Barrett, M., Robertson-Waters, E., Whitehouse, M., Blom, A., & Berstock, J. (2019). Trochanteric spurs and surface irregularities on plain radiography are not predictive of greater trochanteric pain syndrome. *Hip International*. https://doi.org/10.1177/1120700019835641

Full title: Trochanteric Spurs and Surface Irregularities on Plain Radiography Are Not Predictive of Greater Trochanteric Pain Syndrome

Short tile: Trochanteric Spurs on Plain Radiography Are Not Predictive of GTPS

Authors:

Matthew C Barrett BSc.

Eve E Robertson-Waters BSc.

Michael R Whitehouse FRCS, PhD, MBChB, MSc, BSc, PGCertMedEd.

Ashley W Blom FRCS, PhD, MD, MBChB.

James R Berstock FRCS, MD, MBChB, PGCertMedEd.

Originating institution (all authors): Musculoskeletal Research Unit, Bristol Medical School, University of Bristol, Level 1 Learning and Research Building, Southmead Hospital, Westbury-on-Trym, Bristol BS10 5NB, UK

Corresponding author: Matthew Barrett

Fax: (+44) 117 41 47924 email: matt.barrett95@hotmail.com Tel: 07447070109

Manuscript type: Original research (diagnostic accuracy study)

Main text word count: 1500 (minus figure and table title/legends)

Abstract

Purpose: Surface irregularities of the greater trochanter have been described as a potential radiographic sign of greater trochanteric pain syndrome (GTPS). We report a diagnostic accuracy study to evaluate the clinical usefulness of trochanteric surface irregularities on plain radiographs in the diagnosis of GTPS.

Methods: We retrospectively identified the AP pelvic radiographs of a consecutive group of 38 patients (representing a 27.5% series prevalence) diagnosed with GTPS (mean age 69.5 years ± 16.1 [standard deviation], 27 females, 11 males) based on clinical symptoms and a positive response to a local anaesthetic and steroid injection. A control group consisted of 100 patients (mean age 73 years ± 17.1 [standard deviation], 67 females, 33 males) with either hip osteoarthritis listed for hip arthroplasty (n=50), or with an intracapsular neck of femur fracture (n=50) both presenting between January and July 2017. Radiographs were cropped to blind observers to the presence of hip osteoarthritis or intracapsular fracture but include the trochanteric region. The radiograph sequence was randomized and separately presented to three orthopaedic surgeons to evaluate the presence of trochanteric surface irregularities.

Results: The inter-observer correlation coefficient agreement was acceptable at 0.75 (95% CI: 0.60-0.84). Trochanteric surface irregularities including frank spurs protruding ≥2 mm were associated with a 24.7% positive predictive value, 64.0% sensitivity, 25.7% specificity, 74.3% false-positive rate, 36.0% false-negative rate, and a 65.3% negative predictive value for clinical GTPS.

Conclusion: Surface irregularities of the greater trochanter are not reliable radiographic indicators for the diagnosis of greater trochanteric pain syndrome.

Key words: enthesophytes; GTPS; hip; radiography; spurs

Introduction

Greater trochanteric pain syndrome (GTPS) is a poorly understood condition of the hip, characterised by disabling pain over the lateral aspect of the greater trochanter (1). Although it is still sometimes referred to as 'trochanteric bursitis', it is now also understood to be caused by tendinopathy or macroscopic tearing of the gluteus medius or minimus tendons. GTPS has a prevalence of approximately 18% in the general population (2), mainly affecting those between the ages of 40-60 years with a 4:1 female preponderance (3,4).

Diagnosis is made by history and clinical examination, however radiographs of the pelvis are routinely performed to exclude differential diagnoses such as degenerative hip joint disease, trochanteric avulsion fracture and periosteal neoplastic lesions or secondary malignancy of the proximal femur (5).

