

Utility of clinical tests to diagnose MRI-confirmed gluteal tendinopathy in patients presenting with lateral hip pain

Alison Grimaldi,^{1,2} Rebecca Mellor,² Phillipa Nicolson,³ Paul Hodges,⁴ Kim Bennell,³ Bill Vicenzino^{2,4}

¹Physiotec, Brisbane, Queensland, Australia

²The University of Queensland, School of Health and Rehabilitation Sciences: Physiotherapy, Brisbane, Queensland, Australia

³Department of Physiotherapy, Centre for Health, Exercise and Sports Medicine, University of Melbourne, Carlton, Victoria, Australia

⁴NHMRC Centre of Clinical Research Excellence in Spinal Pain, Injury and Health, The University of Queensland, Brisbane, Queensland, Australia

Correspondence to

Professor Bill Vicenzino, School of Health and Rehabilitation Sciences: Physiotherapy, The University of Queensland, Brisbane, QLD 4072, Australia; b.vicenzino@uq.edu.au

Accepted 24 August 2016
Published Online First
15 September 2016

ABSTRACT

Purpose Gluteal tendinopathy (GT) is a source of lateral hip pain, yet common clinical diagnostic tests have limited validity. Patients with GT are often misdiagnosed, resulting in inappropriate management, including surgery. This study determined the diagnostic utility of clinical tests for GT, using MRI as the reference standard.

Methods 65 participants with lateral hip pain were examined to evaluate the ability of clinical tests to detect MRI-determined GT (an increase in intratendinous signal intensity on T2-weighted images). Palpation of the greater trochanter and several clinical pain provocation tests applying compressive and tensile loads on the gluteal tendons were investigated. MRI of the painful hip was examined by a radiologist, blind to clinical findings.

Results Pain reported within 30 s of standing on the affected limb conclusively moves a (nominal) 50% pretest probability of GT presence on MRI to a post-test probability of 98% (specificity 100%, positive likelihood ratio ~12), whereas no pain on palpation (80% sensitivity) would rule out its presence. 20 participants (31%) had GT on MRI but clinically negative (ie, not positive on palpation and another test).

Conclusions Keeping in mind that the sample size was small (ie, possibly underpowered for indices of diagnostic utility with low precision), the results of this study indicate that a patient who reports lateral hip pain within 30 s of single-leg-standing is very likely to have GT. Patients with lateral hip pain who are not palpably tender over the greater trochanter are unlikely to have MRI-detected GT.

INTRODUCTION

Lateral hip pain can present a diagnostic dilemma for healthcare professionals as soft tissue pathologies around the greater trochanter may coexist with or mimic other conditions.¹ Surgical, histological and imaging studies have demonstrated that pain and local tenderness over the greater trochanter can most commonly be attributed to tendinopathy of the gluteus medius (GMed) and/or minimus (GMin) tendons (gluteal tendinopathy (GT)), with or without bursal distension.^{2–4} Lateral hip pain can be a challenge to manage if it is misdiagnosed clinically as lumbar spine referral^{5–7} or hip joint osteoarthritis,⁸ resulting in delayed or inappropriate intervention.

A meta-analysis (14 studies on diagnostic tests for hip conditions) pooled data from 3 large studies assessing the accuracy of 4 clinical tests used to diagnose GT, and reported that no single clinical

test could convincingly predict findings from MRI.⁹ The test with most promise was the resisted external de-rotation test from a hip flexed and externally rotated position. This test is designed to stress the gluteal tendons by longitudinal (tensile) and compressive loads.¹⁰ Limitations of previous studies assessing diagnostic utility of clinical tests were identified in Reiman *et al's*⁹ systematic review. The primary issues were low participant numbers and the use of healthy controls without hip pain, who would not ordinarily present to a clinic or be considered for differential diagnosis of hip pain.⁹ The aim of this study was to determine the diagnostic utility of clinical tests in determining whether those presenting with lateral hip pain have evidence of GT on MRI.

METHODS

We conducted a diagnostic utility study of people with lateral hip pain in which practitioners who were not aware of each other's findings performed the clinical and MRI examinations.

Participants

Participants were recruited through advertisements (local print and social media) from Brisbane and Melbourne metropolitan regions. A preliminary phone screening was undertaken to confirm the presence of lateral hip pain and eligibility for the study. Participants were included if they were aged between 35 and 70 years, and had been experiencing pain in the lateral hip area, of an intensity of $\geq 4/10$ on an 11-point Numeric Rating Scale on most days within at least the past 3 months. Potential participants were excluded if they had received a corticosteroid injection in the hip region in the past 12 months, had any hip or lower back surgery, had any known neurological disorders, any other relevant medical conditions (eg, fibromyalgia) or any contraindications to MRI (eg, cardiac pacemaker, metal implants). Eligible participants were then booked in for clinical examination at a university clinical research laboratory and MRI examination at a radiology practice within 2 weeks of each other. The Institutional Medical Research Ethics Committee approved the study and all participants gave written informed consent.

