Title

The clinical utility of imaging in osteoarthritis and its importance in future prediction of total hip replacement; a nested case-control study within the AGES-Reykjavik cohort

Authors' Names

Kenneth E S Poole, BM PhD, Department of Medicine, University of Cambridge, kesp2@cam.ac.uk

Ilya S Burkov, MSc, PhD Department of Medicine, University of Cambridge, ib352@cam.ac.uk;

Graham M Treece, MA PhD, Department of Engineering, University of Cambridge, gmt11@eng.cam.ac.uk;

Andrew H Gee, MA PhD, Department of Engineering, University of Cambridge, ahg@eng.cam.ac.uk;

Fjola Johannesdottir, PhD, Department of Medicine, University of Cambridge, fjohanne@bidmc.harvard.edu;

Simona D'Amore, MD PhD, Department of Medicine, University of Cambridge, simonadamore2@gmail.com;

Stephen K Kaptoge, PhD**,** Department of Public Health and Primary Care, University of Cambridge, skk22@medschl.cam.ac.uk

Sigurdur Sigurdsson, MSc, The Icelandic Heart Association, sigurdur@hjarta.is;

Thor Aspelund, PhD Professor, Public Health Sciences, University of Iceland & The Icelandic Heart Association, thor@hi.is;

Tamara B Harris, MD, Laboratory of Epidemiology and Population Sciences, Intramural Research Program, National Institute on Aging, harris99@nia.nih.gov;

Helgi Jonsson, MD PhD, Landspitali University Hospital & University of Iceland, helgijon@landspitali.is;

Vilmundur Gudnason, MD PhD, Faculty of Medicine, University of Iceland & The Icelandic Heart Association, v.gudnason@hjarta.is;

Thomas D Turmezei, MPhil MA BMBCh PhD, Department of Medicine, University of Cambridge & Department of Radiology, Norfolk and Norwich Hospital University Hospital, tom.turmezei@nnuh.nhs.uk;

Corresponding Author
Kenneth E S Poole Department of Medicine, Addenbrooke's Hospital Box 157, Level 5, Cambridge UK, CB2 0QQ kenneth.poole@nhs.net, +44 (0)1223 245151

Data availability
The data underlying this article will be shared on reasonable request to the corresponding author.

Funding
Kenneth E S Poole acknowledges funding and infrastructure support from the NIHR Cambridge Biomedical Research Centre, Addenbrookes Charitable Trust and Versus Arthritis Tissue Engineering Centre for this study. AGES-Reykjavik was funded originally by National Institutes of Health contract N01-AG-012100, the National Institute on Aging Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

Conflict of interest disclosure
The authors declare no conflict of interest.

Ethics approval
The study was approved (VSN 00-063) by the National Bioethics Committee in Iceland as well as the National Institute of Aging with. The full protocol for this sub-study was approved by the Icelandic Heart Association steering group.

Patient consent
All participants provided written informed consent.

ABSTRACT
Current guidelines are split on the role that imaging has in the clinical assessment of osteoarthritis, yet clinical computed tomography (CT) imaging has now revealed how a 3D approach can improve prediction of total hip replacement (THR) over 2D measures alone. We applied 2D grading and measurement along with 3D cortical bone mapping to ordinary clinical CT imaging of the pelvis in a cohort of healthy older people, aiming to discover which of these features had clinical utility in predicting total hip replacement (THR) within 8 years and which were related to baseline hip pain. Using a nested case-control design in the AGES-Reykjavik study, 74 future THR cases were age and sex-matched with 184 controls from the cohort (age 74±5yrs). Baseline assessment involved a validated hip pain questionnaire and pelvic CT. The following were performance-tested using ROC analysis and Clinical Utility Index: (i) hip pain; (ii) Kellgren and Lawrence grade (K&L grade), (iii) minimum joint space width (mJSW); and (iv) 3D cortical bone thickness (CTh). The clinical utility index for prediction of future THR from baseline pain was poor at 0.28, with the inclusion of imaging improving this to 0.79 (K&L grade) and 0.82 (3D CTh). Self-reported hip pain at baseline was also a poor-to-marginal predictor of THR (AUC=0.63), but 3D cortical thickening at the femoral head was predictive of future THR (0.81). Having radiographic osteoarthritis strongly predicted THR irrespective of hip pain (0.85). Combining hip pain, K&L grade and 3D cortical thickness gave optimal prediction (0.88). Ascertainment bias may have occurred if primary care physicians requested their own radiographs of their patients' hips. Imaging features from standard clinical CT identifies patients at high risk of progression to surgery for osteoarthritis, regardless of baseline pain.

LAY SUMMARY
Hip osteoarthritis is a common and debilitating degenerative joint disease that often results in chronic pain and disability affecting individual's quality of life and potentially leading to other conditions such as obesity, diabetes and cardiovascular disease. Early diagnosis and staging of arthritis offers the opportunity to tailor treatment plans to an individual patient's circumstances, particularly helping the general practitioner decide who to refer for hospital surgeons to consider definitive replacement of the arthritic hip joint. Our study shows that timely hip x-ray measurements, whether they are standard front-to-back x-rays or more complex 3D scans of older, healthy people with hip pain would greatly increase the general practitioners' ability to predict those patients who will require future hip replacement for severe osteoarthritis. Conversely, the same simple image results would allow the general practitioner to reassure most patients with hip pain that they will not require surgery for severe osteoarthritis in the coming years. We suspect that in Iceland, where this work was done, doctors

and nurses already follow these simple principles so we now suggest looking at outomes in a country such as the UK where general practitioners are strongly advised by national guidelines not to ask for x-rays in their patients.

ABBREVIATIONS

CBM = cortical bone mapping; CT = computed tomography; CTh = cortical thickness; K&L = Kellgren & Lawrence; mJSW = minimum joint space width; OA = osteoarthritis; OR = odds ratio; ROC = receiver operating characteristics; THR = total hip replacement; UI = utility index.

INTRODUCTION

Hip osteoarthritis (OA) is a highly prevalent and disabling joint disease with a significant individual and societal impact[1][2]. In the USA there has been a 20% increase in reported cases of OA in the decade since 2002[3], and from 2005 to 2030 total hip replacements (THRs) for OA are projected to increase by 174% [4]. Patients destined for THR typically have groin or medial thigh pain, stiffness, reduced hip movement on examination, but also a narrowed radiographic joint space[5], [6], [7]. Although the decision to offer THR depends partly upon patient-reported measures such as pain and functional limitation, the quality of secondary care referral of patients with hip OA might be enhanced if researchers could define specific and sensitive OA imaging features that have clinical utility[8], [9].

Much is already known about the utility of imaging in the clinical assessment of osteoarthritis. Seven years before THR, the baseline grade of radiographic hip OA was the strongest predictor of clinical progression, with the odds ratio (OR) for THR increasing from 6 (irrespective of baseline pain) to 24 (95%CI, 11–52) in those having both pain and radiographic OA[10]. In Iceland, where we have focused our studies, the hazard ratio of radiographic hip OA for incident THR over 11–28 years was 12.9 (95%CI, 7.9–21), regardless of symptoms[11]. Liviense et al. found that the radiographic structural features of OA such as narrowing of the hip joint space were the key predictors of eventual THR (alongside age >60yrs, pain and restricted hip movement) in a large sample of adults presenting to primary care with hip pain and followed for 3-6 years[7]. Predicting individuals at risk of progression to future surgical intervention is also an important step in identifying patients that might stand to benefit from currently available non-surgical therapies such as stem cell treatments.

