Radiomic MRI Features for Predictive Modelling of Knee Osteoarthritis: A Deep Learning Approach

by

Edward Peake, MSc, BSc



Thesis submitted to The University of Nottingham for the degree of Doctor of Philosophy

August 2023

I want to express my heartfelt gratitude to my supervisors, with particular thanks to Prof. Dorothee Auer for her unwavering guidance and support throughout my PhD. This thesis is dedicated to my family, who have shared this journey with me.

Forward

This thesis was completed at the Nottingham National Institute of Health Research (NIHR) Biomedical Research Centre. NIHR Biomedical Research Centres aims to accelerate translational research to improve clinical practice and policy. Research is established into themes for each centre, representing areas of local excellence, with collaborations spanning university and NHS organisations.

The set theme for this PhD was musculoskeletal research, focusing on osteoarthritis. The thesis relies on MRI to investigate possible advancements for better patient outcomes in osteoarthritis. As well as providing excellent tissue contrast, MRI was the modality of choice given the expertise of local departments, including the Sir Peter Mansfield Imaging Centre and Precision Imaging Beacon of Excellence.

Training and development were supported through engagement with the NIRH Research Academy and the University of Nottingham's N-trans programme. Funding from a diverse range of sources, including the University of Nottingham, the NIHR, and the Institute of Physics and Engineering in Medicine, facilitated my attendance at several international conferences, such as the International Society for Magnetic Resonance in Medicine's Joint Annual Meetings in 2019,2020, 2021 and 2022, as well as attendance at the Osteoarthritis Research Society International's World Congress in 2019. As well as an opportunity to present research as part of this PhD, these meetings facilitated engagement with the broader research community in both OA and MRI.

This thesis explores deep learning and radiomics to create new predictive models that may expand our understanding of knee osteoarthritis risk factors, improve biomarkers for clinical trials and identify high-risk individuals early to increase treatment options.

Abstract

Background

The study of osteoarthritis (OA) with artificial intelligence applied to magnetic resonance imaging (MRI) has opened new avenues for clinical practice and driven considerable investment in research. This includes the development of new mechanistic biomarkers for better prognosis, improved clinical endpoints for clinical trials and personalised medicines.

Proposal

For this thesis, deep learning is investigated using radiomic features extracted from segmented knee MRI in patients with osteoarthritis. To segment knee cartilage a fully automated ensemble learning approach may be used to examine the association between knee symptoms and MRI cartilage features. Also proposed is the use of MRI cartilage features to build a prediction model of time to knee replacement. The thesis also aims to investigate the ability of MRI cartilage features and additional structural targets to predict knee pain in osteoarthritis.

Materials and Method

For the development of a cartilage segmentation model, an ensemble learning approach is proposed to improve sensitivity for knee cartilage segmentation. The model combines predictions from a two-label network for identifying cartilage and a multi-label network for classifying specific cartilage compartments. Training, validation and testing were performed on MR images acquired with Double Echo Steady State (DESS) and manual segmentations for 6 different compartments from 88 participants with knee osteoarthritis from the Osteoarthritis Initiative (OAI). Performance is evaluated using Dice coefficient and Bland Altman plots. The method was applied to baseline knee images from the OAI (n = 9383) and investigated the association between segmentation-derived features including cartilage volume and intensity range, and symptoms including WOMAC scores for pain, stiffness and function using Pearson's correlation. Covariates of interest were not explored for these features, as the objective of this first investigation was to probe the utility of simple cartilage features before more complete predictive modelling.

For the development of an MRI prediction model of time to knee replacement, a retrospective study was performed using data from the Osteoarthritis Initiative collected between 2006-2015. Cartilage segmentations were automatically generated using the ensemble segmentation model for 1370 participants (mean age 61±9, 57% female) in the Progression Cohort: participants with symptomatic knee OA and 3D double echo steady state MRI. Elastic net regression was used to develop a radiomics signature from MRI features, followed by Cox proportional hazard regression to evaluate risk prediction of future knee replacement. The radiomics signature and Cox risk model were compared with clinical and demographic models using Kaplan-Meier curves, proportional hazard ratio (HR) estimates, and c-statistics. Pipeline and model testing was performed in the Incident Cohort of the OAI (n = 6628): participants at increased risk of developing symptomatic knee OA.