Clinical examination

A physiotherapist performed a series of standardised tests during a clinical examination for GT. Diagnostic clinical tests were employed to apply loads/stresses that are pathognomic of GT. Tensile



To cite: Grimaldi A, Mellor R, Nicolson P, *et al.* *Br J Sports Med* 2017;**51**:519–524.

(longitudinal) and compressive (transverse) loads have been implicated in the development of tendinopathy,^{10–13} which in the gluteal tendons involves positions of hip adduction and contraction of the GMed and GMin muscles.^{10–14} The following six clinical tests employing either or both tensile and compressive loading were performed sequentially as described below, and rated positive if pain was reproduced in the region of the greater trochanter to a score of >2 on an 11-point Numeric Rating Scale (with 0=no pain, 10=most pain imaginable):

1. Hip Flexion, Adduction, External Rotation (FADER): with the participant lying supine, the hip was passively flexed to 90°, adducted and externally rotated to end of range (figure 1). This test seeks to increase both tensile and compressive load on the GMed and GMin tendons at the greater trochanter through positioning of the hip and the overlying iliotibial band.
2. Hip FADER with resisted isometric internal rotation at end of range (FADER-R): in the FADER position, the participant performed resisted isometric internal rotation against a force applied by the physiotherapist (figure 1). The test is designed to increase tensile and compressive loading of the gluteal tendons through active contraction of the GMed and GMin muscles, which in this position are internal rotators.¹⁵ This is a modification of the resisted external de-rotation test described by Lequesne *et al*¹⁶ and reported by Reiman *et al*⁹ to have pooled sensitivity of 71% (95% CI 51% to 87%) and specificity of 84% (71% to 93; 2 studies; n=79). Hip adduction was added to the flexion/external rotation position to add further tensile loading on the tendons, as well as increasing the compressive loading element exerted by the overlying iliotibial band at the greater trochanter.
3. Hip Flexion, Abduction, External Rotation (FABER): the lateral malleolus of the test leg was placed above the patella of the contralateral leg, the pelvis stabilised via the opposite anterior superior iliac spine and the knee passively lowered so the hip moves into abduction and external rotation. This test is expected to place the anterior portions of the GMed and GMin on tensile load due to the fact that these portions of the muscles have an internal rotation function.¹⁵ Lateral hip pain with a FABER test has been shown to have high sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV; 82.9%, 90%, 94.4% and 72%, respectively) for differentiating between greater trochanteric pain syndrome and hip osteoarthritis.¹⁷

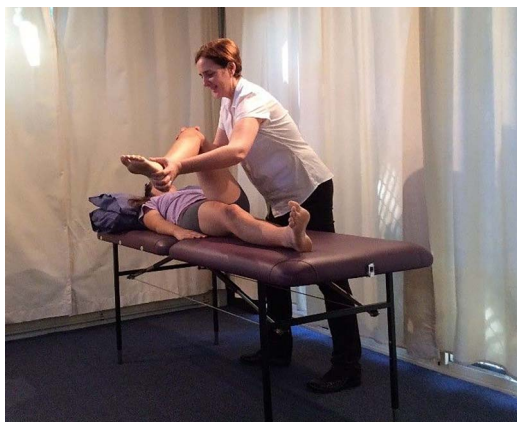


Figure 1 FADER and FADER-R test position. FADER, Flexion, Adduction, External Rotation; FADER-R, Flexion, Adduction, External Rotation with resisted isometric hip internal rotation.



Figure 2 ADD and ADD-R test position. ADD, adduction; ADD-R, adduction with resisted isometric hip abduction.

4. Passive Hip Adduction in Side Lying (ADD): the participant was placed diagonally across the bed in side-lying, with the underneath hip and knee flexed 80–90°, and the uppermost leg supported by the examiner with the knee extended, in neutral rotation, and the femur in line with the trunk (0° hip extension). The anterior superior iliac spines were maintained perpendicular to the examination table. The examiner passively moved the hip through a pure frontal plane motion into end range hip adduction with overpressure, while stabilising the pelvis with the other hand (figure 2). This test places the lateral insertions of the gluteal tendons under tensile and compressive load.
5. ADD with resisted isometric abduction (ADD-R): in the ADD position, the participant was asked to push the thigh up, against the resistance of the examiner's hand at the lateral knee. This adds an active tension component to the passive tensile and compressive loads imposed on the tendons of GMed and GMin at end range hip adduction.
6. Single Leg Stance (SLS): the participant stood side-on to a wall with the affected limb furthest from the wall. A finger of the unaffected side touched the wall at shoulder height for balance. The foot nearest the wall was then raised so that the hip remained in neutral with the knee flexed to 90°. This SLS position was then maintained for up to 30 s. A positive test is reproduction of the patient's lateral hip pain within this 30 s period. The SLS test has been shown to discriminate participants with MRI-diagnosed GT/trochanteric bursitis from those with normal non-symptomatic hips (100% sensitivity and 97.3% specificity).¹⁶
7. Palpation: the participant was positioned in side-lying with the symptomatic side uppermost, hips flexed ~60° and knees together. The lateral hip region was systematically palpated for tenderness as per Falvey *et al*.¹ A positive test was pain with palpation of the GMed and/or GMin tendon insertions (anterior, lateral or posterosuperior facets of the greater trochanter).