Irregularities of the surface of the greater trochanter have been described as a radiographic sign of GTPS (1,4). Surface irregularities include tendinous calcification, exostoses or enthesophytes of the greater trochanter, which may result from abductor tendinopathy or bursal inflammation (6,7). Steinert *et al.* found that 90% of hips with trochanteric enthesophytes protruding ≥2 mm from the cortical surface were associated with abductor tendon abnormalities and peritendinous edema on magnetic resonance imaging (8). It is unclear however, whether these surface irregularities of the greater trochanter are of any clinical significance. We report a

diagnostic accuracy study to evaluate whether surface irregularities of the greater trochanter on plain radiographs have any clinical utility in the diagnosis of GTPS.

Materials and Methods

Patient selection

We retrospectively identified the radiographs of a consecutive series of patients listed for surgical treatment of GTPS. These patients had all responded successfully, albeit temporarily, to an injection of local anaesthetic and steroid infiltrated around the trochanteric bursa region. Patients were included if an anteroposterior (AP) pelvic radiograph was available from the time of diagnosis of GTPS and fulfilled our radiograph selection criteria, as detailed below. These 56 patients have been reported in a previous study (9) that evaluated patient reported outcomes following trochanteric bursectomy and transposition of the gluteal fascia. The patients in this previous study were diagnosed with GTPS based on clinical history and greater trochanteric tenderness. They all underwent surgery when symptoms had persisted for over twelve months and a transient response to conservative measures, including corticosteroid injection and physiotherapy was observed.

For comparison, we used AP pelvic radiographs from 2 control groups both treated at our institution between January and July 2017. The first control group consisted of a consecutive series of patients with hip osteoarthritis listed for primary total hip replacement. The osteoarthritis group included patients of a similar age to the GTPS group (68.0 vs 69.5 years); as it is a possibility that osteoarthritis may overlap with GTPS, we included a second control group. The second control group were patients admitted with an intracapsular hip fracture. The clinical records for both control

groups were accessed through our institution's electronic database, BlueSpier (Bluespier, Droitwich, Worcester, UK). The clinical notes were reviewed to exclude patients reporting clinical features suggestive of GTPS.

Radiograph selection

Radiographs and radiological reports were accessed from our Picture Archiving Communication Service (PACS). All pelvic radiographs were standardised in the AP orientation. Where possible, patients were supine with internal rotation of the legs to 20°. The beam was not angled. All radiographs were anonymised and to blind observers to the underlying diagnosis, a standardised high-resolution section of each radiograph was cropped for radiographic interpretation by the observers. These sections permitted sufficient visualisation of the superior and lateral borders of the greater trochanter only, corresponding to the insertion points of the gluteus medius, minimus and vastus lateralis tendons. Patients were excluded if radiographs were not taken in a standardised AP orientation, if laterality was uncertain and if the standardised section of radiograph contained any evidence of the underlying diagnosis. Furthermore, patients in the control group were excluded if they underwent primary hip arthroplasty for an indication other than osteoarthritis.

Randomisation

We identified 50 patients with osteoarthritis, 50 patients with an intracapsular neck of femur fracture and 38 patients with GTPS, that met the eligibility criteria (see Figure I). Demographic information for the study group is displayed in Table I. Cropped radiographs for each patient were transferred to separate slides of a PowerPoint presentation (Microsoft®). Each slide was designated an arbitrary rank and a

random number generator (RANDOM.ORG; Dublin, IR) was used to randomly re-sequence the order of the radiographs throughout the PowerPoint presentation.

Analysis and statistical methods

Radiographs were presented sequentially and independently to three orthopaedic surgeons with subspecialist interest in the treatment of disorders of the hip. Due to the randomisation process, neither surgeon or researcher knew which group the patient belonged to during analysis. For each radiograph, observers were asked to assign each hip into one of three categories, similar to the methods of Steinert *et al.* (2010). Radiographs were described as 'normal' in the absence of trochanteric irregularities; 'small' indicated subtle cortical irregularities or enthesophytes protruding <2 mm from the cortical surface; 'large' indicated enthesophytes protruding ≥2 mm, see Figure II. All observers were blinded to the underlying pathology of each hip and the prevalence of GTPS within the series.

