeletal clinicians and primary care clinicians) showed significant differences in the results obtained using both the Thessaly Test and Apley's Test to diagnose meniscal tears (*Table 17*). Better agreement was found with McMurray's Test and the joint line tenderness Test. Note that this analysis only assesses whether or not the two clinicians agreed and if they were correct. Assessment of diagnostic accuracy shows that specialist musculoskeletal clinicians were consistently better at diagnosing meniscal tears based on both physical tests and on clinical history (*Table 18*).

TABLE 17 Comparison of musculoskeletal clinicians and primary care clinicians use of physical tests and clinical history to diagnose meniscal tears (McNemar's Test)

Clinical test	*p*-value
Thessaly Test	0.001
McMurray's Test	0.91
Apley's Test	< 0.001
Joint line tenderness Test	0.525
Recorded clinical history	0.076

TABLE 18 Comparison of diagnostic accuracy between clinician groups using the Thessaly Test, McMurray's Test and clinical history

Clinical test	Primary care clinician	Musculoskeletal clinician	Difference
Thessaly Test	54%	59%	+5%
McMurray's Test	54%	63%	+9%
Apley's Test	53%	58%	+5%
Joint line tenderness Test	54%	64%	+5%
Recorded clinical history	55%	69%	+14%

Data taken from *Tables 9* and *13*.

Influence of the presence of osteoarthritis and other patient factors on the accuracy of the Thessaly Test (and other physical tests)

We have examined the influence of the following patient characteristics on the sensitivity and specificity of all four physical tests in the hands of both musculoskeletal and primary care clinicians: sex, age, BMI, osteoarthritis, sports injury, knee injury, ACL tear and previous surgery on the affected knee.

The sensitivity of the Thessaly Test when performed by primary care clinicians was influenced by age and presence of sports injuries (*Table 19*). The sensitivity of Apley's Test was influenced by age and other knee injuries. The specificity of McMurray's Test was also influenced by age and other injuries.

The presence of an injury influenced the sensitivity of all tests performed by musculoskeletal clinicians. In addition, the joint line tenderness Test was also influenced by age (*Table 20*). No factors influenced specificity.

The presence of osteoarthritis of the knee did not significantly influence the sensitivity or specificity of any of the tests, performed by either group of clinicians.

TABLE 19 Influence of subject characteristics, arthritis and medical history on sensitivity of physical tests in knee pathology patients (*p*-value of logistic regressions)

	Musculoskeletal clinicians				Primary care clinicians			
Subject characteristic	Thessaly Test	McMurray's Test	Apley's Test	Joint line tenderness Test	Thessaly Test	McMurray's Test	Apley's Test	Joint line tenderness Test
Sex	0.671	0.655	0.441	0.855	0.42	0.915	0.886	0.086
Age (years)	**0.022**	0.194	0.132	**0.023**	**0.009**	0.214	**0.01**	0.073
BMI (kg/m^2)	0.892	0.27	0.936	0.591	0.82	0.561	0.355	0.635
Arthritis	0.563	0.923	0.366	**0.047**	0.109	0.808	0.082	0.739
Other injury	**0.023**	**0.017**	**0.028**	**0.02**	0.112	0.083	**0.016**	0.304
Sports injury	0.245	0.064	0.279	0.301	**0.004**	0.796	0.122	0.399
ACL	0.99	0.326	0.274	1	0.383	0.604	0.892	0.944
Past arthritis	0.899	0.8	0.495	0.126	0.094	0.996	0.463	0.9
Previous surgery	0.971	0.347	0.913	0.994	0.422	0.413	0.825	0.696

Significant *p*-values are in bold.

TABLE 20 Influence of subject characteristics, arthritis and medical history on specificity of physical tests (*p*-value of logistic regressions)

Subject characteristic	Musculoskeletal clinicians: Thessaly Test	Musculoskeletal clinicians: McMurray's Test	Musculoskeletal clinicians: Apley's Test	Musculoskeletal clinicians: Joint line tenderness Test	Primary care clinicians: Thessaly Test	Primary care clinicians: McMurray's Test	Primary care clinicians: Apley's Test	Primary care clinicians: Joint line tenderness Test
Sex	0.636	0.823	0.564	0.188	0.216	0.095	0.281	0.052
Age (years)	0.275	0.09	0.291	0.384	0.91	**0.05**	0.159	0.485
BMI (kg/m^2)	0.39	0.752	0.118	0.222	0.319	0.576	0.873	0.911
Arthritis	0.549	0.478	0.839	0.253	0.696	0.015	0.643	0.978
Other injury	0.072	0.494	0.607	0.487	0.101	**0.01**	0.113	0.085
Sports injury	0.481	0.377	0.305	0.552	0.836	0.78	0.836	0.853
ACL	0.095	0.237	0.74	0.573	0.294	0.652	0.646	0.377
Past arthritis	0.571	0.92	0.576	0.447	0.646	0.44	0.76	0.486
Previous surgery	0.366	0.94	0.39	0.515	0.993	0.906	0.868	0.422

Significant *p*-values are in bold.

Comparison of patient subgroup recruited directly from a general practitioner practice compared with patients recruited from within an orthopaedic department

Eleven patients were recruited directly from their GP (PD) in order to allow us to verify that the cohort of patients recruited through the orthopaedic department at Glasgow Royal Infirmary were representative of the overall target population (patients attending GP with knee pathologies). The patient groups were well matched for age and BMI (*Table 21*), with only minimal differences that were not of statistical clinical relevance. There was a lower proportion of females in the group recruited directly from the GP practice (18% vs. 37%), but this was not statistically significant ($p = 0.338$). The proportion of patients in each group that had a meniscal tear diagnosed by MRI was almost the same, 57% for the group recruited directly from the GP practice and 55% for the group recruited through the orthopaedic department ($p = 1.0$).

TABLE 21 Age and BMI comparison between patients recruited directly from a GP practice and those recruited through an orthopaedic department

Patient source		Age (years)	BMI (kg/m^2)
All patients ($n = 292$)	Mean (SD)	42.6 (13.4)	27.8 (4.8)
	Min.–max.	18–71	18–47
Recruited from a GP practice ($n = 11$)	Mean (SD)	44 (15.4)	26 (2.8)
	Min.–max.	21–71	21.5–31.5
Not recruited from a GP practice ($n = 281$)	Mean (SD)	42.7 (13.3)	27.9 (4.9)
	Min.–max.	18–71	18–47
p-value	*t*-test	0.794	0.069

Max., maximum; min., minimum; SD, standard deviation.

Although the overall proportion of patients recruited directly from their GP was smaller than we had anticipated (4% vs. 10%), the group demographics and final diagnoses are very similar indicating that our main cohort is representative of the target population.

Do combinations of physical tests provide better specificity and sensitivity than a single physical test?

Analysis of individual tests for primary care clinicians is given in *Table 22*. Only McMurray's Test, when performed by a primary care clinicians, was predictive of MRI outcome, no other test or the clinical history taken by a primary care clinician was a significant predictor. Nor was the order that the tests were performed in significant (see *Table 22*). In contrast, all physical tests and the clinical history taken by musculoskeletal clinicians were significant predictors of MRI (see *Table 23*). The order that the tests were performed in was not found to be significant (*Table 23*).

The simple analysis of individual test fails to take account of any potential covariates and therefore logistic regression was undertaken. We have used a logistic regression model including all relevant factors and their interactions with a stepwise procedure to remove unimportant variables. Logistic regression analysis including all four tests, the clinical history and relevant covariates (BMI, age, sex, osteoarthritis and previous surgery) reveal a different result to the individual analyses. When all relevant factors are taken into account the only factors for primary care clinicians that are predictive of a MRI result are age and a past history of osteoarthritis (*Table 24*). Although this may be beneficial in aiding diagnosis of meniscal tears in the elderly that are associated with degenerative disease of the knee, neither of these factors are helpful in diagnosing acute traumatic injuries in young individuals participating in sports. As degenerative tears tend to be associated with knee conditions such as osteoarthritis, few degenerative tears are ever treated. In contrast, treatment is regularly offered for young patients with acute traumatic or sporting injuries.

TABLE 22 Logistic regression: four models (one model per test) with the effect of the test and its order (*p*-values and AUC), for primary care clinicians

Statistical test		Thessaly Test	McMurray's Test	Apley's Test	Joint line tenderness Test	Recorded clinical history
Effect *p*-value	Test	0.366	**0.018**	0.280	0.489	0.154
	Order	0.138	0.874	0.711	0.956	0.528
AUC		0.644	0.671	0.660	0.654	0.668

AUC, area under the curve.
Significant *p*-values are in bold.

TABLE 23 Logistic regression: four models (one model per test) with the effect of the test and its order (*p*-values and AUC), for musculoskeletal clinicians

Statistical test		Thessaly Test	McMurray's Test	Apley's Test	Joint line tenderness Test	Recorded clinical history
Effect *p*-value	Test	**0.009**	**< 0.001**	**0.019**	**< 0.001**	**< 0.001**
	Order	0.210	0.920	0.257	0.670	0.514
AUC		0.621	0.649	0.611	0.644	0.71

AUC, area under the curve.
Significant *p*-values are in bold.

TABLE 24 Multivariable logistic regression with all factors included and stepwise logistic regression for MRI outcome (tests performed by primary care clinicians)

Factor		Multivariable analysis: all factors included	Stepwise analysis
Effect *p*-value	Thessaly Test	0.486	
	McMurray's Test	0.268	
	Apley's Test	0.427	
	Joint line tenderness Test	0.545	
	Clinical history	0.283	
	Age	**< 0.001**	**0.002**
	BMI	0.652	
	Knee injury	0.437	
	Past arthritis	**0.037**	**0.041**
	Previous surgery	0.325	
	Sex	0.128	
	Arthritis	0.126	
AUC		0.722	0.651

AUC, area under the curve.
Significant *p*-values are in bold.

A variety of different clinicians of participated in this study and we had originally intended to include experience of the clinician undertaking the assessment within our models. However, classifying or grading clinician experience is not a simple process. It cannot simply be based on years since qualification as this takes no count of part-time working, maternity leave, training quality, or the number of appropriate cases each clinician has been previously exposed to. We were unable to produce a reliable method of classifying clinicians to include in our models. Instead, within the logistic regression models we included clinician as a random effect to take account of the variability between clinicians. However, this could not be specified with the stepwise regression models, where this factor was left unspecified and therefore included by default in the residual variability not accounted for by the model.

The results for musculoskeletal clinicians are potentially more clinically relevant than those observed for primary care clinicians (*Table 25*). In this clinician group the predicative factors are clinical history (history taken from the patient regarding recent symptoms and, if relevant, injury and mechanism of injury) and age.

The area under the curve (AUC) for the stepwise logistic regression (where clinician variability is accounted for in residual variability) is nearly the same as the logistic regression model with all the effects included and clinician specified as random effect (AUC 0.762 vs. 0.761), suggesting that there was nearly no difference between the different musculoskeletal clinicians. This was not the case for the primary care clinicians (AUC 0.722 vs. 0.651), suggesting a higher degree of variability between clinicians in this group.

Odds ratio for musculoskeletal clinicians diagnosis of meniscal tear based on taking a clinical history, and also for the influence of age, have been calculated and are presented in *Table 26*. A graphical representation of these probabilities for standard adult age range is shown in *Figure 11*.

TABLE 25 Multivariable logistic regression with all factors included and stepwise logistic regression for MRI outcome (tests performed by musculoskeletal clinicians)

Factor		Multivariable analysis – all factors included (*p*-value)	Stepwise analysis (*p*-value)
Effect *p*-value	Thessaly Test	0.702	
	McMurray's Test	0.322	
	Apley's Test	0.793	
	Joint line tenderness Test	0.073	
	Clinical history	**0.004**	**< 0.001**
	Age	**0.025**	**0.019**
	BMI	0.629	
	Knee injury	0.207	
	Past arthritis	0.986	
	Previous surgery	0.655	
	Sex	0.088	
	Arthritis	0.788	
AUC		0.762	0.761

AUC, area under the curve.
Significant *p*-values are in bold.