MRI examination

Participants underwent an MRI examination of their gluteal tendons and hip at a participating radiology clinic. The MRI images were acquired on a MAGNETOM Espree 1.5 T scanner (Siemens AG), and the imaging protocol included the following sequences: axial PD fat sat (TR 3130, TE 23, 23×3.5 mm slices, Base res 384); coronal PD fat sat (TR 3050, TE 42,

19×3.5 mm slices, Base res 384); coronal T1 (TR 459, TE 12, 19×3.5 mm slices, Base res 384); sagittal PD fat sat (TR 3560, TE 39, 23×3.5 mm slices, Base res 384); sagittal PD (TR 2470, TE 31, 23×3.5 mm slices, Base res 384), all with an 0.7 mm×0.5 mm matrix.

The MRIs were examined by an experienced radiologist who was blind to the clinical examination findings. The radiologist looked for the presence of GT, bursitis, osteoarthritis, labral pathology or any other observed pathology. In particular, each participant's images were evaluated for any area of T2 hyperintensity (representing oedema or fluid) around the greater trochanter. If T2 hyperintensity was present, the size, shape and location of the hyperintensity was described. Intratendinous T2 hyperintensity was required to fulfil the diagnostic criteria for GT. Partial or full thickness tears of the GMed and GMin tendons were also identified when present. A partial thickness tear was diagnosed if the tendon was irregular, thinned or focally discontinuous on T1-weighted images, with hyperintensity in the corresponding area on T2-weighted images. A complete tear was diagnosed when discontinuity and/or retraction of the torn tendon was seen on T1-weighted images, with a marked increase in signal on T2-weighted images.

Data analysis

A positive clinical diagnosis of GT from the clinical examination was defined as positive to palpation over the greater trochanter, plus at least one other positive result from the remaining clinical tests described above. From the MRI reports, the presence of GMed or GMin tendinopathy was defined as an intratendinous increase in signal intensity on T2-weighted images, based on the classification system from a previous study.¹⁸

Statistical analysis was performed using IBM SPSS Statistics, V.22 (SPSS, Chicago, Illinois, USA). In order to be able to determine the diagnostic utility of the clinical tests (defined as the accuracy with which GT can be detected on clinical examination), we compared the physical examination results with the MRI diagnosis. A series of 2×2 contingency tables were generated, using the MRI results (positive or negative for GT) as the reference standard. Sensitivity (the proportion of true positives that are correctly identified by a diagnostic test, ie, how good the test is at detecting the condition), specificity (the proportion of true negatives correctly identified by a diagnostic test, ie, how good the test is at ruling out a condition that is not present), PPV (probability that the disease is present when the test is positive), NPV (probability that the disease is not present when the test is negative) and their 95% CIs were calculated for each of the clinical tests as well as for the positive clinical diagnosis (ie, palpation plus at least one other positive test). Accuracy (the proportion of true results either positive or negative) was also calculated. While these are indicators of diagnostic utility, they do not provide an indication of the shift in probability of the condition being present should the test be positive or negative.¹⁹ Likelihood ratios (LRs) combine sensitivity and specificity to provide such an indicator (point estimate of effect). LR can be used to calculate post-test probabilities of a condition being present (ie, the probability that the condition is present after the test is applied), based on pretest probabilities (the probability that the condition is present based on information the clinician has at hand before applying the test). A clinical example will be used to illustrate the clinical application of findings from this study. Since LR values can range from 0 to infinity, almost conclusive shifts in post-test probability of the condition being present or not are likely to occur for positive LR (LR+) ≥10 or

negative LR (LR-) ≤0.1, respectively.²⁰ Statistical significance (p<0.05) will be achieved when 95% CIs do not contain 1.

RESULTS

Participants

Sixty-five participants (45 (69%) females) with a mean age of 54.9±9 years were recruited for this study (table 1). Figure 3 outlines the flow of participants throughout the study. All participants completed both clinical and MRI assessments without adverse effect.

Diagnostic accuracy

Thirty-four participants (52%) had a positive clinical diagnosis (ie, positive palpation plus ≥1 other clinical test), whereas 50 (77%) of the MRI reports identified the presence of GT. Of the clinical diagnoses, there were 30 (46%) true positives, 11 (16.9%) true negatives, 4 (6%) false positives and 20 (30.7%) false negatives (table 2).²¹ The clinical examination had moderate sensitivity and specificity (60% and 73%, respectively), a high PPV (88%) and a reasonable accuracy rate of 63%.

Palpation over the greater trochanter was the primary clinical diagnostic criterion and it had the highest sensitivity (80%) and accuracy rate (72%), but the lowest specificity (47%). Of the 65 participants, there were 40 true-positive palpation tests (positive on palpation and GT present on MRI) and only 8 false-positive diagnoses (ie, 8 tested positive on palpation, yet GT was not present on MRI). Palpation had the lowest negative LR (0.43), indicating that if palpation was negative on clinical testing, it was the single best of all the individual tests used in this study for ruling out the presence of GT.