Given the different diagnostic 'cut-offs' the area under the receiver operating characteristic curve was determined. Statistical analysis was performed by Excel and IBM SPSS Statistics for Windows v22.0 (IBM Corp; Armonk, NY).

Results

Inter-observer agreement

A two-way mixed model intra-class correlation coefficient between the three observers was calculated to determine inter-observer agreement. The inter-observer correlation coefficient for three grades of trochanteric abnormality (normal, small surface irregularity or a frank spur ≥2 mm) was 0.75 (95% CI 0.60-0.84). This value

represents acceptable internal consistency of observations among the three observers.

Normal versus small or large enthesophytes

A total of 138 radiographs were included for assessment by the three observers. The prevalence of patients with clinical GTPS in this series was 27.5%. Overall, the positive predictive value of trochanteric surface irregularities or enthesophytes protruding ≥2 mm seen on plain radiographs was poor (24.7%) for clinical GTPS. Furthermore, such abnormalities were associated with a 64.0% sensitivity, 25.7% specificity, 74.3% false positive rate, a 36.0% false negative rate, and a 65.3% negative predictive value for clinical GTPS.

The diagnostic accuracy of detecting GTPS varied between observers, when comparing normal cortical surfaces of the greater trochanter to those with both ‘small’ and ‘large’ cortical irregularities, see Table II. Positive likelihood ratios were 0.83 for Observer 1, 0.82 for Observer 2 and 0.91 for Observer 3. Negative likelihood ratios were 1.40, 1.25 and 2.64 for each observer, respectively.

Normal or small vs large enthesophytes

The mean number of large enthesophytes recorded for each group was 7 (range 6-8) in the GTPS group, 10 (range 9-13) in the osteoarthritis group and 8 (range 3-17) in the hip fracture group. Refining the diagnostic criteria to include only large enthesophytes did not improve the diagnostic accuracy, see Table III.

Diagnostic Test Receiver Operating Characteristics

The area under the curve for the presence of cortical irregularities and frank spurs was 0.45 (95% CI 0.39-.52).

Discussion

Our results indicate that surface irregularities of the greater trochanter on plain radiographs are not reliable radiological indicators for the diagnosis of GTPS.

'Degenerative enthesopathy' is a term coined by Resnick and Niwayama (1983) to describe the enthesophytes that develop at tendinous insertions into bone. Such spurs may be associated with aging and factors such as microtrauma, muscular activity and local ischaemia (10). Steinert et *al.* reported that spurs ≥2 mm were associated with peritendinous edema at the insertions of gluteus medius and minimus on magnetic resonance imaging. In the series by Steinert et al., spurs were found to be 96-98% specific but only 18-28% sensitive, for edematous tendon abnormalities (8). It is noteworthy that peritendinous edema is not specific for bursal inflammation or symptomatic hip pain. Blakenbaker *et al.* report that 88% of 240 patients without trochanteric pain had peritrochanteric abnormalities (11). Thus, large spurs, although associated with peritrochanteric edema, may not be associated with clinically symptomatic GTPS.

Our study blinded observers to the underlying diagnosis and the prevalence of the target condition in the series. The observed radiographic images were carefully cropped in a standardised manner followed by randomisation of the order in which they were presented to observers to preserve blinding. Inter-observer agreement was acceptable, suggesting that cortical irregularities are reliably identifiable on

radiographs. We acknowledge small inherent age differences between the case and control groups. Accordingly, a greater prevalence of enthesopathy among the older hip fracture group was observed. In addition, minor changes in leg rotation at the time of performing the radiograph may rotate surface irregularities out of view. Controlling rotation following hip fracture is particularly difficult due to discomfort, and likewise internal rotation is usually restricted in patients with hip arthritis. Excessive external rotation within both arms of the control group may rotate trochanteric surface irregularities out of view. This would however serve to reduce any observed trochanteric abnormalities in the control group and artificially increase any association in the GTPS group. Our diagnostic accuracy data therefore represents a best-case scenario; it is possible that trochanteric abnormalities may be of even less use for diagnosing GTPS than we report.