TABLE 26 Odds ratio estimates for clinical history and age

Effect	Point estimate	95% Wald confidence limits	
Clinical history taken by a musculoskeletal clinician: no vs. yes	5.910	2.954	11.826
Age	0.968	0.943	0.995

Exp, exponential function; MSC, musculoskeletal clinician; P, probability of event; RCH, recorded clinical history.
Clinical history: a patient assessed by a musculoskeletal clinician as not having a meniscal tear is 5.91 times more likely to have a negative MRI scan in comparison with a patient who has been assessed by a musculoskeletal clinician as having a meniscal tear. The probability modelled was MRI negative, but the same will be true for MRI positive.
Age: a patient is 1.033 (= 1/0.968) times more likely to be MRI positive for a meniscal tear per additional year of age.
P(MRI = 'Yes') = 1 – P(MRI = 'No')
P(MRI = 'No' if RCH_MSC = 'No') = exp(1.108 + 0.884–0.0322 × Age)/[1 + exp(1.108 + 0.884–0.0322 × Age)]
P(MRI = 'No' if RCH_MSC = 'Yes') = exp(1.108 – 0.884 – 0.0322 × Age)/[1 + exp(1.108 – 0.884 – 0.0322 × Age)]

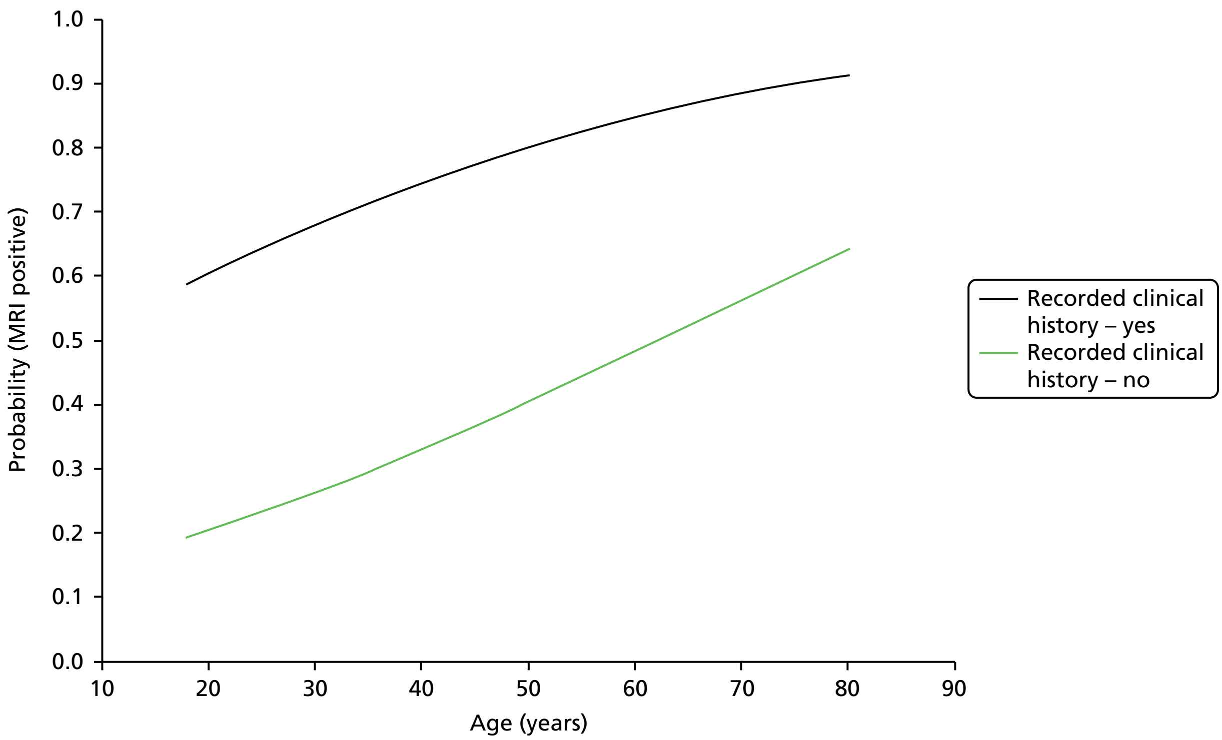


FIGURE 11 Influence of age on the probability of a meniscal tear diagnosis by MRI when a musculoskeletal clinician diagnoses a meniscal tear based on a clinical history.

Validation of magnetic resonance imaging diagnosis for meniscal tears using knee arthroscopy

Seventy-seven patients had an arthroscopy as part of their routine clinical care. When arthroscopic diagnosis of meniscal tears is taken as the gold standard and compared with results obtained with MRI, the sensitivity of MRI was 1.0, whereas the specificity was 0.53 (*Table 27*).

The specificity of 0.53 (95% CI 0.28 to 0.77) is lower than reported in the literature for MRI diagnosis of meniscal tears. Only eight patients had a disagreement between their MRI diagnosis and their arthroscopic findings. In three patients the radiologist's report of the MRI scan reported a 'possible' meniscal tear. We have included in the analysis any patient with these 'possible' meniscal tears, because this reflects the real life situation where diagnosis, even on MRI scan, is not clear cut. Diagnosis on MRI is an interpretation of scan findings and although the results tend to be reported using a binary approach, MRI itself is certainly not binary and is highly dependent on the experience of the radiologist reporting the scan.

TABLE 27 Magnetic resonance imaging vs. arthroscopy (95% CI)

Statistical test	MRI/arthroscopy
Sensitivity	1.00 (0.93 to 1.00)
Specificity	0.53 (0.28 to 0.77)
LR+	2.12 (1.28 to 3.52)
LR–	0 (not estimable)
OR	Not estimable
PPV	0.40 (0.32 to 0.49)
NPV	1.00 (0.66 to 1.00)

LR–, likelihood ratio for negative test; LR+, likelihood ratio for positive test; OR, odds ratio.

In a further three cases the patients had previously received meniscal knee surgery. We believe that the explanation for the discordant results in these patients is that the MRI scan still shows an abnormal signal in the areas of the knee that previously underwent surgery; however, on arthroscopic visualisation any previous meniscal damage appears to have been repaired and therefore a result of no meniscal tear was recorded. Removing patients who have had previous knee surgery improved the specificity from 0.53 to 0.60 (95% CI 0.32 to 0.84). Similarly, including only those patients who were classified as not having a meniscal tear based on diagnosis by a musculoskeletal clinician improved the specificity to 0.67. This suggests that indiscriminate use of MRI to diagnose meniscal injuries is unlikely to be beneficial and a more targeted approach to use of MRI will yield more accurate and beneficial data.

Patient and public involvement

Patients were not directly involved in the design or analysis of the study. We had intended to recruit a suitable patient to the trial management group prior to initiation of the study. However, we were unable to recruit a candidate with appropriate patient feedback, and opinion was sought on the individual physical tests used within the study. Each patient who underwent examination was asked to give their view on the tests (the Thessaly Test, McMurray's Test, Apley's Test and the joint line tenderness Test), particularly in relation to the pain experienced during the test.

Overall, 72% of patients said that one or more tests were painful when performed. In particular the joint line tenderness Test (57.1% of patients) and the Thessaly Test (50.5%) were found to cause pain during the test procedure. In comparison McMurray's Test and Apley's Test were reported as causing pain less frequently, 37.7% and 33.3% respectively.

Chapter 5 Discussion

In the UK approximately 30% of GP appointments are for musculoskeletal complaints, with knee problems being one of the most common problems.[33] Complicated forces are exerted through the knee during various activities, making it susceptible to a range of acute and chronic injuries. The mean annual incidence of meniscal tears in the knee may be as high as 60–70 per 1000 population, with a 2.5 to 4 times male to female predominance.[3] The age distribution of patients with meniscal tears shows a bimodal peak with acute injuries peaking in young active patients in their mid-to-late twenties and chronic degenerative injuries becoming more prevalent in old age. Patients with meniscal tears present both acutely to the emergency department at the time of injury, but also commonly post injury to their GP. The latter scenario usually occurs if the initial injury has failed to settle and symptoms remain.

This study was designed to evaluate whether or not the Thessaly Test, alone or in combination with other physical tests, could be used by primary care clinicians (GPs, community physiotherapists, etc.) to rule out a meniscal tear in patients presenting with knee pain. This would allow targeted onward referral to secondary care and minimising the utilisation of expensive and overburdened diagnostic imaging such as MRI scanning.

Several previous studies have examined the diagnostic accuracy of the Thessaly Test since the original publication by Karachalios *et al.*[25] In a retrospective study of 116 patients, Harrison *et al.*[27] validated the Thessaly Test against arthroscopy concluding that the Thessaly Test is a valid a reproducible test for diagnosis of meniscal tears. Harrison *et al.*[27] found a sensitivity of 90% and a specificity of 98%. However, the study patient group was preselected by an experienced musculoskeletal clinician and all patients enrolled in the study were classified as having a suspicion of a meniscal tear and were on a waiting list for arthroscopic surgery. Importantly, this study group is therefore not comparable with our desired target group which encompasses all knee patients likely to be encountered by a non-specialist primary care clinician.

In a similar study, Konan *et al.*[28] reviewed 121 patients and used both MRI imaging and arthroscopy, with arthroscopy set as the 'gold standard'. Sensitivity and specificity for medial meniscal tears was 83% and 76%, respectively, though specificity dropped in the presence of ACL injury to just 56%. They conclude that the Thessaly Test alone is not beneficial in diagnosis of meniscal tears, but that in combination with McMurray's Test it may be useful. As with the cohort of patients selected for Harrison *et al.*'s study,[27] the cohort used by Konan *et al.*[28] were preselected by an experienced musculoskeletal clinician and all patients enrolled in the study were classified as having a high suspicion of a meniscal tear.

Finally, Mirzatolooei *et al.*[29] have examined the specificity and sensitivity of the Thessaly Test in diagnosing meniscal tears in a cohort of patients with ACL injuries ($n = 80$). They found a sensitivity of 79%, but low specificity of just 40%. They conclude that low specificity of the Thessaly Test renders it unsuitable for this cohort of patients with significant additional knee pathology.

There are several limitations to our study. The use of MRI as the diagnostic 'gold standard' has limitations as diagnostic accuracy of MRI for meniscal tears is not 100%. The only truly accurate method of diagnosing meniscal tears is to perform an arthroscopy, but for obvious ethical reasons it was not feasible to perform an invasive arthroscopy of the knee on all patients with suspected knee pathology. The data generated during this study suggest that when MRI is used to detect meniscal tears in a diverse group of knee pathology patients the accuracy of MRI is lower than is often reported in the literature. Other authors have also commented on this issue with Mirzatolooei *et al.*[29] estimating that relying blindly on MRI to determine surgical intervention could result in inappropriate treatment in up to 35% of the cases.

Although the recruitment to the study reached its target, there were a significant number of patients who did not attend for MRI of the knee resulting in us being unable to include their data in the final analysis. The dropout rate for the knee pathology group was just 18%, but the control group had a dropout rate of 43%. In order to compensate for this we recruited an additional 25 control subjects, giving us usable data on 43 subjects, which was close to the target value of 50 control subjects. We believe that there were two reasons for the large discrepancy in the dropout rate between the two study groups. First, the control group were unlikely to derive any personal benefit from attending for their MRI scan and therefore were more likely to dropout. Second, around a third of our control subjects had their MRI scan delayed due to a lack of capacity in imaging reporting. This undoubtedly affected the willingness of volunteers to attend for scanning when appointments were rearranged. Although the number of patients with data available for final analysis was below our target, post-hoc power calculations have demonstrated that the study is still appropriately powered.

We had originally intended to recruit our control group from only an orthopaedic hand clinic. However, recruitment rates were slower than anticipated and so we utilised local adverts in order to increase the recruitment rate for control subjects. We did not undertake any comparison between those recruited through the clinics and those responding to the adverts.

The final potential weakness of the study was undertaking the study in a secondary care setting when the target population was patients in primary care. Although our study was set in a secondary care environment, we have attempted to replicate the typical patient group that a primary care clinician might encounter (i.e. a diverse cohort of patients with multiple knee pathologies in whom a non-specialist may struggle to differentiate meniscal tears). This methodology was chosen in order to allow us to recruit sufficient numbers of patients to perform the study and to avoid the need for a large network of GP practices to participate in the study, each of whom would only have contributed a few patients. In order to validate that our cohort was representative of the types of patients who a typical GP may encounter, we asked one GP to directly recruit all patients who presented to him during the study period with soft tissue knee pain. This group represented just under 5% of the overall cohort and were well matched to the main cohort recruited from the orthopaedic department at Glasgow Royal Infirmary indicating that the main group recruited in secondary care were representative of the target population. Although we recruited fewer patients directly from GP practices than we had specified in the original study protocol, it is unlikely that this influenced the overall study results. This element of the study was simply to ensure that demographics of patients recruited in secondary care matched those that would have been recruited from primary care had the study been run in this setting. If anything, the lower than anticipated number of patients recruited from GP practice vindicates the decision to run the study in secondary care and serves to highlight the issues that would have been encountered if we had opted to use a large network of primary care units for recruitment rather than a secondary care setting.