In contrast to palpation, each of the other individual clinical examination tests showed high specificity (SLS 100% to FABER 80%). The SLS test had a specificity of 100%, sensitivity of 38% and PPV of 100%, indicating that if the test was positive there was an extremely high likelihood that GT was present on MRI. That is, all participants who tested positive on SLS had GT on MRI, but there were 31 participants who were negative on the test who had GT identified on MRI.

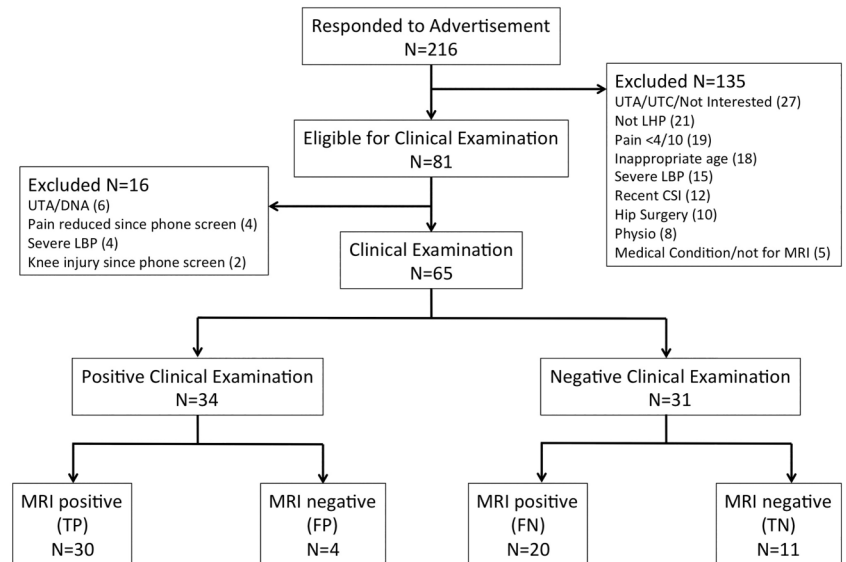
The FADER and ADD tests both had high specificity (86.67%) and high PPV (88.24% and 83.33%, respectively). Of particular interest is that when resisted muscle contraction is added to each of these tests (ie, FADER-R and ADD-R), the sensitivity, specificity, PPV and NPV of both of these tests improved. The positive LR of the FADER test (2.25) increased to 6.6 for the FADER-R test, and the positive LR of the ADD test (1.5) increased to 5.7 for the ADD-R test. This tends to suggest that if manual resistance is added to these tests, which effectively adds an active tension component (muscle tension) to the passive tension and compressive component of each test (positioning of the tendon), then diagnostic accuracy is increased substantially.

Table 1 Participant demographics (mean±SD)

	All participants (N=65)	Positive clinical (N=34)	Negative clinical (N=31)
Sex (%)	45 female (69)	28 female (82)	17 female (55)
Age (years)	54.9±9	55.88±8.66	53.9±9.9
Height (m)	1.68±0.07	1.66±0.06	1.7±0.08
Weight (kg)	80.82±18.8	79.69±16.57	82.09±21.25
BMI (kg/m ²)	28.36±5.54	28.55±5.08	28.13±6.1

BMI, body mass index.

Figure 3 Participant flow chart. CSI, corticosteroid injection; DNA, did not attend; FP, false positive; FN, false negative; LBP, low back pain; LHP, lateral hip pain; TN, true negative; TP, true positive; UTA, unable to attend; UTC, unable to contact.



All clinical tests with the exception of palpation had a high number of false-negative tests (28 (43%) to 40 (62%)), which far outnumbers any of the other decisions (0–34% for false and true positive and true negative). That is, the MRI has revealed the presence of GT on imaging, yet the clinical test did not detect this.

DISCUSSION

This study provides guidance for the clinician regarding which tests might be most useful in diagnosing GT as observed on MRI in a population presenting with lateral hip pain. The clinical tests chosen for this study were ones that have been considered to be positive for GT, so it is not that surprising that the PPV were high and the NPV low (table 2). The data indicate that combining palpation with its high sensitivity and accuracy with any of the other individual clinical examination tests, which have higher specificity and PPVs, might be best to identify those patients who have GT on MRI. Of all the physical tests, SLS when positive (ie, pain within 30 s of standing on one limb) is the one that significantly increases the probability that there is GT on MRI. As is the case with many other conditions,²² MRI detects substantially more abnormalities (in this case tendinopathy) than is identifiable on clinical examination. To make a diagnosis of symptomatic GT, clinicians should rely on MRI, as well as include palpation of the greater trochanter and at least one positive physical test of the clinical examination studied here.

Our data infer that if palpation is positive, there is a high likelihood of MRI-identified GT (ie, PPV of 83%), and if it is negative the patient is unlikely to have a positive MRI because negative results on tests with high sensitivity (ie, 80% sensitivity for palpation) are useful in ruling out pathology. This was supported by an LR⁻ of 0.43, which signifies that the odds of GT being present are more than halved when palpation is pain-free. While palpation had the highest combined true positives and negatives (72%), it had the highest number of positive results when there was no GT on MRI and the lowest specificity (46.67%). Even in Lequesne *et al*'s¹⁶ pain-free control group, specificity of palpation was only 66%, reflecting a significant number of asymptomatic yet tender greater trochanters. This reinforces that pairing palpation with one of the other tests with greater specificity will have greater utility in identifying those with GT on MRI compared with palpation alone.