For the investigation of MRI features for the prediction of pain in knee osteoarthritis, the correlation between baseline knee cartilage MRI features, knee pain and established 9-year pain trajectories was examined using the progression cohort of the OAI. These features were then used to predict knee pain and pain phenotypes. The addition of new structural markers was introduced to explore if these would improve the classification of knee pain in osteoarthritis. To provide the additional segmentation a knee atlas was constructed and used for registration before extraction of radiomic features. As well as the classification results, the effect of atlas-based registration on the image features were explored to examine if this improved corrections of individual features with knee pain.

Results

The knee cartilage segmentation model showed excellent accuracy for all 6 types of cartilage: femoral 0.88, medial tibial 0.84, lateral tibial 0.88, patellar 0.85, medial meniscal 0.85 and lateral meniscal 0.90. The average volume correlation was 0.988, overestimating volume by 9% ± 14% over all compartments. Intensity range and cartilage volume were negatively associated with clinical WOMAC scores for pain, function and stiffness but only weakly (Range -0.08 to -0.15). For WOMAC pain, function and stiffness, the intensity range was more strongly associated with WOMAC scores in femoral, medial tibial and lateral meniscal cartilage when compared to volume measurements ($p < 0.01$).

The MRI prediction model of time to knee replacement showed good predictive ability of MRI. In the Progression Cohort, 171 participants underwent knee joint replacements within 5 years of baseline MRI. The radiomics cartilage signature had a C-index of 0.81 (0.77 – 0.84) and HR of 7.5 (7.1 – 7.9, $p< .001$). In the Incident Cohort, 63 participants underwent knee replacement surgery and the radiomics cartilage signature at baseline had a C-index of 0.82 (0.75 – 0.86) and HR of 9.6 (9.2 – 10.1, $p< .001$).

The correlation of cartilage MRI features derived from automated knee segmentations and pain in knee osteoarthritis showed small but significant associations between cartilage volume and baseline pain scores (r-values -0.01 to -0.10) and 9-year pain trajectories (r-values 0.01 to -0.13). Cartilage volume was in the 44th centile for the prediction of baseline pain, and 56th centile for the prediction of pain trajectories when compared to other radiomic features.

Using cartilage MRI features, neural network classification was most effective for the prediction of outcomes with an AUC of 0.63 for baseline pain and an AUC of 0.64 for pain trajectory. The inclusion of additional segmentation for the infrapatellar fat pad, and bone marrow lesion areas in the tibiofemoral joint led to marginal improvements in the prediction of baseline pain (AUC 0.65). The atlas used for image registration was found to have a normalizing effect of radiomic features, reducing the standard deviation in cartilage volume by over 30%; this improved the correlation between cartilage volumes and baseline pain scores by an average of 90%.

An examination of the association of radiomics features from new structures with pain scores showed that radiomic features from the infrapatellar fat pad were more strongly associated with pain scores compared to other structures, while regions for bone marrow

lesions performed poorly. However, this had little impact on the overall predictive power for the classification of pain scores.

Conclusion

The cartilage segmentation model produces anatomically plausible segmentations with high dice coefficients when compared to expert manual annotations. The results of correlation analysis of radiomic features suggest features other than cartilage volume were more closely associated with knee pain.

The proposed radiomics signature for the prediction of knee replacement within 5 years improves the predictive ability of standard clinical and demographic features in both symptomatic and incident cases. Individualized risk score and cartilage visualisation for knee osteoarthritis based on MRI may provide new outcome metrics for clinical trials and could improve dialogue between patients and clinicians.

The examination of radiomics features for the prediction of knee pain in osteoarthritis showed limited performance when using radiomic features derived from cartilage alone. The inclusion of MRI features from other joint structures had only a small effect on the prediction of knee pain. Registration of images to an anatomical atlas did improve the association between radiomic features and knee pain and appeared to have a normalizing effect on cartilage volumes. Of the new anatomical targets tested, the infrapatellar fat pad showed the strongest association with knee pain compared with other targets. This adds to a growing body of evidence for an inflammatory component of osteoarthritis with potential new markers that can be derived via MR imaging.