Contrary to previous work, we conclude that cortical irregularities of the greater trochanter, including frank spurs, have no clinical utility in the diagnosis of the greater trochanteric pain syndrome.

Acknowledgements

No conflict of interest to declare. This paper was supported by the NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

References

1. Shbeeb MI, Matteson EL. Trochanteric bursitis (greater trochanteric pain syndrome). Mayo Clin Proc 1996;71(6):565-569.
2. Segal NA, Felson DT, Torner JC, et al. Greater trochanteric pain syndrome: epidemiology and associated factors. Arch Phys Med Rehabil 2007;88(8):988-992.
3. Williams BS, Cohen SP. Greater trochanteric pain syndrome: a review of anatomy, diagnosis and treatment. Anesth Analg 2009;108(5):1662-1670.
4. Chowdhury R, Naaseri S, Lee J, Rajeswaran G. Imaging and management of greater trochanteric pain syndrome. Postgrad Med J 2014;90(1068):576-581.
5. Ho GW, Howard TM. Greater trochanteric pain syndrome: more than bursitis and iliotibial tract friction. Curr Sports Med Rep 2012;11(5):232-238.
6. Gordon EJ. Trochanteric bursitis and tendinitis. Clin Orthop Relat Res 1961;20:193-202.
7. Fraser PR, Taylor V II, Reeves RE, Guttmann GD. Case study on the pathology of the greater trochanter. JSM Clin Case Rep 2014;2(6):1061.
8. Steinert L, Zanetti M, Holder J, Pfirmann CWA, Dora C, Saupe N. Are radiographic trochanteric surface irregularities associated with abductor tendon abnormalities? Radiology 2010;257(3):754-763.
9. Robertson-Waters E, Berstock JR, Whitehouse MR, Blom AW. Surgery for greater trochanteric pain syndrome after total hip replacement confers poor outcome. Int Orthop 2017;[Epub ahead of print]

https://link.springer.com/article/10.1007%2Fs00264-017-3546-7 Published 28 July 2017. Accessed 10 September 2017.

10. Resnick D, Niwayama G. Entheses and enthesopathy: anatomical, pathological, and radiological correlation. Radiology 1983;146(1):1-9.
11. Blakenbaker DG, Ullrick SR, Davis KW, De Smet AA, Haaland B, Fine JP. Correlation of MRI findings with clinical findings of trochanteric pain syndrome. Skeletal Radiol 2008;37(10):903-909.

Table I: Demographic details for each group

Group	Number of hips	Mean age (± SD)	Number of males (%)	Number of females (%)	Laterality (Rt vs Lt)
GTPS	38	69.5 (±16.1)	11 (29)	27 (71)	22 vs. 16
Osteoarthritis	50	68.0 (±12.9)	16 (32)	34 (68)	24 vs. 26
Hip fracture	50	78.6 (±19.3)	17 (34)	33 (66)	23 vs. 27

GTPS = greater trochanteric pain syndrome.

Table II: Diagnostic values for the detection of GTPS using any radiographic abnormality (small surface irregularities and spurs ≥2 mm).

Observer	Positive predictive value (%)	Sensitivity (%)	Specificity (%)	False-positive rate (%)	False-negative rate (%)	Negative predictive value (%)
Observer 1	23.9	57.9	30.0	70.0	42.1	65.2
Observer 2	23.7	47.4	42.0	58.0	52.6	67.7
Observer 3	25.8	86.8	5.0	95.0	13.2	50.0
Overall	**24.7**	**64.0**	**25.7**	**74.3**	**36.0**	**65.3**

Table III: Diagnostic values for the detection of GTPS using only large spurs ≥2mm as the diagnostic criteria.