Of all the physical tests studied, those that involve an active muscle contraction component, such as FADER-R, ADD-R and SLS, appear more useful than FADER, FABER and ADD for identifying GT on MRI. The FADER-R and ADD-R tests use active muscle contraction to intensify the tensile loading component from positions in which the gluteal tendons have been exposed to maximal compression and passive tension. Comparisons with previous studies find our results in agreement with Lequesne *et al*¹⁶ that when positive, both SLS and resisted internal rotation in a position of relative tendon compression (resisted external de-rotation test and FADER-R) are useful tests in the diagnosis of GT. Specificity of these tests was high for both studies, but the true negatives returned by the current study were gathered from a population with lateral hip pain rather than the completely pain-free controls of the previous study. This new finding provides greater confidence in the specificity of such tests. Regarding sensitivity, our findings are in contrast with those of Lequesne *et al*¹⁶ who reported high sensitivity for SLS (100%) and the resisted external de-rotation test (88%). This might have been due to the use of a no pain control group being used as the comparison and the GT group being included on the basis of a positive palpation of the greater trochanter and hip external rotation at 90° of hip flexion.

Trendelenburg-like SLS assessments that test the ability of the hip abductor muscles to control pelvic posture rather than testing pain provocation have been shown to be specific but not sensitive to MRI-detected GT²³ and reasonably sensitive and specific for detection of gluteal tendon tears (72.7%, 76.9%, respectively).⁴ We found SLS to have high specificity but low sensitivity, possibly indicating that observation of postural control of the pelvis during SLS is not required in detecting GT on MRI. In contrast, identification of gluteal tendon tears on MRI might require both pain reproduction and observation of postural control on SLS.

Stronger conclusions may have been possible with higher participant numbers, even though with 65 participants this is the largest study of its kind until now. Increasing sample size in future studies might reduce the CIs around the indices of diagnostic utility, thereby providing greater confidence in conclusions from this study. The other limitation within the study design was the limitation inherent in using MRI as a reference standard, as MRI findings tend to be poorly correlated with

Table 2 Clinical test results and indices of diagnostic utility in the diagnosis of gluteal tendinopathy using MRI as the reference standard

Test	True +	True -	False +	False -	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	LR+(95% CI)	LR-(95% CI)	Accuracy (%)
Clinical diagnosis of GT	30	11	4	20	60.0 (30/50) (45.2 to 3.6)	73.3 (11/15) (44.9 to 92.2)	88.2 (30/34) (72.5 to 96.7)	35.0 (11/31) (19.2 to 54.6)	2.25 (0.94 to 5.37)	0.55 (0.35 to 0.86)†	63
PALP	40	7	8	10	80.0 (40/50) (66.3 to 89.9)	46.7 (7/15) (21.3 to 73.4)	83.3 (40/48) (69.8 to 92.5)	41.2 (7/17) (18.4 to 67.1)	1.5 (0.92 to 2.46)	0.43 (0.20 to 0.93)†	72
SLS	19	15	0	31	38.0 (19/50) (24.6 to 52.8)	100 (15/15)	100 (19/19) (78.2 to 100)	32.6 (15/46) (19.5 to 48.0)	12.2 (0.8 to 191.5)*	0.62 (0.50 to 0.77)†	52
FADER	15	13	2	35	30.0 (15/50) (17.9 to 44.6)	86.7 (13/15) (59.5 to 98.3)	88.2 (15/17) (63.6 to 98.5)	27.1 (13/48) (15.3 to 41.8)	2.25 (0.58 to 8.75)	0.80 (0.62 to 1.06)	43
FADER-R	22	14	1	28	44.0 (22/50) (29.9 to 58.7)	93.3 (14/15) (68.1 to 99.8)	95.6 (22/23) (78.0 to 99.9)	33.3 (14/42) (19.6 to 49.5)	6.6 (0.97 to 44.98)	0.60 (0.45 to 0.79)†	55
FABER	21	12	3	29	42.0 (21/50) (28.2 to 56.8)	80.0 (12/15) (51.9 to 95.7)	87.5 (21/24) (67.6 to 97.3)	29.3 (12/41) (16.1 to 45.5)	2.10 (0.73 to 6.08)	0.73 (0.52 to 1.02)	51
ADD	10	13	2	40	20.0 (10/50) (10.0 to 33.7)	86.7 (13/15) (59.5 to 98.3)	83.3 (10/12) (51.6 to 97.9)	24.5 (13/53) (13.8 to 31.3)	1.5 (0.37 to 6.11)	0.92 (0.72 to 1.18)	35
ADD-R	19	14	1	31	38.0 (19/50) (24.6 to 52.8)	93.3 (14/15) (68.1 to 99.8)	95.0 (19/20) (75.1 to 99.9)	31.1 (14/45) (18.2 to 46.6)	5.7 (0.83 to 39.13)	0.66 (0.51 to 0.86)†	51

Positive clinical diagnosis of GT=positive on palpation over greater trochanter, and positive on at least one other clinical test.