Recommendations concerning how and when to request radiographs in diagnosis and referral vary across the major evidence-based guidelines [6], [12], [13], [14]. Some guidelines recommend an AP hip radiograph before, and to help guide surgical referral [12], [15], while others recommend imaging for patients with atypical presentations and unexpected rapid progression to identify alternative diagnoses[14]. Conversely, the authors of the UK NICE Osteoarthritis Quality Standard 87 (linked to guideline CG177) describe radiographs as *unnecessary*, *potentially harmful*, and *costly*, mandating that doctors do not request *any* imaging for hip OA diagnosis, management, or when considering referral [16]. Despite this, there remains discordance between guideline and practice; many clinicians (cognisant perhaps, of *spectrum bias*[17]) do find hip radiographs to be useful when deciding about referral and treatment.

Researchers have shown that radiological hip osteoarthritis is characterised by highly focal bony changes on the surface of the femoral head[18], [19], and proposed that increased Cortical thickness (CTh) could predate the development of painful disabling hip OA[20]. Recently, a narrow 3D joint space (JSW) mapped from computed tomography (CT) images showed optimal prediction of future THR in a cohort of healthy older adults when combined with 3D shape information and radiographic Kellgren & Lawrence grading (K&L)[21]. In the ongoing debate over the role of imaging in the care of patients with osteoarthritis, these studies suggest that CT has a facilitatory role in the clinical assessment of hip osteoarthritis.

In the present study, we aimed to discover which imaging features of hip OA (if any) had clinical utility in determining future THR for hip pain. Using a large, prospective, nested case-control study design within a healthy random sample of Icelandic men and women, we investigated 2D imaging measures (such as minimal joint space width (mJSW) and K&L grade along with 3D cortical bone thickness (CTh) from standard clinical CT imaging.

MATERIALS AND METHODS

Study Population

Participants (n=3133) were volunteers from Age, Gene/Environment Susceptibility-Reykjavik (AGES-Reykjavik), a single-centre prospective population study, which has been well described[22]. AGES participants were invited to attend the Icelandic Heart Association (Hjartavernd) for extensive phenotyping, including CT imaging. Study subjects were drawn from the 5,764 (58% female) survivors of the original sample of the population of Reykjavik born in 1907-1935, on whom studies began in 1967. Discharge records from all hospitals performing THR in Iceland were examined. Cases were reviewed by a rheumatologist (HJ) who confirmed that THR was for OA only. The nested case-control flowchart is shown in figure 1. A similar selection of participants was also used in the 2020 study by Turmezei et al. looking at the role of 3D JSW in prediction of future THR [21].

Baseline assessment of all 3133 volunteers included pelvic CT (Somaton Sensation 4, Siemens Medical Systems, Erlangen, Germany), with imaging of both hips from 1cm above acetabular margin to >3mm below the lesser trochanter acquired in the supine position. Spiral acquisition parameters were: 120kVp, 140mAs, pitch=1. Axial frames were reconstructed with 1mm slice thickness, 512x512 matrix in a 50cm field of view, and x-y pixel dimensions of 0.977mm with a standard soft tissue kernel. Hip-pain specific questions were from the Western Ontario and McMaster Universities WOMAC© and included "*HEALHIP12*. *In the past 12 months, have you had pain in or around either hip joint, including the buttock, groin or either side of the upper thigh, lasting at least one month?")*. During a maximum of 8 years (average 5 years) following participants who underwent baseline CT, 95 THR were reported. After screening, 74 subjects remained (48 F and 26 M), who had undergone 90 THR operations (16 bilateral THRs). Using a random selection procedure, at least two controls were selected for each THR case matched for year of recruitment, sex and age. After screening, 184 subjects remained (116 women and 68 men), giving a total of 258 included participants (figure 1). K&L grade assessment, mJSW measurement, and 3D osteoarthritis disease feature grading were performed on digitally reconstructed radiographs (table 1)[18], [23]. Digitally reconstructed radiograph (DRR) creation and CBM methods are shown in table 1[24], [25], [26]. The single operator (TDT) performing the mJSW scores previously showed substantial intra-observer reliability (0.63 95%CI 0.37-0.90) weighted kappa statistic for this methodology [18].

The study was approved (VSN 00-063) by the National Bioethics Committee in Iceland as well as the National Institute of Aging with all participants having provided written informed consent. The full protocol for this sub-study was approved by the Icelandic Heart Association steering group.

Statistical Methods

Statistical analyses were conducted using MATLAB R2016b (The MathsWork, Inc.), JMP® (v13.0 SAS Institute Inc.) and Stata Statistical Software (v14.2 StataCorp). Baseline measures were hip pain (HEALHIP12, K&L Grade, mJSW and mean 3D cortical thickness (CTh) from the significant ROI patch from statistical parametric mapping (SPM)–see table 1. Logistic regression determined the odds ratios (OR) for THR expressed per each SD increase in CTh or decrease in mJSW and per increment in grade of K&L with adjustment for age, sex, height, baseline pain, femur shape, and using leave-one-out cross-validation. The area under the receiver operating characteristic (AUC, ROC) curve was used to assess the performance of the model. ROC curves cut-offs were derived with JMP and the relationships between the various OA grades or increasing structural/imaging evidence of OA and hip pain were examined with nominal logistic fit methods. MATLAB was used to perform leave-one-out cross-validation of AUC values. Performance of the various logistic regression models, including imaging performance with and without hip pain was compared in STATA using the roccomp command and chi square analysis. Contingency analysis and positive and negative Clinical Utility index (*UI+* and *UI-*) was used to estimate the performance of diagnostic cut offs in regard to imaging utility in individuals presenting with baseline pain [9]. We modelled THR as a binary outcome using logistic regression because our primary aim was to compare baseline discriminative performance (AUC, UI+/-) of imaging/clinical measures to inform clinical decision-making and to simulate how imaging could aid patient stratification at the time of assessment. Although follow-up was variable, cases and their matched controls were selected to be comparable on recruitment year, age and sex, which reduces systematic differences in follow-up distribution between groups. We acknowledge that formal Cox modelling would provide estimates of hazard over time, but THR was an uncommon outcome in the whole cohort, with only 74/3133 (2.4%) cases recorded over 8 years of follow up. Therefore, it would be likely that time-to-event analysis would give similar results as logistic regression.

RESULTS

Patients were well matched at baseline (table 2). Two thirds of the participants destined for THR were female. The mean age of participants was 74.3±4.7 years and mean body mass index (BMI) was 27.7±4.2 kg/m^2. Cases underwent THR 35.8 months (±23.7, range 1–96 months) after their baseline CT. There were no significant differences in age, height, weight and BMI between the control group and THR group, respectively (table 2). Considering case hip DRRs, 65% of those graded had OA by K&L≥2 at baseline versus 10% of control hip DRRs (table 2).

CBM

Patients heading for THR had specific, focal and highly-conserved 3D structural changes of femoral head thickening by up to 70% at superior contact surface/articular margin, closely matching changes that we found previously in patients with OA from the UK[20] (figure 2).