Observer	Positive predictive value (%)	Sensitivity (%)	Specificity (%)	False-positive rate (%)	False-negative rate (%)	Negative predictive value (%)
Observer 1	36.4	21.1	86.0	14.0	78.9	74.1
Observer 2	25.0	15.8	82.0	18.0	84.2	71.9
Observer 3	21.1	21.1	70.0	30.0	78.9	70.0
Overall	**26.2**	**19.3**	**79.3**	**20.7**	**80.7**	**72.1**

Figure Captions

Figure I: Flowchart of patient selection. Other exclusion criteria in the control group included patients that underwent primary hip arthroplasty for an indication other than osteoarthritis, including hypochondroplasia, avascular necrosis and rheumatoid arthritis. Other exclusion criteria in the GTPS group included a case of avascular necrosis and peri-prosthetic fracture. GTPS = greater trochanteric pain syndrome; THR = total hip replacement; NOF = neck of femur.

Figure II: Standardised AP plain radiograph sections of the right greater trochanter. A) normal cortical surface in a 77-year-old male with GTPS. B) subtle cortical irregularities (arrows) in an 88-year-old male with an intracapsular neck of femur fracture; C) large superolateral enthesophyte (arrows) in an 84-year-old female with GTPS.

STARD checklist

Section & Topic	No.	×	Item
TITLE OR ABSTRACT			
Abstract page	1	x	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)
ABSTRACT			
Page 2	2	x	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)
INTRODUCTION			
Page 3	3	x	Scientific & clinical background, including the intended use and clinical role of the index test
Page 3, para 3, line 8	4	x	Study objectives and hypotheses
METHODS			
Study design ***Page 4, para 1, line 1***	5	x	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)
Participants ***Page 4, para 3, Fig 1***	6	x	Eligibility criteria
Page 4, para 1	7	x	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)
Page 4, para 1 and 2	8	x	Where and when potentially eligible participants were identified (setting, location and dates)
Page 4, para 2, line 3	9	x	Whether participants formed a consecutive, random or convenience series
Test methods	10a	x	Index test, in sufficient detail to allow replication
Page 5, para 3	10b	x	Reference standard, in sufficient detail to allow replication
Page, 5, para 3	11	x	Rationale for choosing the reference standard (if alternatives exist)
Page, 5, para 3	12a	x	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory
Page, 5, para 3	12b	x	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory
Page 5, para 3, line 3	13a	x	Whether clinical information and reference standard results were available to the performers/readers of the index test
Page 5, para 3, line 3	13b	x	Whether clinical information and index test results were available to the assessors of the reference standard
Analysis ***Page 6, para 1, line 7***	14	x	Methods for estimating or comparing measures of diagnostic accuracy
	15	**N/A**	How indeterminate index test or reference standard results were handled
	16	**N/A**	How missing data on the index test and reference standard were handled
	17	**N/A**	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory
	18	**N/A**	Intended sample size and how it was determined
RESULTS			
Participants ***Fig 1, page 6, para 2***	19	x	Flow of participants, using a diagram
Page 6, para 2, line 3	20	x	Baseline demographic and clinical characteristics of participants
	21a	**N/A**	Distribution of severity of disease in those with the target condition
Page 5, para1 (part of exclusion criteria)	21b	x	Distribution of alternative diagnoses in those without the target condition
	22	**N/A**	Time interval and any clinical interventions between index test and reference standard
Test results	23	**N/A**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard
Page 6, para 3	24	x	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)
	25	**N/A**	Any adverse events from performing the index test or the reference standard

DISCUSSION			
Page 8, para 2	26	**x**	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability
Page 8 para 2, page 9	27	**x**	Implications for practice, including the intended use and clinical role of the index test
OTHER INFORMATION			
	28	**N/A**	Registration number and name of registry
	29	**N/A**	Where the full study protocol can be accessed
	30	**N/A**	Sources of funding and other support; role of funders