* Calculation of LR+ for SLS required the addition of 0.5 to each cell,²¹ because all participants who tested positive on this test had positive MRI findings, that is, 100% specificity.

† Significant negative LRs (ie, CI does not contain 1).

‡ Significant negative LRs (ie, CI does not contain 1). ADD, adduction; ADD-R, adduction with resisted isometric hip abduction; FABER, Flexion, Abduction, External Rotation; FADER, Flexion, Adduction, External Rotation; FADER-R, Flexion, Adduction, External Rotation with resisted isometric hip internal rotation; False -, negative clinical result and positive MRI; False +, positive clinical result and negative MRI; GT, gluteal tendinopathy; LR-, negative likelihood ratio; LR+, positive likelihood ratio; NPV, negative predictive value; PALP, palpation; PPV, positive predictive value; SLS, Single Leg Stance; True -, negative clinical result and negative MRI; True +, positive clinical result and positive MRI.

symptomatic presentation.²² Blankenbaker *et al*¹⁸ reported that all patients in their cohort with trochanteric pain and 88% of patients without trochanteric pain had peritrochanteric abnormalities on MRI. Woodley *et al*²³ also discussed the limitations of MRI as a diagnostic tool for GT. The fact that the presence of GT on MRI does not always correlate with the presence of significant pain and dysfunction is also supported somewhat by the large proportion of false-negative findings in this study, both for the clinical examination as well as each individual test. With respect to clinical implications, the low sensitivity and accuracy of many of our clinical tests reflects that MRI detects the presence of pathology within the gluteal tendons but does not determine well whether this pathology is a source of pain/symptoms. Forming a diagnosis of GT would optimally require considering both clinical examination and imaging findings.

Clinical implications

To improve the relevancy of our data to clinical practice, we present an example with reference to [figure 4](#).²¹ Envisage a middle-aged overweight female who presents with lateral hip pain. The possible source of this lateral hip pain could be from GT, but it could also be due to hip joint, pelvic or lumbar spine pathology.¹ Let us assume that the probability of this lateral hip pain being associated with GT is 50%. If the clinician applied the tests studied here, when positive some might provide an increased probability of GT being present while when negative they would indicate a lower probability of GT. For example, [figure 4](#) shows that FADER-R, when positive, increases the probability of the patient having MRI identified GT by ~37%, indicating that this test is helpful in confirming a diagnosis of GT. In contrast, when the test is negative, it does not rule out the likelihood of GT being present as the shift in probability is much lower at 12%. So a test might well be only good for helping to confirm a provisional diagnosis and not to refute it. This probably underpins the tendency of clinicians to use several tests, both those which confirm or refute a diagnosis, in coming to a more definitive working diagnosis. In this regard, envisage that if FADER-R was not positive; then palpation might be helpful for the clinician to rule out the presence of GT, because when palpation is negative it reduces the likelihood of GT by 20%.

An important consideration for the clinician in applying the LR findings from this study in practice is the notion that as further information becomes available during the clinical examination, the probability of a diagnosis being present or not varies. In the preceding example, we assumed that the (pretest) probability that the clinician assumes/determines on current knowledge of the patient before applying the test was 50%. A salient matter to understand is that the amount to which the probability changes when applying an LR (either positive or negative) is dependent on the assumed/predicted pretest probability. For example, if the pretest probability of MRI identified tendinopathy was 30% rather than 50%, then a positive FADER-R would shift the post-test probability by 44% (30% to 74%, compared with 50% to 87%). If the pretest probability was higher at 75%, then the shift in probability would be less (20%), but by that stage of the clinical examination there would be a 95% probability that MRI identified tendinopathy was present. This continuous nature of LR and their implication on shifting probabilities in a condition being present or not can be graphically illustrated on a nomogram ([figure 5](#)) which could also be used in clinic.²⁴

Somewhat surprisingly we found that the combination of palpation and any one of the other tests, defined as the clinical

Test	Clinical Diagnosis*		Palpation		FADER-R#		SLS^	
	+ve	-ve	+ve	-ve	+ve	-ve	+ve	-ve
Pre-test probability	50%	50%	50%	50%	50%	50%	50%	50%
	↓	↓	↓	↓	↓	↓	↓	↓
Likelihood Ratio +/-	2.25	0.55	1.5	0.43	6.6	0.6	12.2	0.62
	↓	↓	↓	↓	↓	↓	↓	↓
Post-test probability~	69%	35%	60%	30%	87%	38%	92%	38%
	↓	↓	↓	↓	↓	↓	↓	↓
Interpretation and implications of the LR	The clinical diagnosis test is slightly better at ruling out the condition (15%) than it is at ruling it in (19%).		A negative palpation test would indicate that there is a low probability of GT on MRI (by 20%), whereas a positive palpation test does not improve the likelihood of GT being present on MRI.		A positive FADER-R improves confidence by approximately 40% that the lateral hip pain is present with MRI identified GT. A negative test is not so useful at ruling out GT being present on MRI (12% reduction).		A positive test conclusively establishes that GT will be present on MRI (~doubles probability), whereas a negative test will not be as useful in negating its presence (only lowers it by 12%).	