Self-reported hip pain performance

The percentage of the entire AGES-Reykjavik cohort (n=3133) reporting hip pain at baseline was 24.3%. Similarly, the percentage reporting hip pain at baseline from our 258-participant sample was 26.0%. Only 45% (33/74) of patients destined for THR reported hip pain at baseline. The presence of hip pain was a poor-to-marginal predictor of THR (cross-validated AUC=0.63). Hip pain alone had poor sensitivity for THR of 44.6% (33.8–55.9), in contrast to any imaging measure (table 4A). The presence of self-reported baseline hip pain demonstrated very poor rule- generaliin accuracy (UI+ 0.28) and good rule-out accuracy for THR (UI- 0.72)[9].

Imaging performance

All imaging investigations ranged in Clinical Utility index from good to excellent (table 4B), except rule-out accuracy for mJSW which was adequate. Adding any radiological measurement to hip pain alone greatly enhanced the prediction of THR (table 4B). Adding a simple distance measure (2D mJSW) improved AUC discrimination from 0.63 to 0.78 (p=0.02). K&L grade alone gave an AUC of 0.85 (OR 6.30 per incremental K&L grade, 95%CI 4.0–10.0), unchanged with or without inclusion of baseline pain. K&L gave better discrimination than mJSW alone (0.85 vs 0.77 p=0.0056). 2D K&L grading and 3D CTh seem to provide different information, as manifested by excellent (and optimal) prediction of THR in a model containing; pain, K&L grade and 3D CTh (AUC=0.88, table 3).

DISCUSSION

Primary care health professionals such as general practitioners managing patients with hip osteoarthritis have a central role in referring individuals who might benefit from the *operation of the century*, as total hip replacement has been billed[27]. It is appreciated that the decision to refer for this operation could be aided by appropriate risk stratification tests with positive clinical utility[9]. Any test with negative clinical utility would also help health professionals reassure patients with hip osteoarthritis who are *not* likely to require THR assessment in the near term, which has utility when choosing non-surgical options for hip OA. A decade after the landmark paper by Kim et al., it is surprising that the NICE osteoarthritis guideline committee — perhaps influenced by cohort findings of discordance between hip pain and radiographic hip OA — elected to remove routine radiography from osteoarthritis care pathways [17]. Our findings do not support the guidance's characterisation of radiographs as "*unnecessary*". Future guideline revisions might benefit from revisiting the final paragraph of the BMJ editorial that accompanied the paper, which prudently delineated the role of imaging in the management of patients with hip pain[28]

When it comes to the role of imaging in assessing patients with hip pain, anteroposterior hip radiographs commonly demonstrate structural features of osteoarthritis — reduced joint space width, osteophytes, and femoral head subchondral sclerosis)[10], [11], [29], [30]. Opinion differs about routine imaging: some guidance advises against routine radiographs for suspected osteoarthritis [16] based on [17] while others continue to recommend radiographs to help differentiate among the common causes of adult hip pain[31] In the present nested case–control study of older adults from Iceland, we quantified clinically and surgically relevant features of hip osteoarthritis using pelvic CT, focusing on: (i) 2D Kellgren–Lawrence grading, (ii) 2D minimum joint space width (mJSW), and (iii) 3D cortical thickness of the proximal femur. In participants in their mid-70s who reported hip pain at baseline, these imaging measures had useful prognostic performance for predicting total hip

replacement within roughly three years. In our sample, a report indicating K&L ≥2 in a painful hip correctly identified most people who went on to THR and correctly reassured most who did not (table 4B). These findings align with larger UK Biobank analyses which identified that osteophyte burden and osteoarthritis grades clearly associated with prevalent hip pain and with later clinical outcomes including hospital-diagnosed OA and THR[32]. Notably, the UK analyses suggested that 2D osteophyte area related more strongly to pain than mJSW alone, implying that different radiographic features contribute variably to symptoms [32][33]

Guidelines differ in emphasis. NICE Quality Standard QS87 (based on NG226) recommends against routine imaging to diagnose osteoarthritis, stating there is limited evidence that imaging improves diagnostic management. [16], [34] We do not contend that every patient with hip pain requires imaging, but our data supports many others in confirming that imaging can materially inform management when the test result would change care[28] — for example, when assessing a patient with hip pain for suspected 'severe osteoarthritis' (a critical part of patient and clinician decision making [35]) , considering referral for surgery [35], excluding alternative pathology, or planning interventions [28]. In these scenarios, an imaging report showing definite radiographic OA in the painful hip (e.g., K&L ≥2) had good confirmatory and screening utility for subsequent THR (table 4B). Based on our data, with imaging reports available, the health professional seeing the individual with hip pain could correctly refer 85% (28/33) of people who would go on to require THR within 3 years and correctly reassure 95% (18/19) that THR would not be performed. Knowing that these patients' K&L score was ≥2 was excellent for confirmatory case finding (positive clinical utility index 0.82 (0.70, 0.94)), and good for screening; i.e. ruling out which patients would not require THR (negative clinical utility index 0.74 (0.62, 0.87)). Simpler measures (performed manually in seconds) had similar clinical utility: a minimum joint space ≤1.24mm for women and ≤1.73mm for men would correctly identify 76% (25/33) and reassure 89% (17/19 table 4B).

The multi-colour patch in the hip model in figure 2 is a critical anatomical zone when considering where the bone surfaces of hip and socket meet. Our 3D mapping technique has therefore uncovered a new marker of surgery-destined severe osteoarthritis. Bone thickness (CTh) values ≥1.20mm for women and ≥1.29mm for men in the patch would correctly identify 82% (27/33) of people with painful hips destined for total hip replacement within roughly 3 years and reassure 95% (18/19 table 4B) that they would not end up having surgery[35] It is not yet known if this focal thickening relates to joint shape. Morphological determinants (sex, size, shape mode) of 3D femoral cortical thickness in older people have previously been identified [36], [37]. In further work it would be interesting to explore relationships between femoroacetabular impingement (FAI)- relevant shapes (such as cam deformity) and the focally thickened areas identified in patients with surgically relevant osteoarthritis.

We retained the full Kellgren–Lawrence (K&L) 0–4 scale because it preserved prognostic discrimination across radiographic severity. Dichotomising K&L grade (for example into <2 versus ≥2) is useful for simple rule-in reporting — the general practitioner requester receiving a report that a painful hip has moderate-severe osteoarthritis (K&L ≥2) can then readily confirm causality and escalate management[32]. However, this binary approach sacrifices information: the risk of future THR increases markedly with each K&L increment (near-exponentially in our data), so models using the full 0–4 scale provide substantially better discrimination for individual risk stratification. Thus, while a dichotomous K&L≥2 flag is practical for immediate clinical decisions, maintaining the 0–4 grading on the report better informs prognosis and allows more nuanced stratification (for example, prioritising K&L 3–4 for expedited referral as per German National guidelines). Unlike NICE QS87 (but like the UK Royal College of Surgeons, 2017) the German 2021 guidelines consider radiographs to be critical, *"THR should be performed solely with radiologically demonstrated advanced osteoarthritis of the hip (K&L grade 3-4), after at least three months of conservative treatment, and in the presence of high subjective distress due to symptoms arising from the affected hip joint."* The rule-out performance of hip imaging is also clinically relevant: Logishetty et al. (2025) reported that patients undergoing total hip arthroplasty with minimal or no radiographic OA had lower postoperative function than typical THA patients and recommended both low-dose CT imaging and diagnostic injection to aid decision-making[38] Since total hip replacement has demonstrated superiority over conservative treatment for 'severe hip osteoarthritis' in a recent large RCT, imaging that helps define 'severe' disease may meaningfully inform referral and treatment prioritisation.[39]

Formal K&L grading is rarely done in clinical practice and is subject to intra- and inter-observer variability. Relying on a simpler descriptive OA grading may not be any better; one of us audited hospital radiology reports and found only 53% reliability from

routine clinical hip radiograph reports in allowing primary care doctors to determine which patients should be referred for orthopaedic surgery assessment (4% with no comment at all, 10% of reports that could have led to missed referral and 33% of reports that could have led to inappropriate referral, TDT *personal communication).* While semi-automated 3D bone mapping from CT has no prerequisite for radiological expertise, the higher radiation dose, availability, processing steps and cost of CT will clearly favour 2D radiographs. Since formal K&L grading requires specific expertise, we also tested the FDA-approved clinical trial gold standard of minimum superior joint space width measure which takes moments using any modern PACS reporting station[40]. This showed mJSW to be highly predictive of THR in symptomatic individuals, although not quite as good as K&L grade (AUC 0.77 compared to 0.85, table 3).