Our results have demonstrated that the Thessaly Test used in isolation is not superior to existing physical tests for diagnosing meniscal tears. None of the physical tests examined is sufficiently sensitive or specific to use as a routine alternative to MRI, or referral to a musculoskeletal specialist. We found that no combination of tests, when used by primary care clinicians, was suitable as an alternative to MRI diagnosis or referral to a musculoskeletal clinician for assessment. Our data demonstrates that even in the hands of experienced musculoskeletal clinicians, physical tests for meniscal tears have limited value if used indiscriminately.

Interestingly, there was a large disparity between the two groups of clinicians with respect to the proportion of patients classified as being too painful to perform the test on. Primary care clinicians opted not to perform the Thessaly Test on 10% of patients compared with 17% for musculoskeletal clinicians. The reason that the test was not completed was not formally recorded as part of the study, but anecdotally it was most commonly secondary to knee pain. We can speculate that experience levels in dealing with musculoskeletal complaints influenced the threshold for stopping the test due to pain. We hypothesise that more experienced clinicians realise that little value will be gained from performing the test in patients with high levels of knee pain. It is also possible that the primary care clinicians who were brought in specifically to undertake study assessments, rather than providing direct clinical care, may have been more likely to adhere to the study protocol. The relatively high rate of 'test unable to be performed' further limits the usefulness of the Thessaly Test as a generic tool.

In addition to testing the accuracy of diagnosis of meniscal tears using physical tests, our protocol also included an assessment of the accuracy of diagnosis based on a standardised clinical history. We developed a simple minimal data set to be collected as part of the routine clinical history, which was designed to aid the correct diagnosis. The design was deliberately kept simple so that if it was successful in the hands of primary care clinicians it could easily be utilised nationally with no need for an expensive educational programme to support the roll out of the concept.

Although both groups of clinicians received basic training in the use of this diagnostic tool our multivariate analysis showed that it was either more effective in the hands of specialist knee clinicians or that it was insufficient on its own as a surrogate for specialist experience.

We have validated our MRI findings in a subset of patients using arthroscopy as the 'gold standard'. Our initial validation of MRI against arthroscopy revealed high sensitivity (1.0), but low specificity (0.53). This was surprising given the good specificity of MRI reported in the literature.[18] The Thessaly Test validation study by Konan *et al.*[28] used both MRI and arthroscopy. Konan *et al.*[28] found that MRI detected 96% of tears which is in line with other reports validating MRI against arthroscopy.[18,28] However, this was in a preselected group of patients with a high suspicion of meniscal tears, rather than the more general group that we have studied. We believe that our study cohort is more representative of the patient cohort that an average primary care clinician may encounter and that our results are of greater relevance to this group of clinicians than previous studies of highly selected patient groups.

We are not the first group to document the lack of specificity of MRI for diagnosis of meniscal tears in the presence of concomitant knee pathology. Several other authors have reported reduced accuracy in the presence of acute ACL rupture.[24,34]

Although MRI accuracy is clearly influenced by multiple knee pathologies, it is also influenced by patient age. England *et al.*[13] have reported high rates of incidental meniscal findings on MRI of the knee in the general population and that these findings tend to increase with age. In middle-aged and elderly patients a lower threshold of suspicion should be applied for meniscal tears as they tend to follow minor trauma and MRI signal changes are significantly higher in the elderly population.

Few studies have examined diagnostic accuracy of MRI in a community-based setting; however, Hardy *et al.*,[35] in the USA, have reported community-based recording of MRI for diagnosis of meniscal tears to have a sensitivity of 73% and specificity of 68%, which suggests that this imaging modality is of less use in a diverse poorly stratified cohort and perhaps corroborates our data.

Our study data shows that the most accurate non-imaged method of diagnosing meniscal tears is to use a clinical history taken by an experience musculoskeletal clinician. Mohan and Gosal[36] have reported similar findings suggesting that the accuracy of clinical diagnosis by an experienced musculoskeletal clinician based on examination and history is 88% for medial meniscal tears and 92% for lateral meniscal tears, with similar results reported by Ercin *et al.*[37] These results are comparable with that achieved by MRI, suggesting that MRI should be reserved for use in more doubtful, difficult or complex knee injuries.[36]

The data generated by this study suggest that the Thessaly Test is no better at diagnosing meniscal tears than other established physical tests. We found that neither the Thessaly Test alone or in combination with other physical tests could be reliably used by primary care clinicians as an alternative, or surrogate, for MRI scanning to diagnose meniscal tears in the knee.

We have also found that MRI diagnosis of meniscal tears in comparison with arthroscopic diagnosis was less accurate than is commonly reported in the literature. This was particularly true when MRI is used indiscriminately rather than in targeted populations of patients with specific 'suspected' knee pathologies. Further well-defined studies are required to determine the true accuracy of MRI for the diagnosis of meniscal tears.

The findings of this study have implications for primary care clinicians who may wish to consider referral of patients with significant knee pain and injuries to a specialist clinician for assessment and treatment rather than automatic referral for MRI of the knee. This approach is likely to maximise the benefit from MRI and minimise the cost of unnecessary or inappropriate imaging, though further cost-effectiveness studies are required to validate this assumption.

Chapter 6 Conclusion

The data generated by this study suggest that the Thessaly Test is no better at diagnosing meniscal tears than other established physical tests. We found that neither the Thessaly Test alone or in combination with other physical tests could be reliably used by primary care clinicians as an alternative, or surrogate, for MRI scanning to diagnose meniscal tears in the knee.

We have also found that MRI diagnosis of meniscal tears in comparison with arthroscopic diagnosis was less accurate than is commonly reported in the literature. This was particularly true when MRI is used indiscriminately rather than in targeted populations of patients with specific 'suspected' knee pathologies. Further well-defined studies are required to determine the true accuracy of MRI for the diagnosis of meniscal tears.

The findings of this study have implications for primary care clinicians who may wish to consider referral of patients with significant knee pain and injuries to a specialist clinician for assessment and treatment rather than automatic referral for MRI of the knee. This approach is likely to maximise the benefit from MRI and minimise the cost of unnecessary or inappropriate imaging.

Recommendation for further research

Further research is required to determine the true accuracy and cost-effectiveness of MRI for the detection of meniscal tears, in a cohort of patients who has not been highly selected by experienced specialists.

Acknowledgements

We gratefully acknowledge the funding received for this study from the National Institutes of Health Research HTA programme, grant reference 09/163/02.

We would like to acknowledge the help and support provided by Miss Sinead O'Donnell and Sister Emma Sharp (Orthopaedic Department, Glasgow Royal Infirmary), Mr Alun Bevan (Medical Illustration Department at Glasgow Royal Infirmary) and the Extended Scope Physiotherapy Team at Glasgow Royal Infirmary (Linda Mercer, Jenny Love, Val Forgie, John O'Toole, Alan Krawcyzk, Brenda Danks, Margot Cohen, Lynn Smyth and Jo Chambers).

We would also like to acknowledge all of the primary care clinicians who participated in this study.

Contributions of authors

Mr Mark Blyth (Consultant Orthopaedic Surgeon, Glasgow Royal Infirmary), as chief investigator had overall responsibility for the study.

Dr Iain Anthony (Senior Clinical Research Manager, Glasgow Royal Infirmary) wrote the initial grant application, wrote the final report and was responsible for study co-ordination and set-up.

Dr Bernard Francq (Statistician, Robertson Centre for Biostatistics, University of Glasgow) conducted the statistical analysis of the study.

Dr Katriona Brooksbank (Research Assistant, Glasgow Royal Infirmary) was responsible for patient recruitment and data collection.

Dr Paul Downie (General Practitioner) was responsible for recruitment of patients directly from GP practice, he also undertook study clinician assessments.

Mr Andrew Powell (Orthopaedic Surgical Trainee, Glasgow Royal Infirmary) was responsible for undertaking osteoarthritis grading of X- ray images.

Mr Bryn Jones (Consultant Orthopaedic Surgeon, Glasgow Royal Infirmary) aided in editing of the final report.

Mr Angus MacLean (Consultant Orthopaedic Surgeon, Glasgow Royal Infirmary) aided in editing of the final report.

Dr Alex McConnachie (Assistant Director of Biostatistics, Robertson Centre for Biostatistics, University of Glasgow) contributed to the development of the statistical analysis plan and supervised the statistical analysis.

Professor John Norrie (Director of the Centre for Randomised Health Care Trials, University of Aberdeen) generated the original statistical analysis plan and aided in report editing.

Data sharing statement

Data can be obtained from the corresponding author on request.

References

1. Messner K, Gao J. The menisci of the knee joint. Anatomical and functional characteristics, and a rationale for clinical treatment. *J Anat* 1998;**193**:161–78. http://dx.doi.org/10.1046/j.1469-7580.1998.19320161.x
2. Greis PE, Bardana DD, Holmstrom MC, Burks RT. Meniscal injury: I. Basic science and evaluation. *J Am Acad Orthop Surg* 2002;**10**:168–76.
3. Maffulli N, Longo UG, Campi S, Denaro V. Meniscal tears. *Open Access J Sports Med* 2010;**1**:45–54. http://dx.doi.org/10.2147/OAJSM.S7753
4. Arnoczky SP, Warren RF. Microvasculature of the human meniscus. *Am J Sports Med* 1982;**10**:90–5. http://dx.doi.org/10.1177/036354658201000205
5. de Albornoz PM, Forriol F. The meniscal healing process. *Muscles Ligaments Tendons J* 2012;**2**:10–18.
6. Fox AJ, Bedi A, Rodeo SA. The basic science of human knee menisci: structure, composition, and function. *Sports Health* 2012;**4**:340–51. http://dx.doi.org/10.1177/1941738111429419
7. Majewski M, Susanne H, Klaus S. Epidemiology of athletic knee injuries: a 10-year study. *Knee* 2006;**13**:184–8. http://dx.doi.org/10.1016/j.knee.2006.01.005
8. Barrett GR, Field MH, Treacy SH, Ruff CG. Clinical results of meniscus repair in patients 40 years and older. *Arthroscopy* 1998;**14**:824–9. http://dx.doi.org/10.1016/S0749-8063(98)70018-0
9. Englund M, Niu J, Guermazi A, Roemer FW, Hunter DJ, Lynch JA, *et al.* Effect of meniscal damage on the development of frequent knee pain, aching, or stiffness. *Arthritis Rheum* 2007;**56**:4048–54. http://dx.doi.org/10.1002/art.23071
10. Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. *Arthritis Rheum* 2003;**48**:2178–87. http://dx.doi.org/10.1002/art.11088
11. Hunter DJ, Zhang YQ, Niu JB, Tu X, Amin S, Clancy M, *et al.* The association of meniscal pathologic changes with cartilage loss in symptomatic knee osteoarthritis. *Arthritis Rheum* 2006;**54**:795–801. http://dx.doi.org/10.1002/art.21724
12. Ding C, Martel-Pelletier J, Pelletier JP, Abram F, Raynauld JP, Cicuttini F, *et al.* Meniscal tear as an osteoarthritis risk factor in a largely non-osteoarthritic cohort: a cross-sectional study. *J Rheumatol* 2007;**34**:776–84.
13. Englund M, Guermazi A, Gale D, Hunter DJ, Aliabadi P, Clancy M, *et al.* Incidental meniscal findings on knee MRI in middle-aged and elderly persons. *N Engl J Med* 2008;**359**:1108–15. http://dx.doi.org/10.1056/NEJMoa0800777
14. Allum R. Complications of arthroscopy of the knee. *J Bone Joint Surg Br* 2002;**84**:937–45. http://dx.doi.org/10.1302/0301-620X.84B7.13637
15. DeLee JC, the Committee on Complications of Arthroscopy Association of North America. Complications of arthroscopy and arthroscopic surgery: results of a national survey. Committee on Complications of Arthroscopy Association of North America. *Arthroscopy* 1985;**1**:214–20. http://dx.doi.org/10.1016/S0749-8063(85)80086-4
16. Armstrong RW, Bolding F, Joseph R. Septic arthritis following arthroscopy: clinical syndromes and analysis of risk factors. *Arthroscopy* 1992;**8**:213–23. http://dx.doi.org/10.1016/0749-8063(92)90039-E

17. 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

18. Oei EH, Nikken JJ, Verstijnen AC, Ginai AZ, Myriam Hunink MG. MR imaging of the menisci and cruciate ligaments: a systematic review. *Radiology* 2003;**226**:837–48. http://dx.doi.org/10.1148/radiol.2263011892

19. Englund M, Guermazi A, Lohmander LS. The meniscus in knee osteoarthritis. *Rheum Dis Clin North Am* 2009;**35**:579–90. http://dx.doi.org/10.1016/j.rdc.2009.08.004

20. Bhattacharyya T, Gale D, Dewire P, Totterman S, Gale ME, McLaughlin S, *et al.* The clinical importance of meniscal tears demonstrated by magnetic resonance imaging in osteoarthritis of the knee. *J Bone Joint Surg Am* 2003;**85**:4–9.