Contents

Forward 2
Abstract 3
Background 3
Proposal 3
Materials and Method 3
Results 4
Conclusion 5
1. Introduction 10
1.1 Definition of osteoarthritis 12
1.2 Symptoms 13
1.3 Prevalence 14
1.4 Incidence 15
1.5 Disease burden 16
1.6 Cost of disease 16
1.7 Diagnosis 17
1.8 Risk Factors 18
1.9 Comorbidity 19
1.10 Pathology 19
1.11 Pain mechanisms 21
1.12 The search for disease-modifying therapeutics 23
1.13 Image Features 24
1.13.1 Articular cartilage 25
1.13.2 Bone marrow lesions and cysts 26
1.13.3 Osteophytes 27
1.13.4 Hoffa synovitis 28
1.13.5 Synovial effusion 29
1.13.6 Ligaments 29
1.13.7 Meniscus 30
1.14 Disease management in knee osteoarthritis 31
1.14.1 Pain Medication 31
1.14.2 Non-pharmacological treatment 32
1.14.3 Joint replacement 33
1.14.4 Knee osteotomy 34
1.14.5 Arthroscopic knee surgery 34

1.15 Unmet needs 34
1.16 Potential added value of advanced MRI. 35
1.16.1 Imaging 35
1.16.2 Artificial Intelligence 37
2 Motivation, Aims and Overview of Thesis 41
3 Methods 44
3.1 Deep learning 44
3.1.1 Supervised learning 45
3.1.2 Formal definition 46
3.2 Radiomics 51
3.2.1 Segmentation 52
3.2.2 Feature Extraction 55
3.2.3 Modelling 55
3.2.4 Validation 59
4 Ensemble Learning for Robust Knee Cartilage Segmentation 61
4.1 Introduction 61
4.2 Materials and Methods 62
4.2.1 Study Dataset 62
4.2.2 Data Preparation 62
4.2.3 Model architecture and training 63
4.2.4 Post-processing 64
4.2.5 Model evaluation 65
4.2.6 Association with clinical outcomes 65
4.3 Results 66
4.4 Discussion 72
4.5 Conclusion 75
5. Prediction of joint replacement in knee osteoarthritis 76
5.1 Introduction 76
5.2 Materials and methods 77
5.2.1 Study Participants and clinical-demographic features 77
5.2.2 MRI Measures 79
5.2.3 MRI Image segmentation 79
5.2.4 MRI Features 80
5.2.5 Radiomics model 80
5.2.6 Clinical and demographic model 80
5.2.7 Combined model 81

5.2.8 Testing in incident cases 81
5.3 Results 82
5.3.1 Radiomics model 82
5.3.2 Clinical and demographic model 85
5.3.3 Combined model 85
5.3.4 Testing in incident cases 85
5.4 Discussion 86
5.5 Conclusion 87
6. Pain Phenotyping with Cartilage 88
6.1 Introduction 88
6.2 Method 89
6.2.1 Study participants 89
6.2.2 Outcome measures 90
6.2.3 Radiomic features 92
6.2.4 Correlation analysis 92
6.2.5 Pain prediction models 93
6.3 Results 93
6.3.1 Correlation analysis 93
6.3.2 Pain prediction models 97
6.4 Discussion 104
6.5 Conclusion 105
7 Pain Phenotyping with other structural targets 106
7.1 Introduction 106
7.2 Method 107
7.2.1 Study participants 107
7.2.2 Knee Atlas 108
7.2.3 Radiomic features 108
7.2.4 Pain prediction model 109
7.3 Results 109
7.3.1 Knee Atlas 109
7.3.2 Radiomic features 113
7.3.2 Pain Prediction Model 117
7.4 Discussion 119
7.5 Conclusion 120
8 Conclusion 122
8.1 Contribution to knowledge 122

8.2 Cartilage segmentation 122
8.3 Prediction of Total Knee Replacement 123
8.4 Pain phenotyping using cartilage MRI 125
8.5 Pain phenotyping part extended methods 126
8.6 Future developments 127
9 Acknowledgements 129
10 References 130

1. Introduction

Osteoarthritis is a degenerative joint disease whose hallmarks are pain, stiffness and a loss of function or movement. It is an insidious disease with symptoms that are often gradual and increase with older age. As well as individual suffering, the disease has a high economic cost, with medical treatments estimated to be between 1 – 2.5% of GDP in developed countries (1). Osteoarthritis is a common disease worldwide, with incident knee OA present in 1 in 5 adults aged 40 years and older (2). As global populations age, osteoarthritis has become one of the fastest-growing causes of disability worldwide (3).