An important caveat is that these joint space measurements are only correct for supine radiographs (DRRs) digitally reconstructed from CT and might be different for plain radiographs; we lacked images from both techniques to allow comparisons. [41]Supine acquisition might yield slightly larger absolute JSW values than weight-bearing imaging [18], [23], [42]associations with THR are unlikely to be driven by supine acquisition alone. We note this as a limitation and recommend external validation against weight-bearing measures. As a clinical strategy this measure may also be limited by inter-subject variability since normal hip mJSW does vary across a 5mm range [43]. Since hip OA imaging is of most utility as a stratification test to aid with referral versus non-referral, a future priority is to extend our analysis to that larger AGES-Reykjavik sample. These aims are in keeping with those of the Quantitative Imaging Biomarker Alliance https://www.rsna.org/QIBA/. It would also be interesting to study the relationships between pain and imaging in the 3113 randomly sampled healthy Icelandic volunteers to compare with Kim et. al.'s younger US volunteers selected for their knee OA risk [13].

There are other limitations to our study. ROC analysis in matched case-control studies may necessitate adjustment for potentially confounding covariates, but we used standard unadjusted research methods. We lack information from clinical examination of the hip in the participants and a lack of data on patient surgical/conservative treatment preference. Finally, Iceland has a very high prevalence (fivefold higher than in southern Scandinavia) of radiological primary hip OA [30]. Hence generalisability beyond Icelandic population is unknown, although this study was the validation of an original discovery made in a UK cohort so it is likely that the findings are robust at least to the UK and Iceland[20]. Since the cohort was racially homogenous, comprised of white Icelandic individuals, this inevitably limits applicability to other groups.[40]

Previously, we demonstrated that the inclusion of detailed 3D joint space parameters alongside K&L grade delivered an 18% improvement on mJSW alone (AUC of 0.86 compared to 0.73 respectively) in a predictive model for the future THR in the same study cohort[44]. In the present study we are now able to propose that the 3D appearance of femoral cortical thickening can help characterise osteoarthritis destined for surgery. Faber et al. [33]found important associations between osteophyte size or location and hip pain [33]. In our cases destined for THR, osteophytes accounted for 25% of the variance in cortical thickness within the 3D ROI (CTh in ROI=1.2–0.01*Osteophyte load, $p<0.0001$, adj $R^2=0.25$, figure 2). While osteophytes did not fully explain cortical thickening, THR prediction in ROC analysis was the same whether 3D osteophyte load or CTh was modelled (AUC 0.83). We suspect that painful osteophyte-derived cortical thickening at the femoral head (blue in figure 2) is a key pathological entity in painful hip osteoarthritis. The 3D bone thickening we observed in figure 2 was particularly prominent at the femoral contact area with the acetabular labrum where osteophytes typically form and in the crescent of bone that is preferentially loaded during walking and sitting[45]. If such features are present in younger patients, these 3D imaging techniques might provide an opportunity to identify those whose nascent osteoarthritis is amenable to emerging, e.g. stem cell therapies [46]. Irrespective of the underlying cause of OA, this characteristic region of femoral cortical thickening not only predicted THR with an OR of 5.0 or each SD thicker (95%CI 3.2–7.7, table 3), but was also predictive of baseline hip pain with OR 2.1 (95%CI 1.5–3.1, table 5).

CONCLUSION

In conclusion, these data support targeted imaging assessment for osteoarthritis in older individuals presenting with hip pain when the result would influence management, since imaging features (including K&L grade, mJSW thresholds and quantitative osteophyte or 3D cortical measures) improve prediction of surgically-relevant disease and can aid referral, reassurance and treatment decisions

AUTHOR CONTRIBUTIONS

ISB, GMT, AHG, TDT, FJ, SS, TA, TBH, HJ, VG and KESP conceived and designed the study. ISB, TDT, SS, TA, TBH, HJ and VG acquired the data. ISB, GMT, AHG, TDT, TA, SKK and KESP analysed and interpreted the data. SDA and SKK provided drafting and writing advice. All authors drafted the article, and all authors were responsible for revision of critical important intellectual content.

ACKNOWLEDGEMENTS

Assistance with statistical analysis from Stephen Sharp, Senior Statistician of the University of Cambridge MRC Epidemiology Unit.

REFERENCES

[1] M. Fu, H. Zhou, Y. Li, H. Jin, and X. Liu, "Global, regional, and national burdens of hip osteoarthritis from 1990 to 2019: estimates from the 2019 Global Burden of Disease Study," *Arthritis Res Ther*, vol. 24, no. 1, p. 8, Dec. 2022, doi: 10.1186/s13075-021-02705-6.

[2] T. Jennison, A. MacGregor, and A. Goldberg, "Hip arthroplasty practice across the Organisation for Economic Co-operation and Development (OECD) over the last decade," *https://doi.org/10.1308/rcsann.2022.0101*, vol. 105, no. 7, pp. 645–652, Aug. 2023, doi: 10.1308/RCSANN.2022.0101.

[3] K. E. Barbour, C. G. Helmick, M. Boring, and T. J. Brady, "Vital Signs: Prevalence of Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation - United States, 2013-2015," *MMWR Morb Mortal Wkly Rep*, vol. 66, no. 9, pp. 246–253, 2017, doi: 10.15585/mmwr.mm6609e1.

[4] S. Kurtz, K. Ong, E. Lau, F. Mowat, and M. Halpern, "Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030," *J Bone Joint Surg Am*, vol. 89, no. 4, pp. 780–785, 2007, doi: 10.2106/JBJS.F.00222.

[5] G. E. Gold *et al.*, "OARSI Clinical Trials Recommendations: Hip imaging in clinical trials in osteoarthritis," *Osteoarthritis Cartilage*, vol. 23, no. 5, pp. 716–731, 2015, doi: 10.1016/j.joca.2015.03.004.

[6] D. Chu Miow Lin, W. M. Reichmann, L. Gossec, E. Losina, P. G. Conaghan, and J. F. Maillefert, "Validity and responsiveness of radiographic joint space width metric measurement in hip osteoarthritis: a systematic review," *Osteoarthritis Cartilage*, vol. 19, no. 5, pp. 543–549, 2011, doi: 10.1016/j.joca.2010.12.014.