21. Kornaat PR, Bloem JL, Ceulemans RY, Riyazi N, Rosendaal FR, Nelissen RG, *et al.* Osteoarthritis of the knee: association between clinical features and MR imaging findings. *Radiology*. 2006;**239**:811–17. http://dx.doi.org/10.1148/radiol.2393050253

22. Kocabey Y, Tetik O, Isbell WM, Atay OA, Johnson DL. The value of clinical examination versus magnetic resonance imaging in the diagnosis of meniscal tears and anterior cruciate ligament rupture. *Arthroscopy* 2004;**20**:696–700. http://dx.doi.org/10.1016/j.arthro.2004.06.008

23. Rossi R, Dettoni F, Bruzzone M, Cottino U, D'Elicio DG, Bonasia DE. Clinical examination of the knee: know your tools for diagnosis of knee injuries. *Sports Med Arthrosc Rehabil Ther Technol* 2011;**3**:25. http://dx.doi.org/10.1186/1758-2555-3-25

24. Fowler PJ, Lubliner JA. The predictive value of five clinical signs in the evaluation of meniscal pathology. *Arthroscopy* 1989;**5**:184–6. http://dx.doi.org/10.1016/0749-8063(89)90168-0

25. 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**:955–62. http://dx.doi.org/10.2106/JBJS.D.02338

26. Hegedus EJ, Cook C, Hasselblad V, Goode A, McCrory DC. Physical examination tests for assessing a torn meniscus in the knee: a systematic review with meta-analysis. *J Orthop Sports Phys Ther* 2007;**37**:541–50. http://dx.doi.org/10.2519/jospt.2007.2560

27. 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

28. Konan S, Rayan F, Haddad FS. Do physical diagnostic tests accurately detect meniscal tears? *Knee Surg Sports Traumatol Arthrosc* 2009;**17**:806–11. http://dx.doi.org/10.1007/s00167-009-0803-3

29. 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**:221–3. http://dx.doi.org/10.1016/j.knee.2009.08.007

30. Apley AG. The diagnosis of meniscus injuries; some new clinical methods. *J Bone Joint Surg Am* 1947;**29**:78–84.

31. Chivers MD, Howitt SD. Anatomy and physical examination of the knee menisci: a narrative review of the orthopedic literature. *J Can Chiropr Assoc* 2009;**53**:319–33.

32. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;**16**:494–502. http://dx.doi.org/10.1136/ard.16.4.494

33. Department of Health. *The Musculoskeletal Services Framework*. 2006. URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4138413 (accessed 8 June 2015).

34. Nam TS, Kim MK, Ahn JH. Efficacy of magnetic resonance imaging evaluation for meniscal tear in acute anterior cruciate ligament injuries. *Arthroscopy* 2014;**30**:475–82. http://dx.doi.org/10.1016/j.arthro.2013.12.016

35. Hardy JC, Evangelista GT, Grana WA, Hunter RE. Accuracy of magnetic resonance imaging of the knee in the community setting. *Sports Health* 2012;**4**:222–31. http://dx.doi.org/10.1177/1941738111434396

36. Mohan BR, Gosal HS. Reliability of clinical diagnosis in meniscal tears. *Int Orthop* 2007;**31**:57–60. http://dx.doi.org/10.1007/s00264-006-0131-x

37. Ercin E, Kaya I, Sungur I, Demirbas E, Ugras AA, Cetinus EM. History, clinical findings, magnetic resonance imaging, and arthroscopic correlation in meniscal lesions. *Knee Surg Sports Traumatol Arthrosc* 2012;**20**:851–6. http://dx.doi.org/10.1007/s00167-011-1636-4

Appendix 1 Statistical analysis plan

DIAGNOSTIC ACCURACY OF THE THESSALY TEST, THE STANDARDISED CLINICAL HISTORY, AND OTHER CLINICAL EXAMINATION TESTS FOR MENISCAL TEARS

STATISTICAL ANALYSIS PLAN

1. INTRODUCTION

1.1. STUDY BACKGROUND

The menisci are two semilunar, fibrocartilaginous disks located between the medial and lateral articular surfaces of the femur and tibia in each knee. The menisci play an important role in the function of the knee providing load bearing, stress distribution and shock absorption across the knee. Tears in the menisci are a common knee injury that can cause pain in the joint. In younger active patients tears are often a result of sports injuries. In older people degenerative meniscal tears are more common. Reliable non-invasive diagnosis of meniscal tears is difficult. There are a number of physical examination tests that diagnose tears but all suffer a lack of specificity (the correct identification of those that do not have a meniscal tear) and sensitivity (the correct identification of those that do have a meniscal tear).

MRI is often referred to as the gold standard for non-invasive diagnosis of meniscal tears. However, incidental meniscal findings on MRI of the knee are common in the general population, increase with age and may not be associated with pain. Meniscal damage is also a frequent finding on MRI of the osteoarthritic knee limiting the value of this diagnostic tool for meniscal tears in this section of the population.

The Thessaly test is a clinical examination used to detect miniscal tears in the knee. Established alternative tests to the Thessaly test include the McMurray test, Apley's test and joint line tenderness test. Previous reports have come to the conclusion that a combination of tests is required to produce accurate diagnoses.

The accuracy of specialist knee clinicians in performing physical examinations of the knee may differ to primary care staff who will inevitably see fewer patients with knee pathology and have less training in performing tests.

1.2. STUDY OBJECTIVES

The objectives of the study are:

- to determine the diagnostic accuracy of the Thessaly test by GPs for meniscal tear in the knee and whether this test can obviate the need for further investigation by arthroscopy or magnetic resonance imaging (MRI);
- to determine how the Thessaly test compares to clinical history and other commonly used physical examinations (McMurray test, Apley's test, joint line tenderness test) in diagnosing meniscal tears by GPs;
- to determine if the presence of arthritis or other knee pathologies influences the accuracy of the Thessaly test;
- to determine if the use of combinations of physical tests (such as the Thessaly test, McMurray test, Apley's test and or joint line tenderness test) by GPs provides better specificity and sensitivity than a single test alone in the diagnosis of meniscal tear;
- to determine the ability of non-specialist General Practitioners (GPs) to use the Thessaly test in comparison to specialist knee clinicians.

1.3. STUDY DESIGN

This is a single centre (Glasgow Royal Infirmary) observational diagnostic study. 300 patients will be attending knee clinics at Glasgow Royal Infirmary, and have suspected knee pathology. 5-10% of this group will be enrolled via a single general practice, and will be all patients presenting at the GP with knee symptoms, and will be sub-analysed to check comparability of the wider group with the primary care population. 50 patients will be attending orthopaedic hand clinics, and have no suspected knee pathology, acting as controls.

All participants (who will attend weekly knee clinics at the Glasgow Royal Infirmary) will be assessed using the Thessaly test, McMurray test, Apley test and joint line tenderness test, independently by orthopaedics specialist clinicians and GPs. The order of these 4 tests will be randomly permuted. Likewise if feasible the order of specialist clinician and GP. All participants with will undergo MRI scan and knee x-ray (to identify the subgroup of patients with arthritis in the knee; control subjects will not have knee x-ray). All participants will have a medical history taken (with half randomly assigned to take the medical history before the tests, half after the tests). Arthroscopy will be performed only on patients who would normally receive this as part of their standard care.

There will be 3 specialist orthopaedic clinicians and 10 general practitioners. Each patient will be assessed by one specialist orthopaedic clinician and one GP. It is expected that each specialist orthopaedic clinician and each GP will assess roughly equal numbers of patients.

The GPs and specialist orthopaedic clinicians will be unaware of each others test results and also the referent gold standard MRI test and the X-ray test to establish arthritis in the knee.

1.4. Sample Size and Power

The following sample size justification is given in the study proposal:

> Assuming the sensitivity of the Thessaly test is around 75%, the study would need around 300 subjects to estimate the sensitivity to within +/- 5%. A similar calculation for the width of the confidence interval for a binomial proportion is appropriate for the specificity – for example, if the specificity was around 90%, the required sample size to estimate the specificity to within +/- 8% would be n=50 participants. The power for the pairwise comparison of tests, or combinations of tests, will depend on the degree of disagreement between the tests – for example, with around 220 pairs of measurements the study would have 90% power to detect a difference in proportions of 0.10 when the proportion of discordant pairs is expected to be 0.15 (using McNemar's test).

1.5. Study Population

1.5.1. Inclusion Criteria

Knee pain group (N=300):

- Patients referred to the knee clinic at Glasgow Royal Infirmary.

Control group (N=50):

- Patients attending the hand clinic at Glasgow Royal Infirmary.

1.5.2. Exclusion Criteria

Knee pain group (N=300):

- Age under 18;
- Unable to give informed consent;
- Previous knee replacement.

Control group (N=50):

- Age under 18;
- Unable to give informed consent;
- Previous knee surgery;
- History of knee pain (last 6 months);
- Osteoarthritis;

- Rheumatoid arthritis.

1.6. Statistical Analysis Plan (SAP)

1.6.1. SAP Objectives

The objective of this SAP is to describe the statistical analyses to be carried out for the study titled "Diagnostic accuracy of the Thessaly Test, the standardised clinical history, and other clinical examination tests for meniscal tears".

1.6.2. Current Protocol

At the time of writing, no formal study protocol has been written. This document is based on the study proposal, submitted as a full proposal to the NIHR HTA Commissioning Board in September 2010, and approved for funding in June 2011. Future development of the protocol will inform subsequent versions of this SAP, which will be updated as necessary.

1.6.3. General Principles

Results will be presented for the study population as a whole and separately for the knee pain and control groups. Within the knee pain group, results will also be presented separately for the subgroup of patients referred from a single general practice in comparison to all other patients, for those with and without arthritis, and in other subgroups according to knee pathology (Anterior Cruciate Ligament (ACL) rupture, or other previous injury or treatment of the knee) and patient characteristics (including the subgroup of predominantly younger patients with sports injury and the subgroup of those with degenerative changes due to age).

Diagnostic performance measures will be calculated for each individual test (Thessaly test, McMurray test, Apley test, joint line tenderness test and clinical history), using evidence of meniscal tear on MRI as the referent (gold standard) test. It is recognised that this MRI gold standard is itself imperfect. However, it is the established best diagnostic tool available on which intervention and treatment decisions are made, and no feasible alternative exists or is available. The primary interest will be the performance of these tests when used by GPs. Results will be reported for the tests performed by Orthopaedic Clinicians, and compared to the performance achieved by GPs.

Combinations of physical tests will be considered, to determine the optimal combination for the diagnosis of meniscal tear. Logistic regression methods will be used to determine whether the addition of patient characteristics to the results of physical tests provides greater discriminatory ability.

1.6.4. Software

Statistical analyses will be carried out using SAS for Windows v9.2, R for Windows v 2.12.1 or SPlus for Windows v8.1, or higher versions of these programs.

2. Analysis

2.1. Study Populations

The numbers of people screened and the numbers and percentages recruited in each study population will be presented, as will the numbers providing data for each diagnostic test. Numbers of participants not completing the study according to the protocol, with reasons for non-completion, will be presented.

2.2. Baseline Characteristics

Summary tables will be presented, describing the baseline characteristics of each study population. Appropriate statistical tests will be used to compare the different populations. Similar summaries and tests will be used to describe population subgroups of particular interest.

2.3. Diagnostic Test Results

2.3.1. Descriptive Statistics

The numbers and percentages of individuals classified as having meniscal tears according to each test will be presented for each study population and in subgroups of particular interest. Results of physical tests performed by GPs and Orthopaedic Clinicians will be presented separately and compared with exact McNemar tests.

2.3.2. Single Tests

The diagnostic properties of the tests will be summarised using standard techniques for diagnostic studies as described by Pepe, 2003.

The sensitivity (Sens) and specificity (Spec) of each physical test will be presented, along with the positive and negative likelihood ratios (LR+ and LR-), and the diagnostic odds ratio (DOR) with exact 95% confidence intervals, treating the MRI result as the true diagnosis. This is for the group with knee symptoms. For the controls from the hand clinic, we do not expect any positive meniscal tear diagnosis on MRI – for this group the objective is to compare specificities.