The lack of existing treatments has accelerated the search for ways to slow or prevent the progression of osteoarthritis. Clinical trials of new therapies and interventions have numerous challenges. Endpoint outcome measures for osteoarthritis, such as surgery, led to long study durations, with the risk of high levels of patient dropout. Clinical endpoints relying on expensive healthcare options may be confounded by societal inequalities that limit access to health services (4).

In this context, imaging-based biomarkers for osteoarthritis may provide an unbiased assessment of structural changes. These can both help accelerate the discovery of new treatments and enable optimised trial design through careful risk stratification of populations. There may also exist routes for these tools into clinical settings if disease phenotypes that might benefit from early or alternative treatments can be identified. This aligns well with NHS England's strategy of improving outcomes through personalised medicines (5).

Magnetic resonance imaging (MRI) gives unique insights into anatomical changes in osteoarthritis; it provides a high-quality assessment of cartilage, bones, soft tissue and ligaments. Traditionally, the use of MRI in clinical settings has been limited due to robust diagnosis of OA from clinical indications and treatment pathways that do not require imaging. In some instances, plain radiographs may be used where the presentation is atypical. However, the rise of machine learning and artificial intelligence assessment of clinical images has opened new avenues for improving patient care using/based on advanced imaging.

In the last decade, the development of high-performance computing, coupled with extensive training datasets, has propelled the growth of artificial intelligence across society. Within medicine, radiology has seen by far the most significant increase in the use of artificial intelligence and machine learning, as demonstrated in data published by the US Food and Drug Administration (Figure 1). These changes may be driven by the large amounts of healthcare data recorded in modern healthcare settings, including rich medical imaging data, combined with rapid innovation transfer across sectors.

Notable examples within radiology include using artificial intelligence to identify breast cancer in mammograms, detection of diabetic retinopathy in retinal fundus photographs, skin cancer classification from mobile phone images and lung cancer screening using low-dose chest computer tomography (6–9).

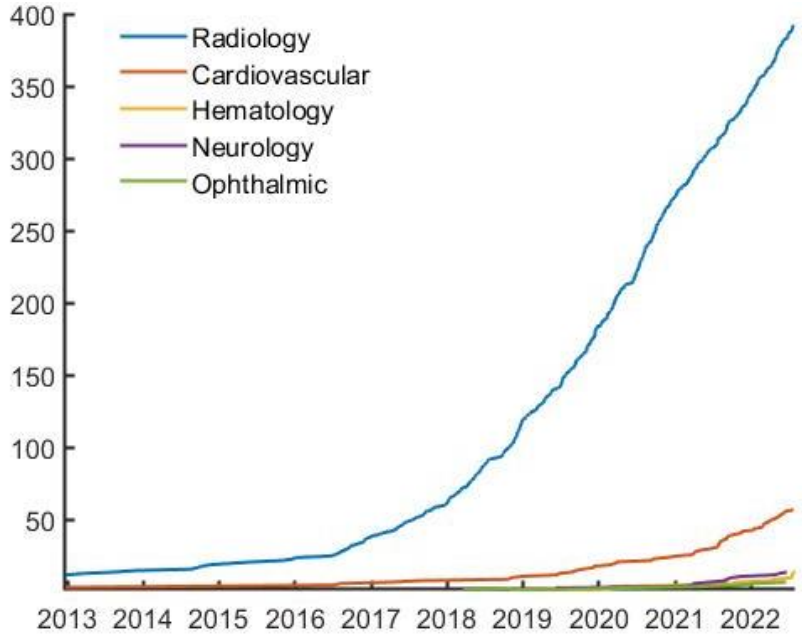


Figure 1: Increasing Adoption of AI and Machine Learning in Medical Devices
This figure, based on data from the US Food and Drug Administration (FDA), illustrates the significant rise in the number of medical devices incorporating artificial intelligence (AI) and machine learning (ML) that have received FDA 510(k) clearance in recent years (10). It reflects the growing impact of AI and ML in healthcare, particularly in radiology. This trend is driven by advances in high-performance computing and the availability of extensive training datasets, which have enabled more sophisticated and precise medical imaging analysis.