Figure 4 Interpreting LRs of some of the studied tests and their implications: an example with 50% pretest probability. *Clinical diagnosis is palpation plus at least one of the other clinical tests. #FADER-R: while lower bound CI is just inclusive of 1 and thus not statistically significant, is used here as an example as the 95% cut-off is an arbitrary convention. ^LR+ is an estimate, which was derived by adding 0.5 to each cell in the contingency table, so as to provide a denominator other than 0 and enable an LR+ to be estimated for the purpose of this example.²¹ ~Post-test probability is derived from: post-test odds, pretest odds×LR. FADER-R, Flexion, Adduction, External Rotation with resisted isometric hip internal rotation; GT, gluteal tendinopathy; LR, likelihood ratio; LR+, positive likelihood ratio; SLS, Single Leg Stance.

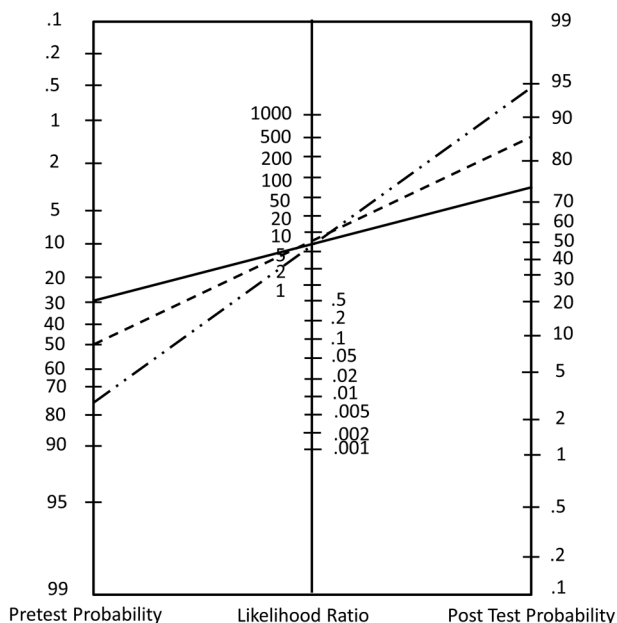


Figure 5 Example of the use of a nomogram to predict post-test probability from pretest probability and likelihood ratio;²⁴ for the FADER-R test, where there is a positive likelihood ratio of 6.6, the post-test probability would be 75% (line), 87% (dashed line) and 95% (dot-dashed line) for pretest probability of 30%, 50% and 75%, respectively. FADER-R, Flexion, Adduction, External Rotation with resisted isometric hip internal rotation.

diagnosis test in this study, at best only shifts the likelihood of the condition being present or not by about 15–20%. For the purposes of this study, we a priori defined the clinical diagnosis, which might not be most optimal in terms of clinical relevance. Another a priori decision that needs to be considered when interpreting the data from this study is the fixed order of testing, which might have influenced pain provocation. Future studies should be designed and powered to develop clinical

prediction rules that elucidate the best combination of a range of clinical tests and their order of testing in a clinical examination.

In conclusion, for the clinician assessing a patient with lateral hip pain, the combination of the most specific tests involving compression and active tension (SLS, FADER-R, ADD-R) and the most sensitive test, palpation, is likely to have greatest diagnostic utility. If a patient is not tender on palpation over the greater trochanter, it is highly advisable to consider other pathologies in the differential diagnosis of a patient with lateral hip pain. Furthermore, the use of a number of tests is advisable because a critical point stemming from this study is that the indices of diagnostic utility are at best modest and lack the conventionally accepted precision (ie, LR+ CI contain 1). In the event that there exists a positive finding of GT on MRI, a positive palpation with confirmation on at least one of the active clinical tests described should be considered necessary prior to the clinician diagnosing GT as the source of the patients' lateral hip pain.

Twitter Follow Alison Grimaldi at @alisongrimaldi

Acknowledgements The authors would like to thank the radiographers and radiologists at XRadiology Australia (Brisbane) and Victoria House Medical Imaging (Melbourne) for their assistance with this study.

Contributors BV, AG and RM contributed to study conception, design, project management, data collection and data analysis. All authors contributed to the interpretation and manuscript preparation and also approved the final manuscript.

Funding The study was undertaken as part of the National Health and Medical Research Council (NHMRC) Programme Grant #1091302. KB and PH are supported by NHMRC Fellowships (#1058440 and #1002190, respectively).

Competing interests None declared.

Ethics approval Human Research Ethics Committee of the University of Queensland and Behavioural and Social Sciences Human Ethics Sub-Committee of the University of Melbourne.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Falvey EC, Franklyn-Miller A, McCrory PR. The greater trochanter triangle; a pathoanatomic approach to the diagnosis of chronic, proximal, lateral, lower limb [corrected] pain in athletes. *Br J Sports Med* 2009;43:146–52.