Likewise, we will also calculate the positive predictive value (PPV) and the negative predictive value (NPV) of the 4 simple tests, along with the appropriate exact 95% confidence intervals, to summarise what a positive and negative test tells us in those that have knee symptoms, assuming this represents the population of patients in primary care with suspected meniscal tear. The appropriateness of the assumption will be assessed by comparison with the subgroup of consecutive patients with knee symptoms referred from a single general practice, and consideration given to reweighting the estimates via an appropriate statistical model to account for any systematic differences between the two populations, if necessary. The characteristics compared will include age, gender, socioeconomic status, and various medical history items. These will be compared using t-tests and chi-squared tests as appropriate.

We will plot the Sensitivity vs. 1 – Specificity for the 4 simple tests in the group with knee symptoms to visualise their relative performance.

The performance of tests performed by GPs and Orthopaedic Clinicians will be presented separately. The main interest is in the performance of the GPs. The performance of these tests at the hands of the specialist orthopaedic clinicians is expected to indicate an upper bound on their potential performance. The GP and specialist orthopaedic clinicians performance will be compared with exact McNemar tests.

Physicians' views on the use of the different physical tests will be summarised and compared between tests.

2.3.3. Combined Tests

The diagnostic performance of alternative combinations of physical tests will be estimated. We will use various methodological approaches as discussed in Knotterus, 2009:

- Logistic regression, with MRI classification (meniscal tear, Yes/No) as the outcome, will be used to build a series of models on the GP's performance to assess the diagnostic properties, as follows:
  - Core model: including 'design' information (indicator variables for the randomised order of the tests, randomised order of taking the medical history), and GP as a random effect.
  - Model Level 1: The Core model with an individual test in isolation (4 models)
  - Model Level 2: The Core model with participant baseline covariates (age, sex, previous history, socioeconomic status, and so on) (1 model)

- Model Level 3: Re-do model Level 1 with Model level 2 covariates (4 models)
- Model Level 4: Explore GP characteristics as influences e.g. age (or time since qualified), gender, specialities, GP status (e.g. partner), GP surgery characteristics (e.g. number of partners), GP practice size, and so on.
- Model Level 5: Stepwise selection model to establish parsimonius model combining GP and patient level predictors to provide Updated Core Model
- Model Level 6: Investigation of combinations of pairs of 2, triplets of 3 and all 4 tests combined in the presence of the Updated Core Model.

All the models will be assessed by their concordance index (c-statistics) measuring the area under the curve. When considering whether an increment in the c-statistic moving from one model to the next is worthwhile, due allowance will be made for the increased complexity of the model.

We will consider adding in the patient-defined subgroups (such as arthritis (yes/no), ACL rupture (yes/no), sports injury (yes/no), degenerative disease (yes/no) as subgroups of particular interest in the development of these models, and formally test for interactions as appropriate.

We will consider re-running this modelling hierarchy for the specialist orthopaedic clinicians data.

.

- Classification And Regression Trees (CART, as implemented in R) will also be used to determine an optimal combination of tests provides better prediction. The advantage of this approach is that it allows complex interactions between the four tests which the logistic regression approach isn't naturally suited for. The disadvantage is that CART is purely data driven, and hence often produces solutions which do not transfer to the next dataset. We will look at 'averaging' trees across split samples using resampling techniques to try to overcome this and produce stable, robust trees.

  As with the logistic regressions, we may will investigate specific subgroups of interest and possibly re-run on the specialist orthopaedic clinicians dataset.

2.3.4. Reporting

The study will be reported to the standards established in the STARD initiative (Bossuyt et al, 2003)

2.3.5. missing data

We do not anticipate any missing data arising from any issues regarding the co-operation or availability of the specialists clinicians or the GPs – the sessions at which the measurements will be taken will be arranged in advance to suit these health professionals. It may happen that not all 300 of those with knee symptoms or all the 50 non-knee symptom controls are not available for all measurements – we will endeavour to make sure we reach these targets. In terms of baseline covariate measurements on both participants and clinicians, these are all simple information and again we do not anticipate not having full information on everyone. As such it is not anticipated that missing data will be an important issue in this study, so we will simply describe what if anything is missing and we have no special plans for dealing with missing data in the analyses.

3. References

Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. OUP (2003).

The Evidence Base of Clinical Diagnosis: Theory and Methods of Diagnostic Research. Editors Knottnerus JA, Buntrix ID. Chapter 8 – Multivariate analyses in diagnostic accuracy studies: what are the possibilities? BMJ Books, Wiley-Blackwell (2009)

Bossuyt PM, Reitsma JB, Bruns DE Gatsonis PP, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, De Vet HC; Standards for Reporting of Diagnostic Accuracy. Toward complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for Reporting of Diagnostic Accuracy. BMJ (2003); 4; 326; 7379; 41-4

Appendix 2 STAndards for the Reporting of Diagnostic accuracy studies diagrams for the Thessaly Test, the joint line tenderness Test, McMurray's Test, Apley's Test and clinical history

Each patient who was examined had four physical tests performed and a clinical history taken. In some patients the injured knee was too painful to perform all of the tests. The STARD diagrams (*Figures 12–21*) report the number patients assessed for each test by each clinician.

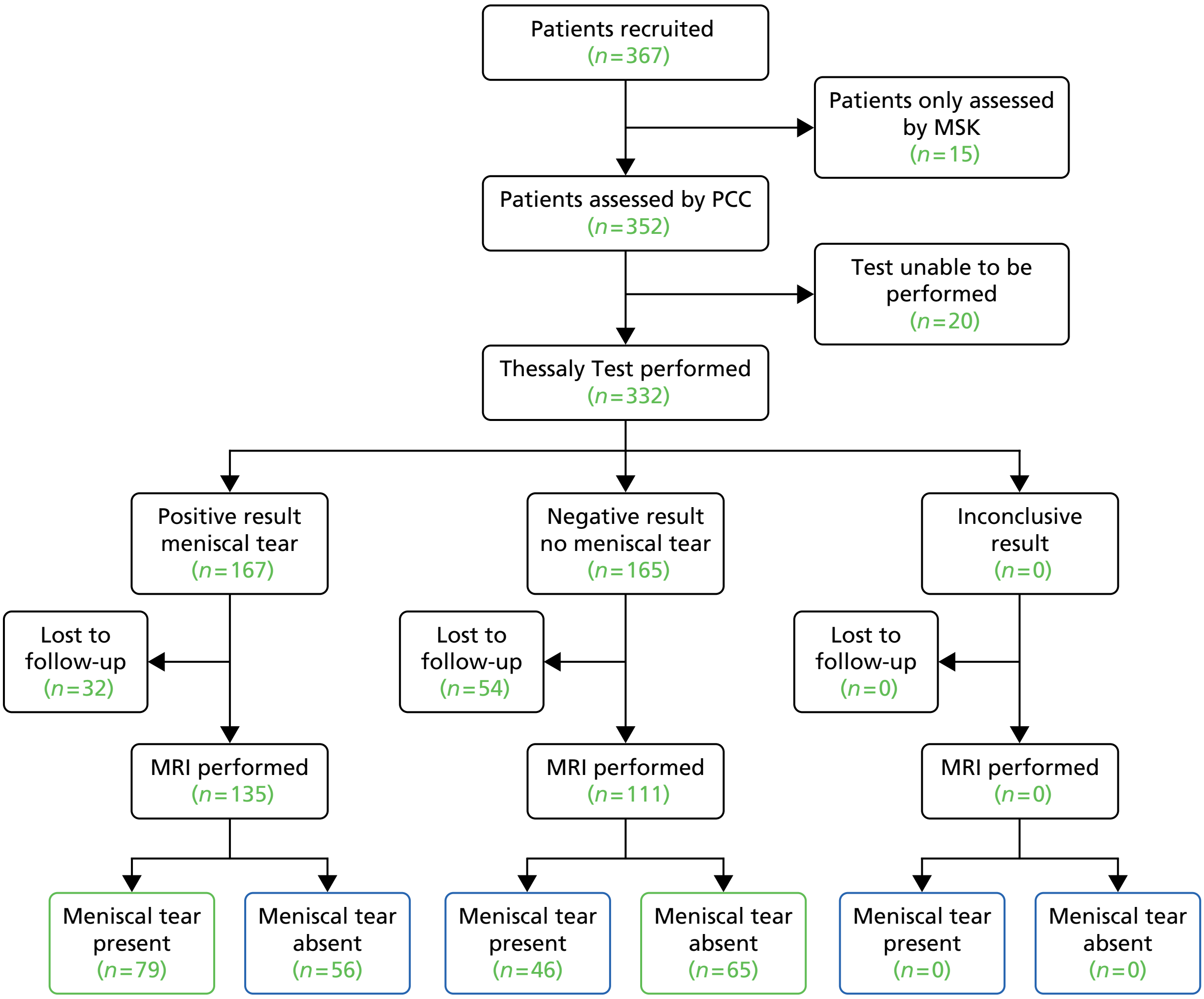


FIGURE 12 STAndards for the Reporting of Diagnostic accuracy studies diagram for the Thessaly Test: primary care clinicians. MSK, musculoskeletal clinician; PCC, primary care clinician.

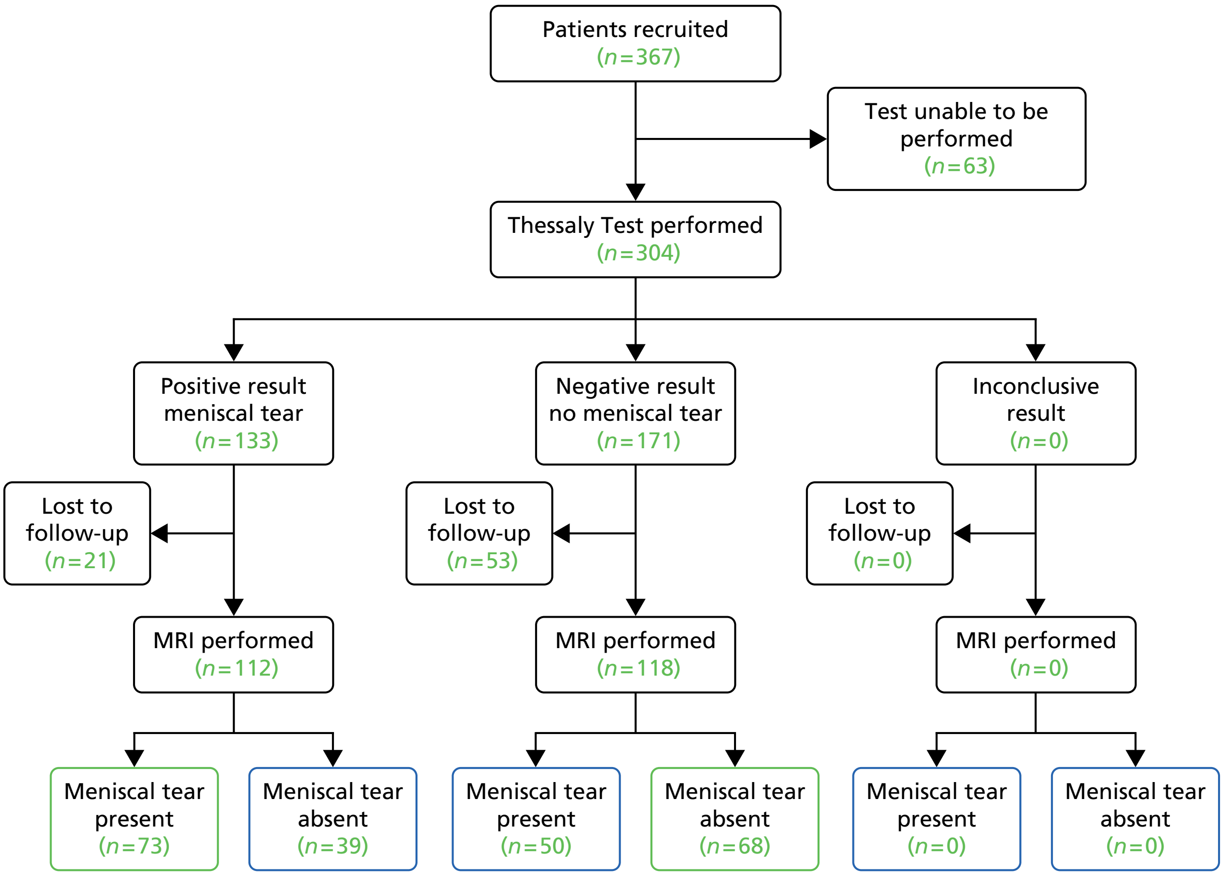


FIGURE 13 STAndards for the Reporting of Diagnostic accuracy studies diagram for the Thessaly Test: musculoskeletal clinicians.