This thesis explores the suitability of MRI for creating robust imaging markers for knee osteoarthritis. A focus on using radiomic features for predicting outcomes such as knee pain and future joint replacement is explored. Leveraging artificial intelligence and machine learning methods now found across radiology, the added benefit of detailed imaging of anatomical changes in OA is examined. In addition to the critical role in facilitating research and future aspirations of personalised treatments, mechanistic outcome markers and AI tools for diagnosis and prognosis may change the clinical landscape.

AI-based reports visualising and quantifying the anatomical changes within OA may provide innovative tools for clinicians to understand their patients' conditions, accelerate diagnosis, personalise healthcare plans, and improve patient-to-practitioner dialogue around their conditions. In insurance-driven health systems such as the US, automated reporting of exams may reduce cost, simplify insurance disbursements and fast-track access to disability-based social security payments.

The introduction to this thesis provides an up-to-date overview of osteoarthritis, focusing on additional clinical information provided by magnetic resonance imaging. The unmet need in osteoarthritis is explored, including possible directions for new OA research. Given the lack of existing therapies, the urgent need for treatments to slow or prevent the progression of osteoarthritis is established. The introduction concludes with the potential benefits of MRI for musculoskeletal imaging, including examples of AI imaging applications.

1.1 Definition of osteoarthritis

Osteoarthritis is a disease that can have significant detrimental effects on individuals and can cause years of pain and joint stiffness. To date, there is no single definition for this disease.

The American College of Rheumatology (ACR) definition for OA is "[a] heterogeneous group of conditions that leads to joint symptoms and signs which are associated with defective integrity of articular cartilage, in addition to related changes in the underlying bone at the joint margins."

The National Institute for Health and Care Excellence (NICE) definition is more detailed: "Osteoarthritis is characterised pathologically by localised loss of cartilage, remodelling of adjacent bone and associated inflammation. Various traumas may trigger the need for a joint to repair itself. Osteoarthritis includes a slow but efficient repair process that often compensates for the initial trauma, resulting in a structurally altered but symptom-free joint. In some people, because of either overwhelming trauma or compromised repair, the process cannot compensate, resulting in eventual presentation with symptomatic osteoarthritis; this might be thought of as 'joint failure'. This explains the extreme variability in clinical presentation and outcome observed between people and at different joints in the same person."

The Osteoarthritis Research Society International (OARSI) definition states: "Osteoarthritis is a disorder involving movable joints characterised by cell stress and extracellular matrix degradation initiated by micro- and macro-injury that activates maladaptive repair responses including pro-inflammatory pathways of innate immunity. The disease manifests first as a molecular derangement (abnormal joint tissue metabolism) followed by anatomic and physiologic derangements (characterised by cartilage degradation, bone remodelling, osteophyte formation, joint inflammation and loss of normal joint function) that can culminate in illness."

The definition used by the ACR describes a disease with a core set of underlying anatomic abnormalities but a wide range of clinical phenotypes. The suggestion that osteoarthritis is a collection of conditions may suggest differing disease pathways.

The NICE definition also includes a description of the anatomical abnormalities common to osteoarthritis. Still, it suggests the aetiology of the disease is traumatic, and progression results from failed repair processes. This disease definition is appealing, as it offers clean transitions from initial injury through pre-clinical disease that can tip into symptomatic OA. However, much of the literature examined in this introduction into osteoarthritis suggests this is a simplified approach due partly to a wide range of risk factors, diverse disease pathways with whole joint involvement, and poor association of structural findings with clinical scores.

The definitions from OARSI, an international community dedicated to osteoarthritis research, are the most specific but would be least accessible to a general audience. It

addresses the complexities of disease initiation with damage on a micro and macro scale, followed by a maladaptive repair process. Similar to the definition from NICE, there is the transition from pre-clinical disease to anatomical abnormalities. However, the definition also recognises that illness does not necessarily result from the progression of the disease.

1.2 Symptoms

The onset of knee pain often indicates underlying osteoarthritis (11). Pain is the most disabling symptom and drives clinical decision-making (12). It is sometimes described as a deep ache (13). In the knee, pain can be localised to the medial or lateral compartments and may be correlated to underlying anatomical changes in those areas; medial pain is most common (13), while lateral pain is also associated with valgus deformity (14). The onset of OA symptoms is often gradual, and symptoms may vary between joints; an overview is given in Table 1.