[7] A. M. Lievense, B. W. Koes, J. A. Verhaar, A. M. Bohnen, and S. M. Bierma-Zeinstra, "Prognosis of hip pain in general practice: a prospective followup study," *Arthritis Rheum*, vol. 57, no. 8, pp. 1368–1374, 2007, doi: 10.1002/art.23094.

[8] M. G. Gademan, S. N. Hofstede, T. P. Vliet Vlieland, R. G. Nelissen, and P. J. Marang-van de Mheen, "Indication criteria for total hip or knee arthroplasty in osteoarthritis: a state-of-the-science overview," *BMC Musculoskelet Disord*, vol. 17, no. 1, p. 463, 2016, doi: 10.1186/s12891-016-1325-z.

[9] A. J. Mitchell, *(letter) Sensitivity x PPV is a recognized test called the clinical utility index (CUI+)*. 2011. doi: 10.1016/0895-4356(91)90128-v\.

[10] M. Reijman, J. M. Hazes, H. A. Pols, R. M. Bernsen, B. W. Koes, and S. M. Bierma-Zeinstra, "Role of radiography in predicting progression of osteoarthritis of the hip: prospective cohort study," *Bmj*, vol. 330, no. 7501, p. 1183, 2005, doi: 10.1136/bmj.38442.457488.8F.

[11] J. Franklin, T. Ingvarsson, M. Englund, O. Ingimarsson, O. Robertsson, and L. S. Lohmander, "Natural history of radiographic hip osteoarthritis: A retrospective cohort study with 11-28 years of followup," *Arthritis Care Res (Hoboken)*, vol. 63, no. 5, pp. 689–695, 2011, doi: 10.1002/acr.20412.

[12] "Painful Hip - Commissioning Guide — Royal College of Surgeons." Accessed: Sep. 23, 2025. [Online]. Available: https://www.rcseng.ac.uk/library-and-publications/rcs-publications/docs/painful-hip-guide/

[13] C. Kim *et al.*, "Association of hip pain with radiographic evidence of hip osteoarthritis: diagnostic test study.," *BMJ*, vol. 351, p. h5983, Dec. 2015, doi: 10.1136/bmj.h5983.

[14] G. Sakellariou *et al.*, "EULAR recommendations for the use of imaging in the clinical management of peripheral joint osteoarthritis," *Ann Rheum Dis*, vol. 76, no. 9, pp. 1484–1494, 2017, doi: 10.1136/annrheumdis-2016-210815.

[15] K.-P. Günther, S. Deckert, C. Lützner, T. Lange, J. Schmitt, and A. Postler, "Total hip replacement for osteoarthritis," *Dtsch Arztebl Int*, Oct. 2021, doi: 10.3238/arztebl.m2021.0323.

[16] "Recommendations | Osteoarthritis in over 16s: diagnosis and management | Guidance | NICE." Accessed: Sep. 23, 2025. [Online]. Available: https://www.nice.org.uk/guidance/ng226/chapter/Recommendations

[17] C. Kim *et al.*, "Association of hip pain with radiographic evidence of hip osteoarthritis: diagnostic test study," *Bmj*, vol. 351, p. h5983, 2015, doi: 10.1136/bmj.h5983.

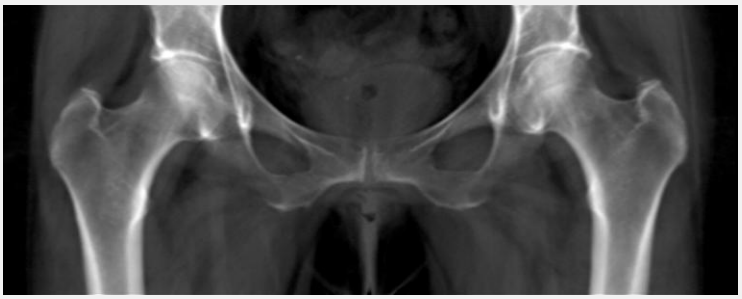
[18] T. D. Turmezei, A. Fotiadou, D. J. Lomas, M. A. Hopper, and K. E. Poole, "A new CT grading system for hip osteoarthritis," *Osteoarthritis Cartilage*, vol. 22, no. 10, pp. 1360–1366, 2014, doi: 10.1016/j.joca.2014.03.008.

[19] T. D. Turmezei, D. J. Lomas, M. A. Hopper, and K. E. Poole, "The development and reliability of a new CT grading system for hip osteoarthritis," *Osteoarthritis Cartilage*, vol. 21, pp. S189–S190, 2013, [Online]. Available: <Go to ISI>://WOS:000317942300390