- 2 Kingzett-Taylor A, Tirman PF, Feller J, *et al.* Tendinosis and tears of gluteus medius and minimus muscles as a cause of hip pain: MR imaging findings. *AJR Am J Roentgenol* 1999;173:1123–6.
- 3 Kong A, Van der Vliet A, Zadow S. MRI and US of gluteal tendinopathy in greater trochanteric pain syndrome. *Eur Radiol* 2007;17:1772–83.
- 4 Bird PA, Oakley SP, Shnier R, *et al.* Prospective evaluation of magnetic resonance imaging and physical examination findings in patients with greater trochanteric pain syndrome. *Arthritis Rheum* 2001;44:2138–45.
- 5 Collee G, Dijkmans BA, Vandenbroucke JP, *et al.* Greater trochanteric pain syndrome (trochanteric bursitis) in low back pain. *Scand J Rheumatol* 1991;20:262–6.
- 6 Sayegh F, Potouprnis M, Kapetanos G. Greater trochanter bursitis pain syndrome in females with chronic low back pain and sciatica. *Acta Orthop Belg* 2004;70:423–8.
- 7 Tortolani PJ, Carbone JJ, Quartararo LG. Greater trochanteric pain syndrome in patients referred to orthopedic spine specialists. *Spine J* 2002;2:251–4.
- 8 Howell GE, Biggs RE, Bourne RB. Prevalence of abductor mechanism tears of the hips in patients with osteoarthritis. *J Arthroplasty* 2001;16:121–3.
- 9 Reiman MP, Goode AP, Hegedus EJ, *et al.* Diagnostic accuracy of clinical tests of the hip: a systematic review with meta-analysis. *Br J Sports Med* 2013;47:893–902.
- 10 Grimaldi A, Mellor R, Hodges P, *et al.* Gluteal tendinopathy: a review of mechanisms, assessment and management. *Sports Med* 2015;45:1107–19.
- 11 Cook JL, Purdam C. Is compressive load a factor in the development of tendinopathy? *Br J Sports Med* 2012;46:163–8.
- 12 Cook JL, Purdam CR. The challenge of managing tendinopathy in competing athletes. *Br J Sports Med* 2014;48:506–9.
- 13 Soslowsky LJ, Thomopoulos S, Esmail A, *et al.* Rotator cuff tendinosis in an animal model: role of extrinsic and overuse factors. *Ann Biomed Eng* 2002;30:1057–63.
- 14 Birnbaum K, Siebert CH, Pandorf T, *et al.* Anatomical and biomechanical investigations of the iliotibial tract. *Surg Radiol Anat* 2004;26:433–46.
- 15 Delp SL, Hess WE, Hungerford DS, *et al.* Variation of rotation moment arms with hip flexion. *J Biomech* 1999;32:493–501.
- 16 Lequesne M, Mathieu P, Vuillemin-Bodaghi V, *et al.* Gluteal tendinopathy in refractory greater trochanter pain syndrome: diagnostic value of two clinical tests. *Arthritis Rheum* 2008;59:241–6.
- 17 Fearon AM, Scarvell JM, Neeman T, *et al.* Greater trochanteric pain syndrome: defining the clinical syndrome. *Br J Sports Med* 2013;47:649–53.
- 18 Blankenbaker DG, Ullrick SR, Davis KW, *et al.* Correlation of MRI findings with clinical findings of trochanteric pain syndrome. *Skeletal Radiol* 2008;37:903–9.
- 19 Cleland J, Koppenhaver S, Su J. The reliability and diagnostic utility of the orthopaedic clinical examination. In: Cleland J, Koppenhaver S, Netter F, eds. *Netter's orthopaedic clinical examination: an evidence-based approach*. Elsevier Health Sciences 2015:1–21.
- 20 Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The evidence-based medicine working group. *JAMA* 1994;271:703–7.
- 21 Deeks JJ, Higgins JP. Statistical algorithms in review manager 5. Statistical Methods Group of The Cochrane Collaboration, The Cochrane Collaboration. Statistical Methods Group of The Cochrane Collaboration, The Cochrane Collaboration 2010.
- 22 Deyle GD. Musculoskeletal imaging in physical therapist practice. *J Orthop Sports Phys Ther* 2005;35:708–21.
- 23 Woodley SJ, Nicholson HD, Livingstone V, *et al.* Lateral hip pain: findings from magnetic resonance imaging and clinical examination. *J Orthop Sports Phys Ther* 2008;38:313–28.
- 24 Akobeng AK. Understanding diagnostic tests 2: likelihood ratios, pre- and post-test probabilities and their use in clinical practice. *Acta Paediatr* 2007;96:487–91.



Utility of clinical tests to diagnose MRI-confirmed gluteal tendinopathy in patients presenting with lateral hip pain

Alison Grimaldi, Rebecca Mellor, Phillipa Nicolson, Paul Hodges, Kim Bennell and Bill Vicenzino

Br J Sports Med 2017 51: 519-524 originally published online September 15, 2016

doi: 10.1136/bjsports-2016-096175

Updated information and services can be found at:
<http://bjsm.bmj.com/content/51/6/519>

References

These include:

This article cites 22 articles, 5 of which you can access for free at:
<http://bjsm.bmj.com/content/51/6/519#BIBL>

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:
<http://group.bmj.com/subscribe/>