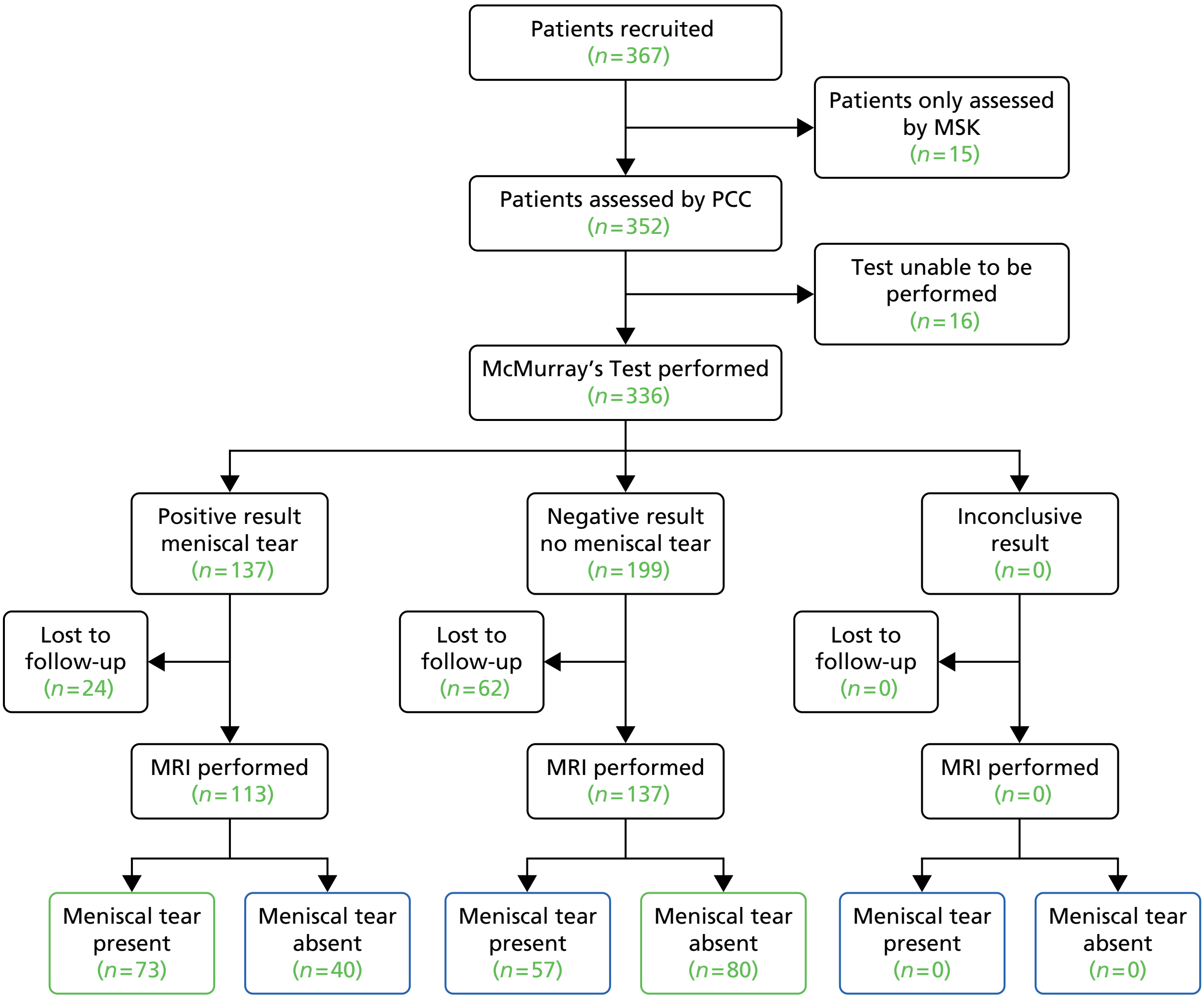


FIGURE 14 STAndards for the Reporting of Diagnostic accuracy studies diagram for McMurray's Test: primary care clinicians. MSK, musculoskeletal clinician; PCC, primary care clinician.

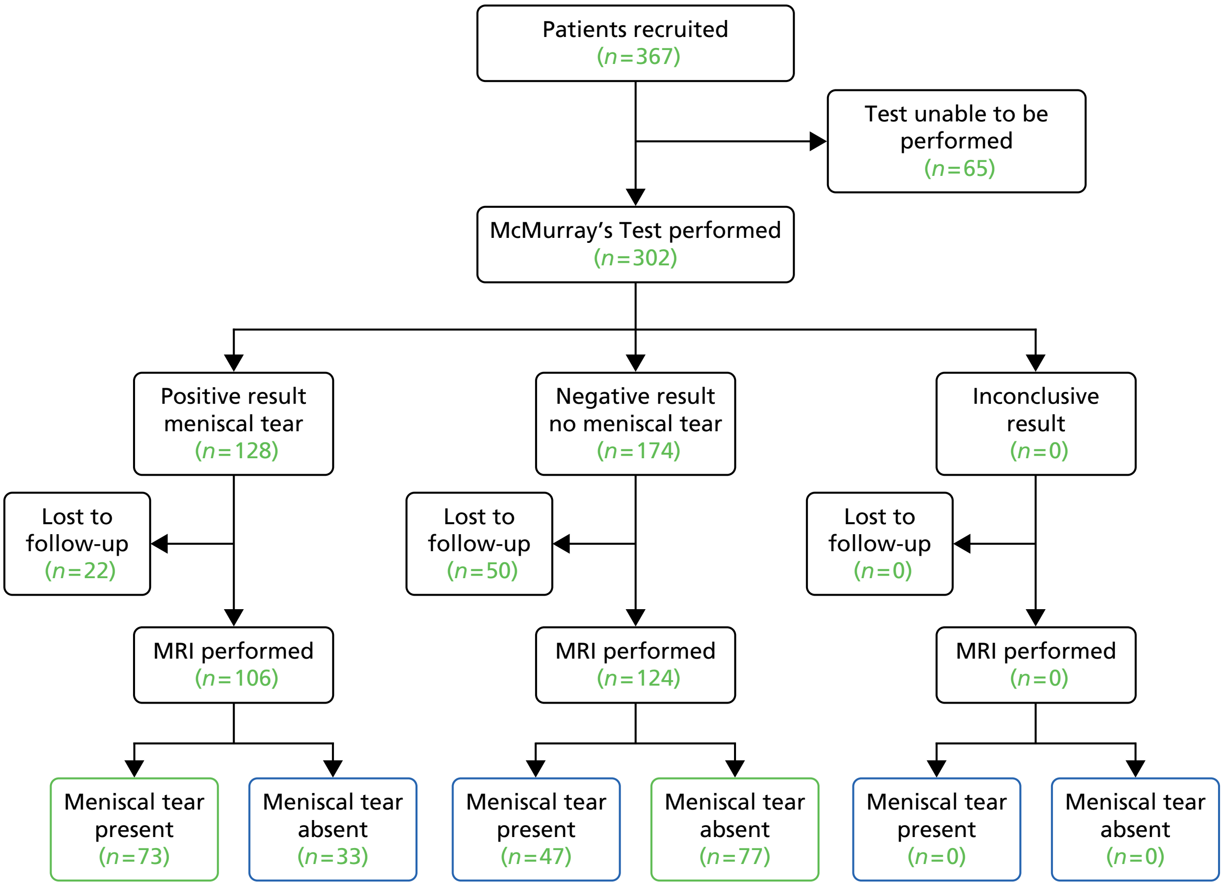


FIGURE 15 STAndards for the Reporting of Diagnostic accuracy studies diagram for McMurray's Test: musculoskeletal clinicians.

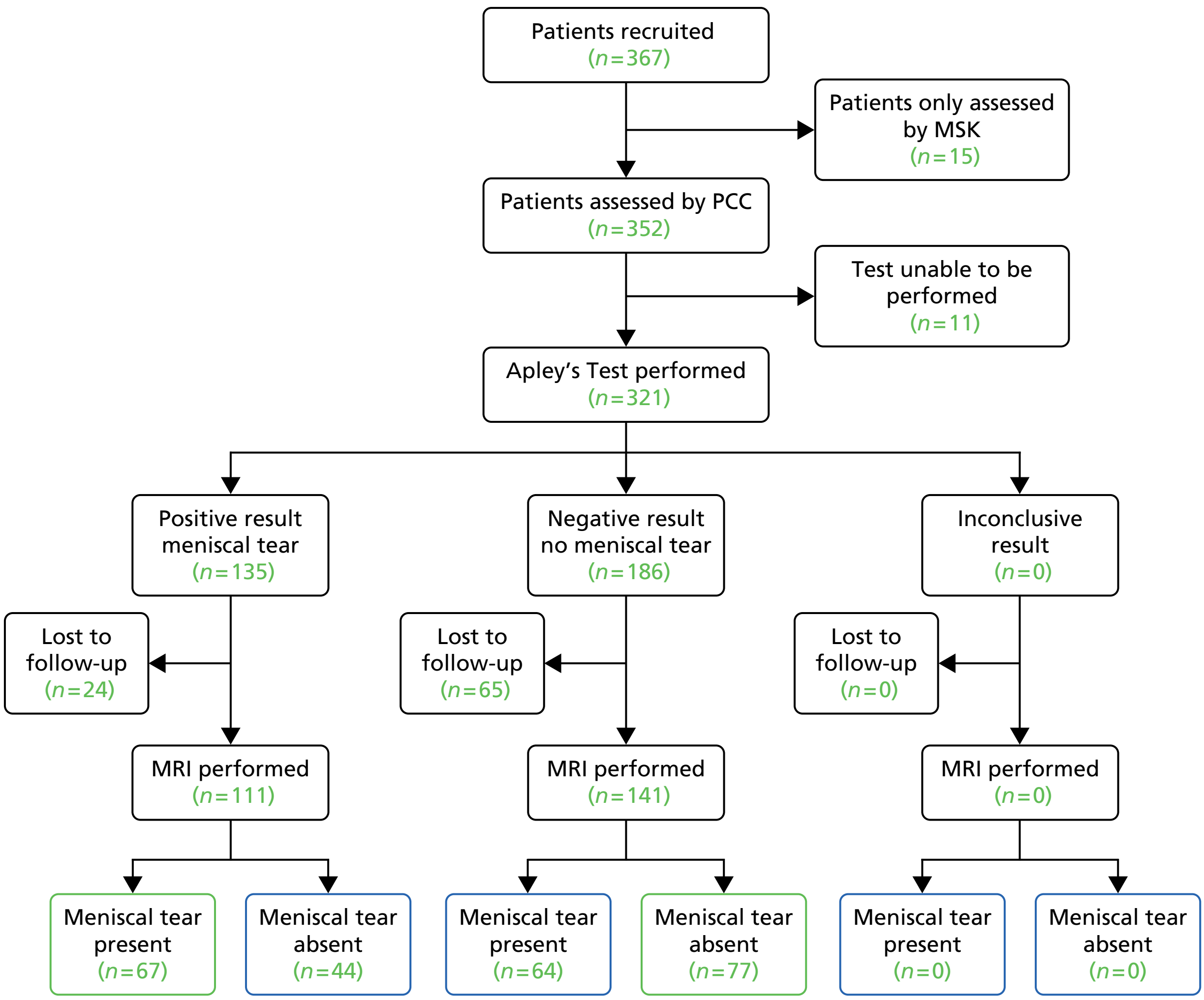


FIGURE 16 STAndards for the Reporting of Diagnostic accuracy studies diagram for Apley's Test: primary care clinicians. MSK, musculoskeletal clinician; PCC, primary care clinician.