[20] T. D. Turmezei, G. M. Treece, A. H. Gee, A. F. Fotiadou, and K. E. Poole, "Quantitative 3D analysis of bone in hip osteoarthritis using clinical computed tomography," *Eur Radiol*, vol. 26, no. 7, pp. 2047–2054, 2016, doi: 10.1007/s00330-015-4048-x.
[21] T. D. Turmezei *et al.*, "Quantitative 3D imaging parameters improve prediction of hip osteoarthritis outcome," *Sci Rep*, vol. 10, no. 1, p. 4127, Mar. 2020, doi: 10.1038/s41598-020-59977-2.
[22] T. B. Harris *et al.*, "Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics," *Am J Epidemiol*, vol. 165, no. 9, pp. 1076–1087, 2007, doi: 10.1093/aje/kwk115.
[23] T. D. Turmezei, D. J. Lomas, M. A. Hopper, and K. E. Poole, "Severity mapping of the proximal femur: a new method for assessing hip osteoarthritis with computed tomography," *Osteoarthritis Cartilage*, vol. 22, no. 10, pp. 1488–1498, 2014, doi: 10.1016/j.joca.2014.03.007.
[24] G. Treece and A. Gee, "Cortical Bone Mapping: Measurement and Statistical Analysis of Localised Skeletal Changes," *Curr Osteoporos Rep*, vol. 16, no. 5, pp. 617–625, 2018, doi: 10.1007/s11914-018-0475-3.
[25] G. M. Treece and A. H. Gee, "Independent measurement of femoral cortical thickness and cortical bone density using clinical CT," *Med Image Anal*, vol. 20, no. 1, pp. 249–264, 2015, doi: 10.1016/j.media.2014.11.012.
[26] G. M. Treece, A. H. Gee, P. M. Mayhew, and K. E. Poole, "High resolution cortical bone thickness measurement from clinical CT data," *Med Image Anal*, vol. 14, no. 3, pp. 276–290, 2010, doi: 10.1016/j.media.2010.01.003.
[27] I. D. Learmonth, C. Young, and C. Rorabeck, "The operation of the century: total hip replacement," *Lancet*, vol. 370, no. 9597, pp. 1508–1519, 2007, doi: 10.1016/S0140-6736(07)60457-7.
[28] M. J. Nieuwenhuijse and R. G. Nelissen, "Hip pain and radiographic signs of osteoarthritis," *BMJ*, p. h6262, Dec. 2015, doi: 10.1136/bmj.h6262.
[29] T. Ingvarsson, "Prevalence and inheritance of hip osteoarthritis in Iceland," *Acta Orthop Scand Suppl*, vol. 298, pp. 1–46, 2000, [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/11338422
[30] T. Ingvarsson, G. Hagglund, and L. S. Lohmander, "Prevalence of hip osteoarthritis in Iceland," *Ann Rheum Dis*, vol. 58, no. 4, pp. 201–207, 1999, doi: 10.1136/ard.58.4.201.
[31] R. Chamberlain, "Hip Pain in Adults: Evaluation and Differential Diagnosis," *Am Fam Physician*, vol. 103, no. 2, pp. 81–89, Jan. 2021, Accessed: Sep. 23, 2025. [Online]. Available: https://www.aafp.org/pubs/afp/issues/2021/0115/p81.html
[32] B. G. Faber *et al.*, "A novel semi-automated classifier of hip osteoarthritis on DXA images shows expected relationships with clinical outcomes in UK Biobank.," *Rheumatology (Oxford)*, vol. 61, no. 9, pp. 3586–3595, Aug. 2022, doi: 10.1093/rheumatology/keab927.
[33] B. G. Faber *et al.*, "Osteophyte size and location on hip DXA scans are associated with hip pain: Findings from a cross sectional study in UK Biobank.," *Bone*, vol. 153, p. 116146, Dec. 2021, doi: 10.1016/j.bone.2021.116146.
[34] "Overview | Osteoarthritis in over 16s | Quality standards | NICE".
[35] "NHS England » Decision support tool: making a decision about hip osteoarthritis." Accessed: Sep. 23, 2025. [Online]. Available: https://www.england.nhs.uk/publication/decision-support-tool-making-a-decision-about-hip-osteoarthritis/
[36] A. H. Gee, G. M. Treece, and K. E. S. Poole, "How does the femoral cortex depend on bone shape? A methodology for the joint analysis of surface texture and shape," *Med Image Anal*, vol. 45, pp. 55–67, 2018, doi: 10.1016/j.media.2018.01.001.
[37] A. H. Gee, G. M. Treece, C. J. Tonkin, D. M. Black, and K. E. S. Poole, "Association between femur size and a focal defect of the superior femoral neck," *Bone*, vol. 81, pp. 60–66, 2015, doi: 10.1016/j.bone.2015.06.024.
[38] K. Logishetty *et al.*, "Does total hip arthroplasty benefit patients with minimal radiological osteoarthritis?," *Bone Jt Open*, vol. 6, no. 3, pp. 328–335, Mar. 2025, doi: 10.1302/2633-1462.63.BJO-2024-0103.R1.
[39] T. Frydendal *et al.*, "Total Hip Replacement or Resistance Training for Severe Hip Osteoarthritis," *New England Journal of Medicine*, vol. 391, no. 17, pp. 1610–1620, Oct. 2024, doi: 10.1056/NEJMoa2400141.
[40] C. Ratzlaff, C. Van Wyngaarden, and J. Duryea, "Location-specific hip joint space width for progression of hip osteoarthritis--data from the osteoarthritis initiative," *Osteoarthritis Cartilage*, vol. 22, no. 10, pp. 1481–1487, 2014, doi: 10.1016/j.joca.2014.05.017.
[41] T. D. Turmezei *et al.*, "Quantitative Three-dimensional Assessment of Knee Joint Space Width from Weight-bearing CT," *Radiology*, vol. 299, no. 3, pp. 649–659, Jun. 2021, doi: 10.1148/radiol.2021203928.
[42] T. D. Turmezei and K. E. Poole, "Computed tomography of subchondral bone and osteophytes in hip osteoarthritis: the shape of things to come?," *Front Endocrinol (Lausanne)*, vol. 2, p. 97, 2011, doi: 10.3389/fendo.2011.00097.
[43] M. Lequesne, J. Malghem, and E. Dion, "The normal hip joint space: variations in width, shape, and architecture on 223 pelvic radiographs," *Ann Rheum Dis*, vol. 63, no. 9, pp. 1145–1151, 2004, doi: 10.1136/ard.2003.018424.
[44] T. D. Turmezei *et al.*, "Quantitative 3D imaging parameters improve prediction of hip osteoarthritis outcome," *Sci Rep*, vol. 10, no. 1, p. 4127, 2020, doi: 10.1038/s41598-020-59977-2.
[45] M. D. Harris, A. E. Anderson, C. R. Henak, B. J. Ellis, C. L. Peters, and J. A. Weiss, "Finite element prediction of cartilage contact stresses in normal human hips," *J Orthop Res*, vol. 30, no. 7, pp. 1133–1139, 2012, doi: 10.1002/jor.22040.
[46] M. Lotz *et al.*, "Value of biomarkers in osteoarthritis: current status and perspectives," *Ann Rheum Dis*, vol. 72, no. 11, pp. 1756–1763, 2013, doi: 10.1136/annrheumdis-2013-203726.

TABLES

Table 1. Imaging Techniques and Osteoarthritis Grading

Digital Reconstructed Radiographs and K&L Grading	Single experienced MSK radiologist (TDT) created Digital Reconstructed Radiograph (DRRs) (13) ● TDT blinded to THR status/clinical outcomes and pain status. ● PACS viewing software (OsiriX, Pixmeo Sarl; v.3.9.3). Multiplanar reformatting (MPR) of mean intensity projection slab at each hip with a soft tissue window. ● Ensured coronal coverage of anterior and posterior joint margins. ● Coronal depth thickness of up to 8cm DRR simulated standard pelvic radiograph (example below). ● K&L grading and mJSW performed at 200% magnification. 
3D Cortical Bone Mapping (CBM)	Cortical thickness measured across the proximal femur using cortical bone mapping (Stradwin software) ● Each proximal femur segmented into a surface from a triangular mesh of vertices **{1-3}** ● Thickness measured to an accuracy of approximately 0.3mm (mean+/-SD error 0.01+/-0.58mm) **{4}** ● Cortical surface measurement locations co-registered onto single canonical femur **{5}** ● Location and magnitude of patches of bone thickening associated with future THR highlighted by statistical parametric mapping, SPM**{6}** ● Detailed statistical methods regarding the 3D CBM technique and methods are in (18-20) ● General linear model (GLM) for SPM allowed for group (case/control), age, sex, weight, height, size (shape mode 0) and the first six non-rigid shape modes (SM), with Surfstat package in MATLAB ● GLM CTh = 1 + group + age + sex + weight + height + SM0 + SM1 + SM2 + SM3 + SM4 + SM5 ● Two-tailed T-testing across the cohort for uncorrected p-values at each measurement location ● Random field theory gave p-values corrected for multiple comparisons (controlling for type I errors). ● Average thickness calculated from the significant 'patch' in each individual used in logistic regression.