Symptoms		**Signs**	
Pain	Weakness	Joint enlargement	Limitation of motion
Altered function	Deformity	Altered gait	Deformity
Stiffness	Grinding/clicking	Tenderness	Instability
Swelling	Instability/buckling	Crepitus	

Table 1: Signs and symptoms of osteoarthritis
The onset of OA is gradual and usually begins in one or a few joints (13).

Stiffness is another common symptom of osteoarthritis and is most frequently experienced in the morning or evening or after long periods of inactivity (15). Unlike other forms of arthritis, stiffness in osteoarthritis can usually be resolved in minutes by joint motion (14).

Loss of movement or function can occur as anatomical changes and joint derangement progress. These symptoms can limit day-to-day activities such as climbing stairs, walking and doing household chores. It is a common symptom and is another reason patients visit their family doctor (16). Symptomatic OA is associated with depression or disturbed sleep, which diminishes the patient's quality of life (17).

Another symptom of OA includes joint enlargement due to effusion or bony swelling. OA can cause crepitus, a sensation of crackling or crunching of the joint, and may be associated with changes in articular cartilage (16,18). Contractures of the soft tissue can lead to changes in alignment, resulting in valgus or varus malalignments, as shown in Figure 2. In valgus alignment, knees are angled outwards, which increases the stress on the lateral side of the knee. In varus alignment, knees are bent inwards, which increases the stress on the medial side of the knee. Both cases can increase the wearing of the joint due to uneven loading of the joint, potentially accelerating OA in a given compartment.

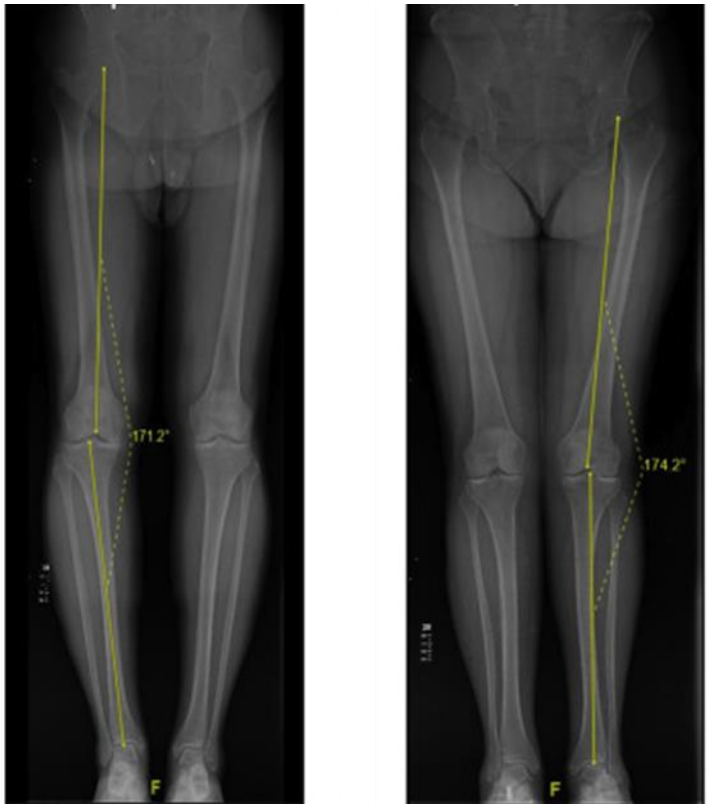


Figure 2: Variations in Knee Alignment in Advanced Osteoarthritis
This X-ray image, adapted from Chapleau et al. (19), demonstrates the anatomical changes in knee alignment due to advanced osteoarthritis, showcasing both valgus (left) and varus (right) alignments. Valgus alignment is characterized by an outward knee angle, stressing the lateral side, while varus alignment features an inward knee angle, increasing medial side stress. These alignment changes can accelerate the wear and tear associated with osteoarthritis.

1.3 Prevalence

Osteoarthritis can affect several sites; the knee, hip and hand are most affected (20,21). While the knee is the most common site, a systematic review showed that the reported prevalence of OA depended on the definition of OA used, the age, country of origin and gender distribution of the study population (22).