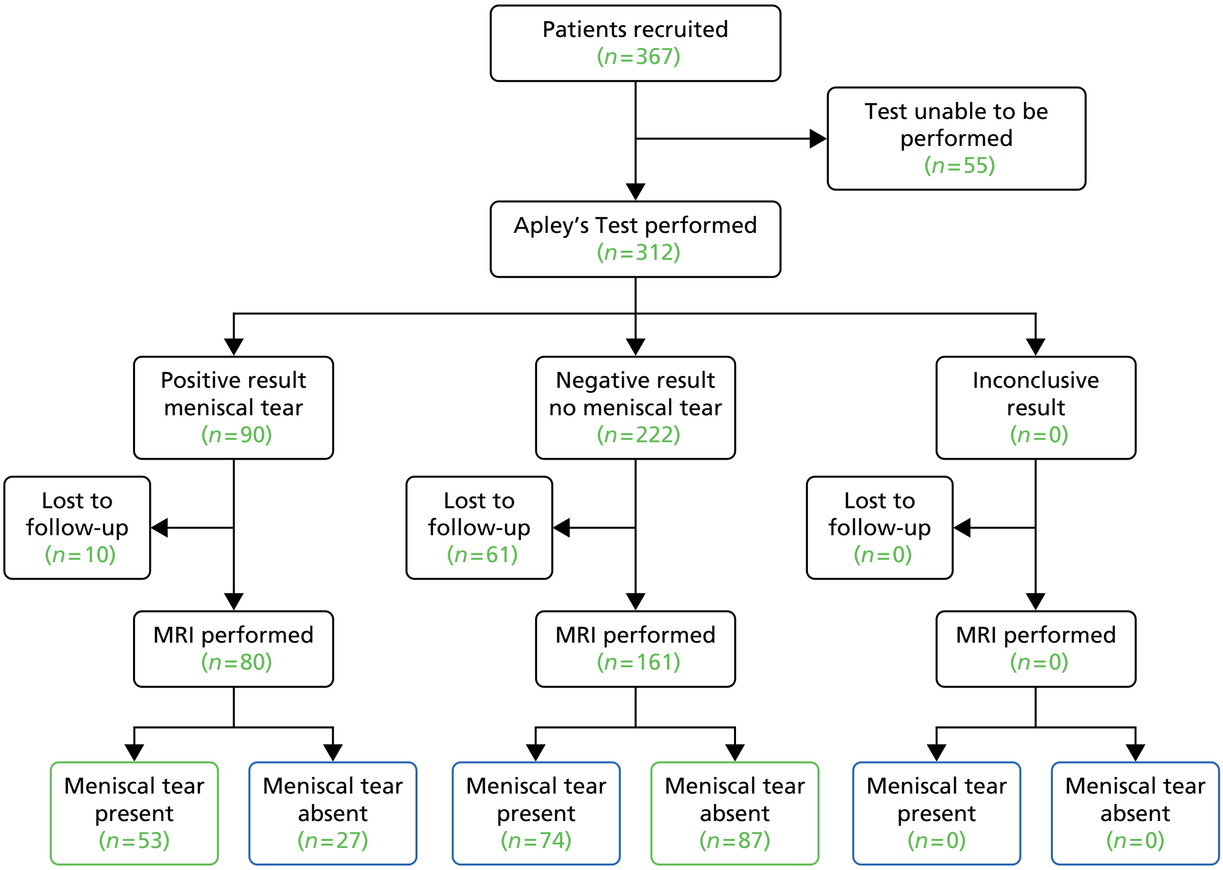


FIGURE 17 STAndards for the Reporting of Diagnostic accuracy studies diagram for Apley's Test: musculoskeletal clinicians.

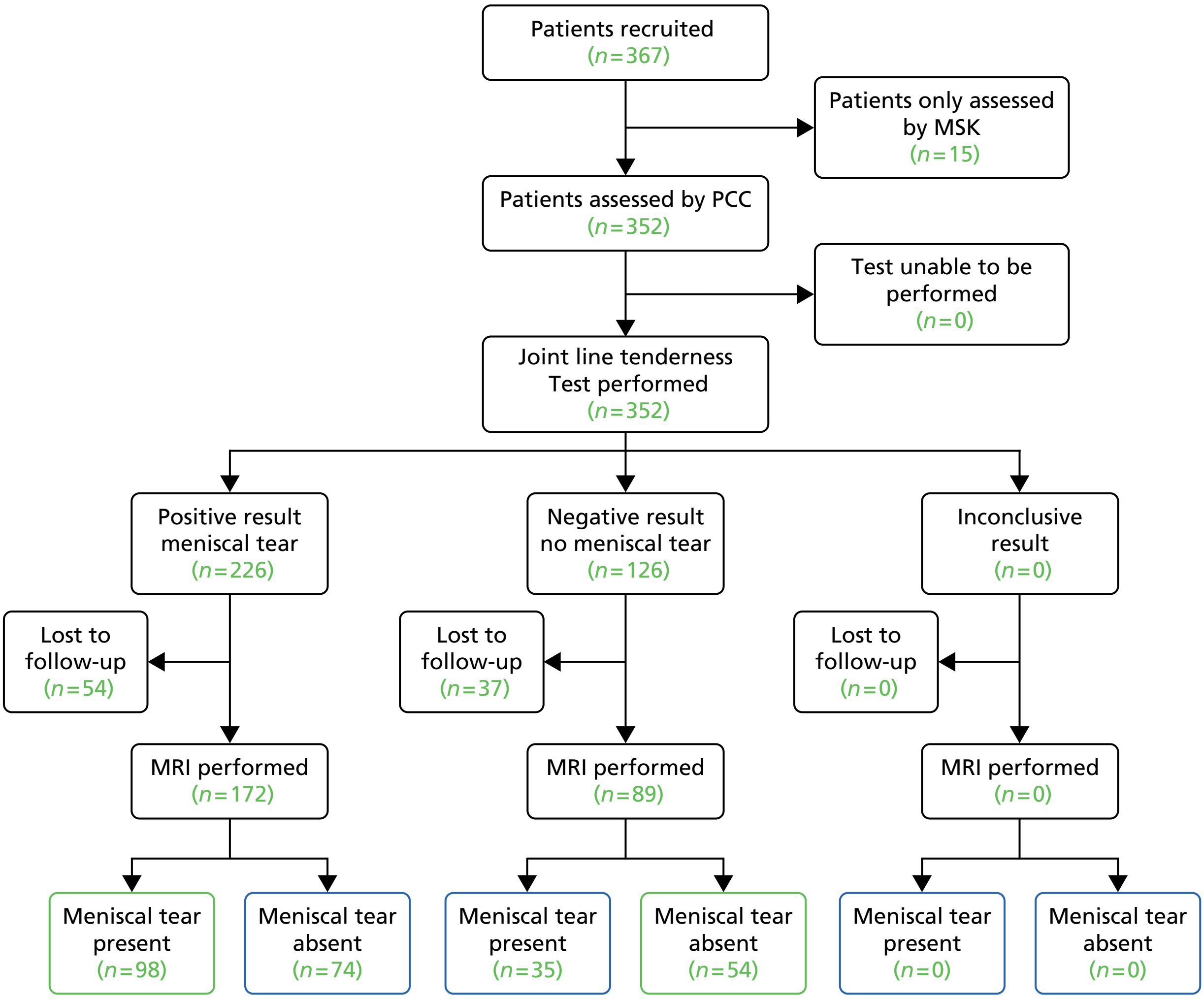


FIGURE 18 STAndards for the Reporting of Diagnostic accuracy studies diagram for the joint line tenderness Test: primary care clinicians. MSK, musculoskeletal clinician; PCC, primary care clinician.

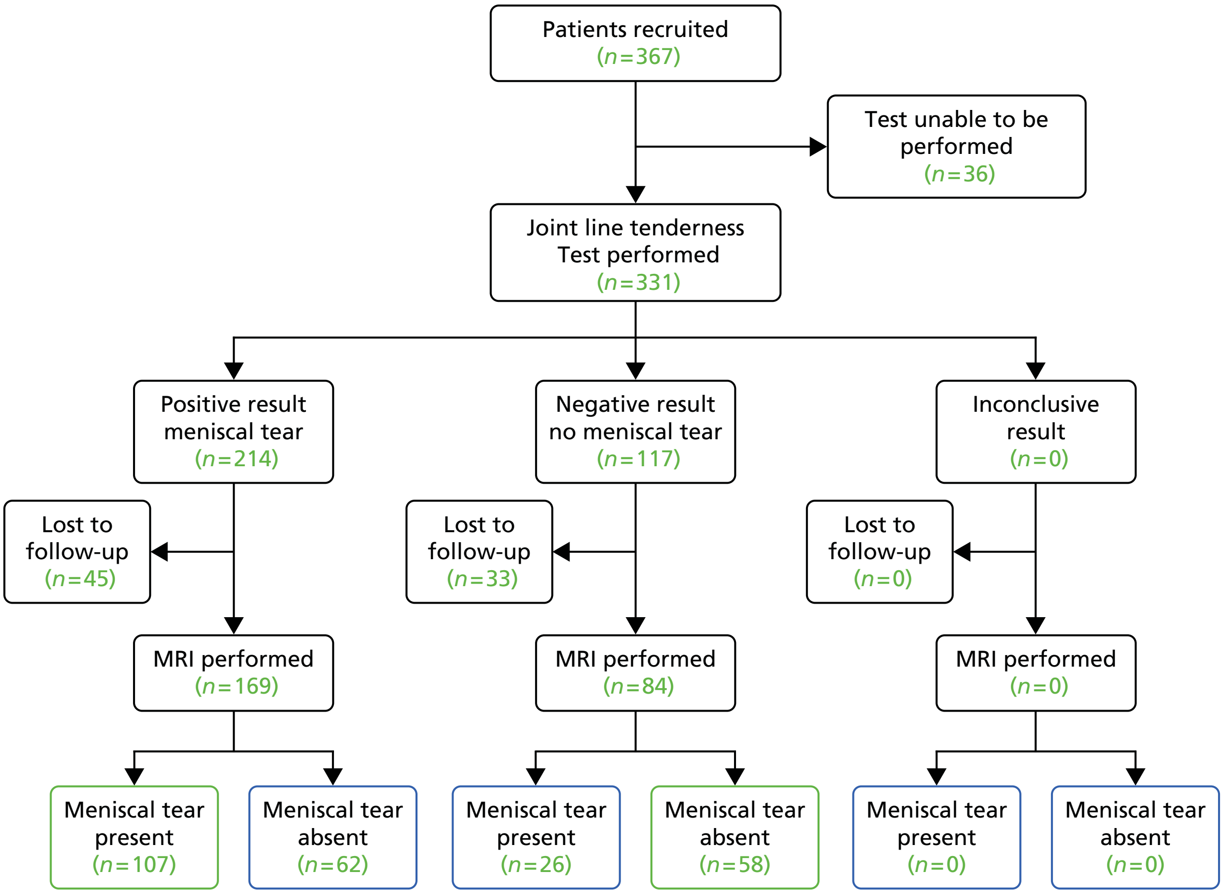


FIGURE 19 STAndards for the Reporting of Diagnostic accuracy studies diagram for the joint line tenderness Test: musculoskeletal clinicians.

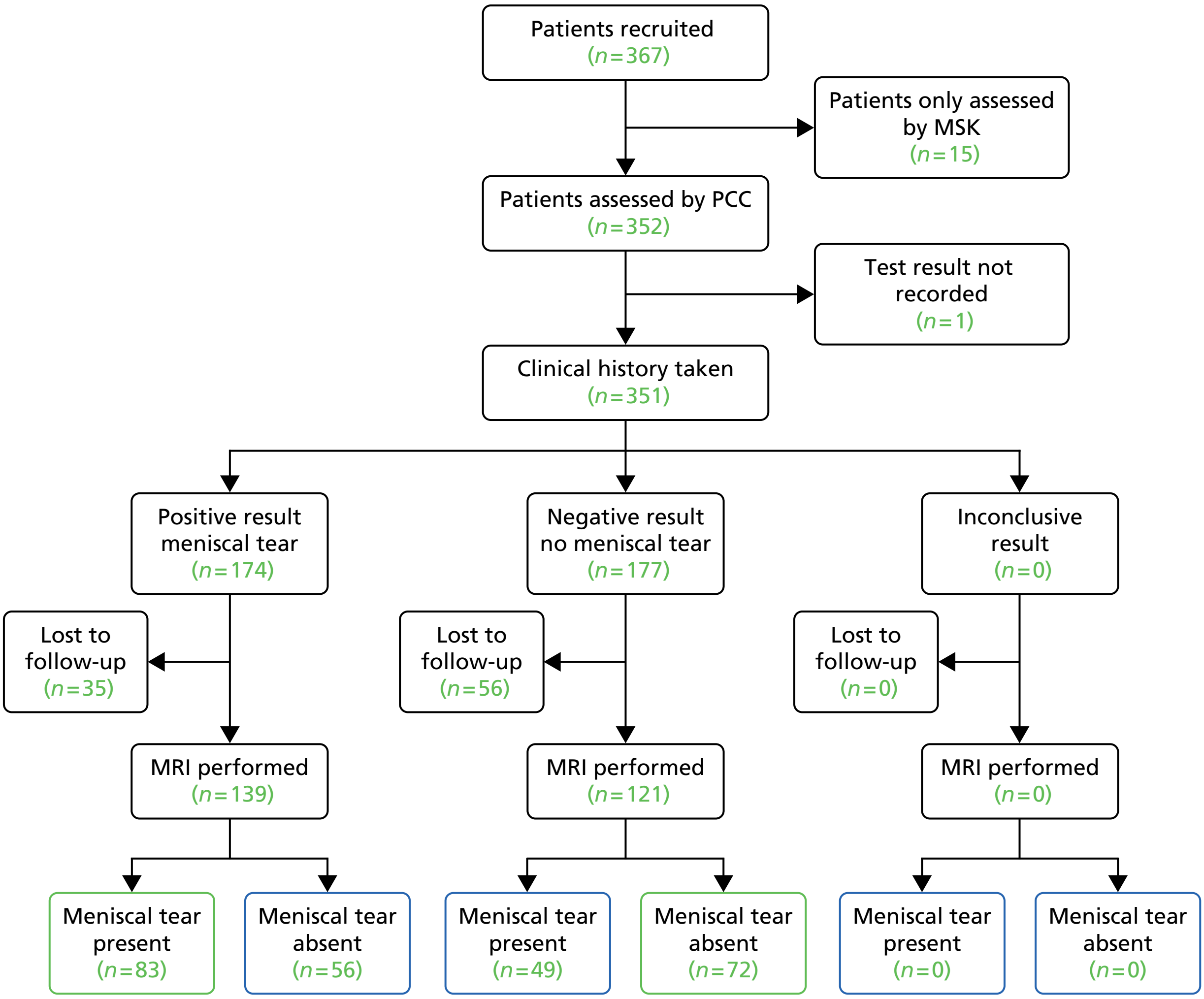


FIGURE 20 STAndards for the Reporting of Diagnostic accuracy studies diagram for clinical history: primary care clinicians. MSK, musculoskeletal clinician; PCC, primary care clinician.