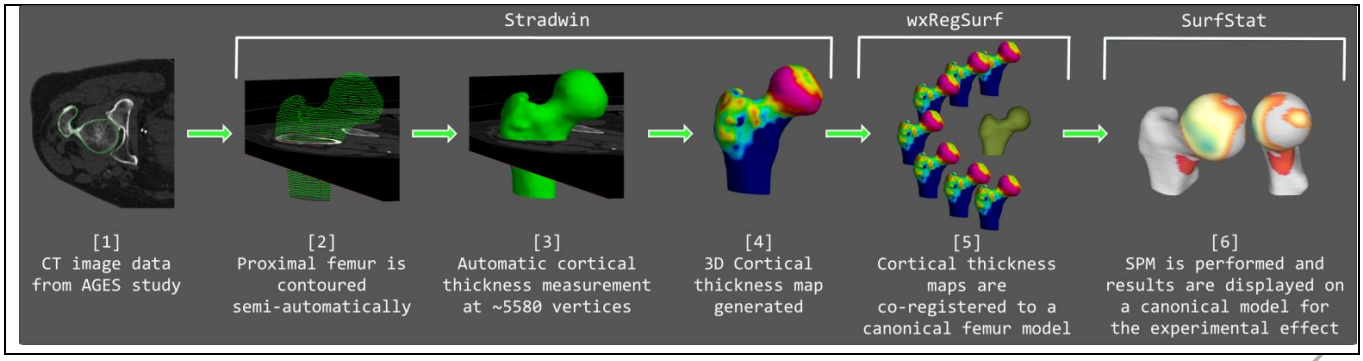


Table 2. Baseline variables in study cases (who eventually underwent THR for osteoarthritis) and matched controls

Baseline Data	**Controls**	**Cases**	*Female Controls*	*Female Cases*	*Male Controls*	*Male Cases*
Number of subjects (%)*	**184 *(71%)***	**74 *(29%)***	116	48	68	26
Age ± SD (years)	**74.0 ± 4.7**	**74.0 ± 4.7**	74.1 ± 5.0	73.6 ± 4.6	74.8 ± 4.6	74.8 ± 4.9
Height ± SD (m)	**1.68 ± 0.09**	**1.67 ± 0.09**	1.62 ± 0.05	1.62 ± 0.06	1.77 ± 0.07	1.75 ± 0.06
Weight ± SD (kg)	**77.5 ± 13.6**	**78.1 ± 12.7**	73.1 ± 12.8	76.5 ± 12.0	85.0 ± 11.6	80.9 ± 13.6
Cortical Bone Thickness in Femoral Patch (*from CBM fig 2*) ± SD (mm)	**1.1 ± 0.1****	**1.3 ± 0.2****	1.1 *(± 0.1)***	1.3 *(± 0.2)***	1.1 *(± 0.1)***	1.4 *(± 0.2)***
mJSW from DRR median + IQR (mm)	**2.3** *(1.8, 2.8)***	**1.1** *(0.6, 1.9)***	2.0 *(1.5, 2.6)***	1.1 *(0.4, 1.8)***	2.6 *(2.0, 3.0)***	1.6 *(1.0, 2.6)***
Hip pain in the 12 months prior to imaging investigation	**19 *(10%)***	**33 *(45%)***	15 *(13%)*	24 *(50%)*	4 *(6%)*	9 *(35%)*
Number of hips, *(baseline DRR analysed)*	**368**	**90**	232	59	136	31
KLG 0; no OA	**279 *(76%)***	**18 *(20%)***	170 (73%)	9 (15%)	109 (80%)	9 (29%)
KLG 1; doubtful OA	**53 *(14%)***	**13 *(15%)***	34 (15%)	8 (14%)	19 (14%)	5 (16%)
KLG 2; mild OA	**26 *(7%)***	**19 *(21%)***	20 (9%)	13 (22%)	6 (4%)	6 (19%)
KLG 3; moderate OA	**10 *(3%)***	**30 *(33%)***	8 (3%)	23 (39%)	2 (2%)	7 (23%)
KLG 4; severe OA	**0 *(%)***	**10 *(11%)***	0 (0%)	6 (10%)	0 (0%)	4 (13%)

**Entire Dataset; **p<0.0001 case vs. control*

Table 3. Predicting total hip replacement for osteoarthritis within 8 years; using baseline hip pain alone versus adding imaging tests. Comparisons of AUC of the various models), model n=258.

Model (n=258) predicting THR	**AUC (95% CI)**	**Cross-validated AUC**	**OR per SD (95% CI)**	**Difference between using Imaging versus hip pain only (χ^2)**	**Imaging versus hip pain only prob > χ^2**
Hip pain only	0.70 (0.62–0.78)	0.63	2.21 (1.68–2.91)	–	–
mJSW only	0.79 (0.72–0.86)	0.77	0.26 (0.17–0.39)	–	–
[i]*hip pain + mJSW*	*0.80 (0.73–0.87)*	*0.78*	–	*5.75*	*0.0164*
K&L grade only	0.87 (0.81–0.92)	0.85	6.30 (3.96–9.99)	–	–

[ii]hip pain + K&L	*0.87 (0.82–0.93)*	*0.85*	*–*	*20.35*	*0.0000*
Cortical thickness in femoral patch (*fig. 1*)	0.83 (0.77–0.89)	0.81	5.00 (3.24–7.71)	–	–
[iii]hip pain + CTh	*0.85 (0.79–0.91)*	*0.83*	*–*	*14.30*	*0.0002*
hip pain + K&L + CTh	*0.90 (0.85–0.95)*	*0.88*	*–*	*26.80*	*0.0000*

ROC = Receiver operating characteristic; AUC = area under the curve; OR = odds ratio per standard deviation (SD); CTh = cortical thickness; mJSW = minimum joint space width; K&L = Kellgren & Lawrence Score; χ^2=chi-square test; i vs ii, χ^2=9.25, p=0.0024; i vs iii, χ^2=2.21, p=0.1375; ii vs iii, χ^2=0.48, p=0.4881.

Table 4. A. Performance of hip pain alone in predicting THR B. Effects of adding any imaging test result to patients presenting with hip pain on Clinical Utility Index (7)a, n=52 (33 from case group and 19 from control group)

A. Baseline pain status predicting THR (n=258)	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**PLR**	**NLR**
Hip pain (WOMAC)	44.6%	89.7%	63.5%	80.1%	4.3	0.60
UI+ 0.28 / UI- 0.72	(33.8–55.9)	(84.4–93.3)	(49.9–75.2)	(74.1–85.0)	(2.6–7.1)	(0.5–0.76)
B. Utility of imaging for those presenting with pain (n=52)	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**PLR**	**NLR**
+ K&L grade*	84.8%	94.7%	96.6%	78.3%	16.1	0.2
UI+ 0.82 / UI- 0.74	(69.1–93.3)	(75.4–99.1)	(82.8–99.4)	(58.1–90.3)	(2.4–109.2)	(0.1–0.4)
+ 2D mJSW**	75.8%	89.5%	92.6%	68.0%	7.2	0.27
UI+ 0.70 / UI- 0.61	(59.0–87.2)	(68.6–97.1)	(76.6–97.9)	(48.4–82.8)	(1.9–27.1)	(0.1–0.5)
+ CT-derived 3D CTh***	81.8%	94.7%	96.4%	75.0%	15.5	0.2
UI+ 0.79 / UI- 0.71*	(65.6–91.4)	(75.4–99.1)	(82.3–99.4)	(55.1–88.8)	(2.3–105.5)	(0.1–0.4)

*PPV = Positive Predictive Value; NPV = Negative Predictive Value; PLR = Positive Likelihood Ratio; NLR = Negative Likelihood Ratio; K&L = Kellgren & Lawrence Score. *Outcome by a K&L grade of ≥ 2. **For females JSW < 1.24mm and males < 1.73mm. ***For females, femoral cortical thickness > 1.2mm and males, > 1.29mm. These diagnostic cut-offs were derived using JMP add-in ('ROC and pAUC Analysis' by D Meintrup). Qualitative interpretation of Utility Index (UI+ in prediction of positive outcome, UI- in prediction of negative outcome) from Mitchell, A (7) ≥0.81 excellent; ≥ 0.64 good; ≥ 0.49 adequate; ≥ 0.36 poor and ≤ 0.36 very poor.*

Table 5. Odds ratios of radiographic osteoarthritis for prevalent hip pain. In addition, multiple linear regression was calculated to predict mJSW based on age, sex and hip pain, replicating Dougaros et al. 1996 (29). A significant regression equation was found (F(3,454)=26.1 p<0.0001), with an R^2 of 0.15. Participants predicted mJSW was equal to 3.12mm–0.35(pain)-0.02(age)+0.22(sex) where pain was coded as 1=yes, 2=no, age was measured in years and sex was coded as 1=male, female=2. mJSW was 0.35mm thinner in those with hip pain and males had 0.22mm wider mJSW than females.