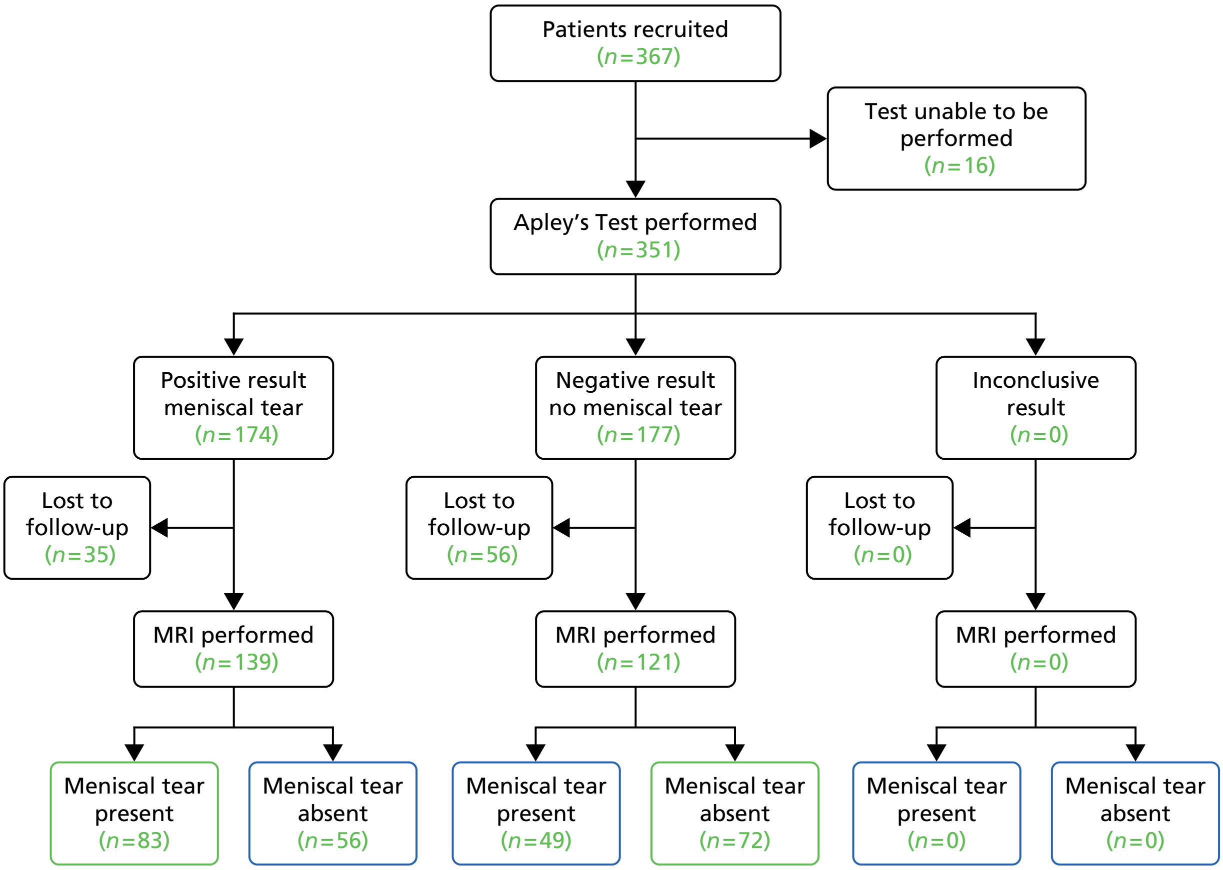


FIGURE 21 STAndards for the Reporting of Diagnostic accuracy studies diagram for clinical history: musculoskeletal clinicians.

Appendix 3 Deviations from the statistical analysis plan

Logistic regressions actually performed in this report in comparison to those specified in the statistical analysis plan:

Core model Not performed, unnecessary.

Model level 1 Five models (four physical tests + clinical history); includes effect of the test and its order.

Model level 2 One model; includes all covariates + GP as random effect.

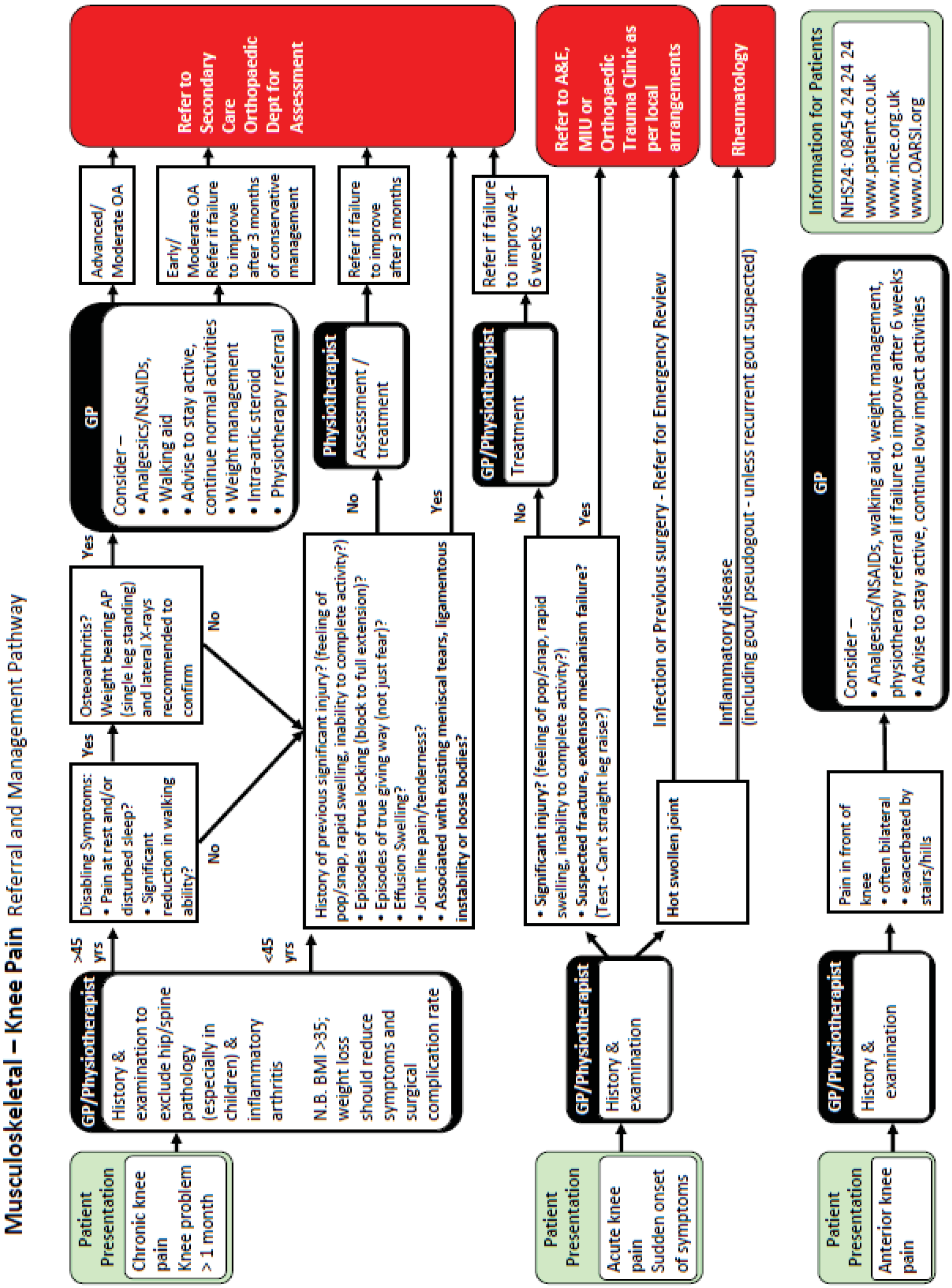
Model level 3 Five models (four physical tests + clinical history); model level 2 + effect of the test.

Model level 4 Not performed, as data not available (GP characteristics).

Model level 5 One model; model level 2 + effects of all tests.

Model level 6 One model, derived by backwards stepwise selection; starting model is model level 5 plus all interactions.

Appendix 4 Patient referral pathway for knee pain



Musculoskeletal – Knee Pain Referral and Management Pathway

These guidelines are not intended to be exhaustive and allow for application of clinical judgement in individual situations.

Referral for Knee Imaging

Plain X-Rays

Indicated in the following situations:

- There is a history of acute trauma and associated knee swelling
- Patient remains unable to bear weight on the affected limb after knee injury
- Other situations where fracture is clinically suspected
- OA is suspected (patient over 45 or clinical signs indicative) N.B. Requires weight-bearing AP (single leg standing) /Lateral X-Rays
- Loose intra-articular body suspected

MRI

MRI is the most accurate non-operative method of diagnosing meniscal and ligament pathology.
It is indicated in the following situations:

- In suspected cases of meniscus and ligament injury
- In suspected cases of patellar dislocation where extensor mechanism injury is suspected
- Where there is doubt about diagnosis

N.B. Direct access to MRI for Knees by GPs and 'straight to test' (at vetting stage, pre-outpatients appointment) options are being piloted at a number of Boards. Once results have been evaluated this pathway may be updated.

Ultrasound Scanning

Ultrasound scanning may be indicated where there is the clinical suggestion of extensor mechanism injury.

Arthroscopy

Arthroscopy is the operative technique of choice for dealing with:

- Meniscus injury, particularly where there are mechanical symptoms or pain in keeping with symptomatic meniscus pathology. N.B. Meniscus pathology can occur at any age, even in the presence of mild OA.
- Cruciate reconstruction
- Other intra-articular pathologies (e.g. loose bodies, localised persistent joint line pain despite conservative management)

Arthroscopy should only be carried out after some form of pre-operative imaging of the knee, usually MRI. Arthroscopy may be appropriate where MRI findings are equivocal or diagnosis remains in doubt after scanning e.g. suspected lateral meniscus tears with persistent symptoms.

N.B. –

- Arthroscopy is not appropriate for meniscus pathology in the presence of moderate to severe OA.
- Arthroscopy for anterior knee pain is rarely indicated.
- Bi-lateral arthroscopy is rarely indicated and would always require pre-operative MRI scanning.
- Arthroscopy should not be routinely used for diagnostic purposes where non-invasive imaging may be more appropriate.
- Evidence is clear that arthroscopy with washout/debridement is not an appropriate treatment for established OA of the knee.

N.B. There is a national project to pilot community MSK service assessment/triage of MSK conditions. If your Board is involved in this, consider whether this condition could be assessed/triaged in this way.

Patient | Primary Care | Secondary Care

Useful Information for Patients

NHS24: 08454 24 24 24
www.patient.co.uk
www.nice.org.uk
www.OARSI.org

Appendix 5 Control patient recruitment poster

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library