K&L Classification	Number of Hips	K&L Grade	OR (95% CI)	P Value
Doubtful	66	1	0.83* (0.24–2.28)	0.7363
Mild	45	2	1.98* (0.69–4.95)	0.1914

Moderate/Severe	50	3+4	12.22* (6.00–25.46)	<0.0001
*Moderate****	*40*	*3*	*8.92 (4.10–19.42)*	*<0.0001*
*Severe****	*10*	*4*	*50.78 (10.05–256.58)*	*<0.0001*
mJSW	458	–	2.68** (1.64–4.36)	*<0.0001*
CTh	258	–	2.14** (1.46–3.12)	*<0.0001*

*OR = odds ratio per standard deviation (SD); CTh = cortical thickness; mJSW = minimum joint space width; K&L = Kellgren & Lawrence Score. *versus none **per SD narrowing (mJSW) or thickening (CTh) ***individual K&L grades*

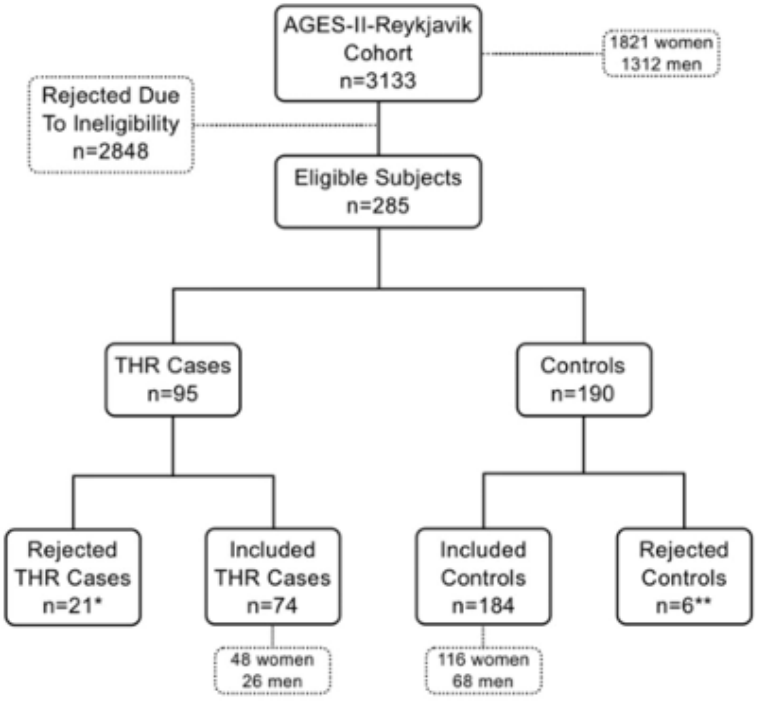


Figure 1. Study Flowchart showing the nested case-control design within the large healthy prospective ageing cohort ** 21 rejected: 18 THR not due to OA, 2x incomplete scans and 1x motion artefacts. ** 6 rejected: 4x motion artefact. 1x incomplete scan, 1x unreliable phantom density value*

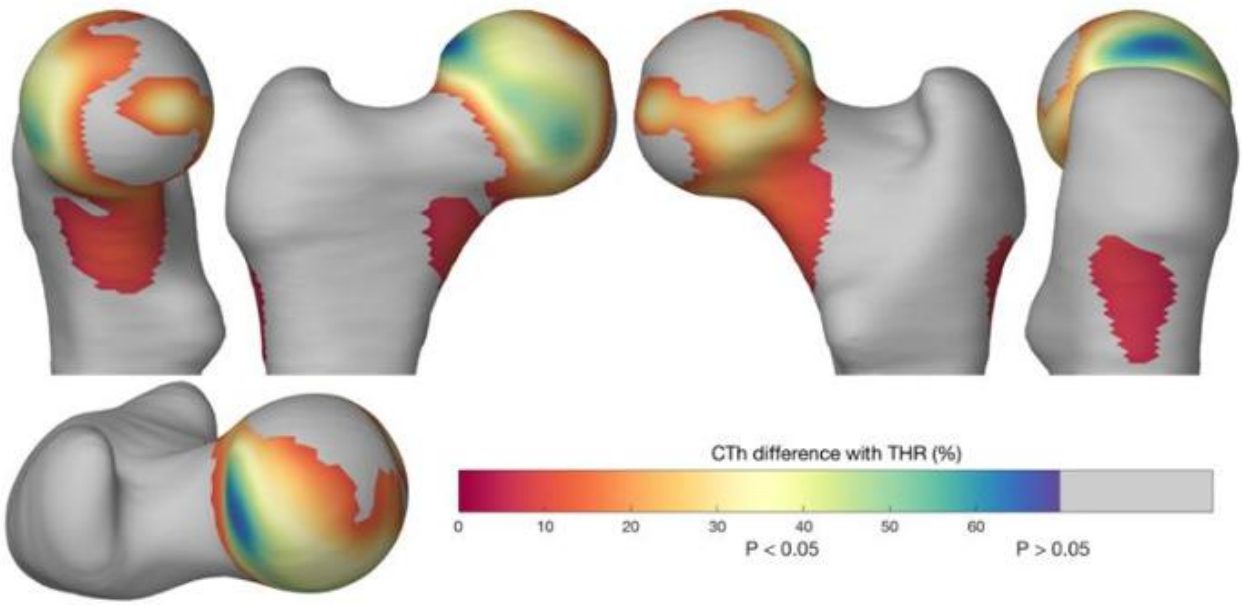


Figure 2. Statistical parametric mapping results.

*Differences in baseline cortical thickness between subjects who went on to have THR and matched controls are displayed as a statistical difference map with a colour scale for the magnitude of the relative differences. Grey areas are regions where there was no statistical significance between THR patients and controls. Blue, yellow and orange areas demonstrate 35–70% thicker femoral head bone in the focal areas. The average cortical thickness from within the multi-coloured ROI was used for further analysis, with values and SD shown in Table 2. The model containing only the average CTh from this ROI (alongside age, sex, height) predicted incident THR (Table 3. A) well, with an AUC of 0.83 (OR 5 per SD thickening, 95%CI 3.2–7.7). Adding baseline pain to the model increased AUC slightly to 0.85. 3D osteophyte load accounted for 25% of the variance in cortical thickness within the ROI (CTh in ROI=1.2–0.01*Osteophyte load, p<0.0001, adj R^2=0.25). While osteophytes did not fully explain cortical thickening, THR prediction in ROC analysis was the same whether 3D osteophyte load or CTh was modelled (AUC 0.83). Thinning of the joint space was only weakly associated with cortical thickening in cases (not controls) explaining 6% of the variance in CTh (CTh in ROI=1.4–0.05*mJSW, p=0.03, adj R2=0.06)*