The prevalence of OA increases with age, and after the age of 50 years, a higher proportion of women are affected compared to men. The disease has a high global prevalence. The worldwide population living with OA has been estimated at 250 million (4% of the total population), with knee OA constituting 83% of cases in 2010; a more recent estimate from 2020 considered the number of individuals with knee OA alone to be 654 million (8% of the total population) (23,24).

In the Rotterdam study (25), a population-based cohort study (n = 3906) of individuals over 55 years, the proportion of women with radiographic hand osteoarthritis was 67%

compared to 55% of men. Similarly, in the knee, in individuals over 80 years, 53% of women and 33% of men had radiographic OA.

The 2020 study on the incidence and prevalence of OA worldwide found substantial variations between countries (24); the distribution by country is shown in Figure 3. Knee osteoarthritis had a global prevalence of 16% (95% CI, 14% - 18%) and an incidence of 20 per 1000 person-years (95% CI, 11% - 33%).

A within-population study of Swedish registry data (primary and secondary care data) estimates an increase in doctor-diagnosed OA from 26.6% to 29.5% for any location, from 13.8% to 15.7% for the knee, and from 5.8 to 6.9% for the hip (21), showing a high prevalence of knee OA.

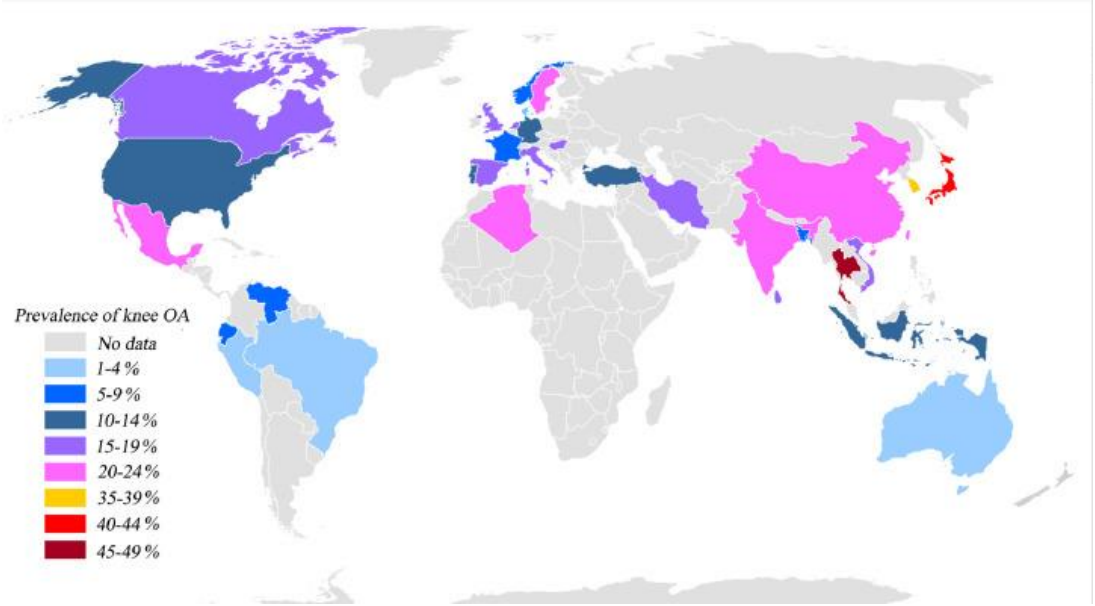


Figure 3: Global distribution and prevalence of knee osteoarthritis
This figure illustrates the worldwide prevalence of knee osteoarthritis based on an extensive study involving over 9 million participants across 34 countries (24). It highlights the variation in knee OA prevalence globally, driven by age, gender, and regional reporting. The infographic highlights the lack of data on the prevalence of OA over much of Africa and some areas of Central Asia and South America.

1.4 Incidence

Spanish and UK general practice data has been used to estimate the incidence of osteoarthritis in the general population, with similar levels between the two populations (20,26). The incidence of hand, hip, and knee OA increased rapidly (much more rapidly in women than men) between 50 and 75 years (Figure 4). The prevalence of osteoarthritis in these studies was 16-17% for the knee, 6% for the hip and 4-5% for the hand (20,26).

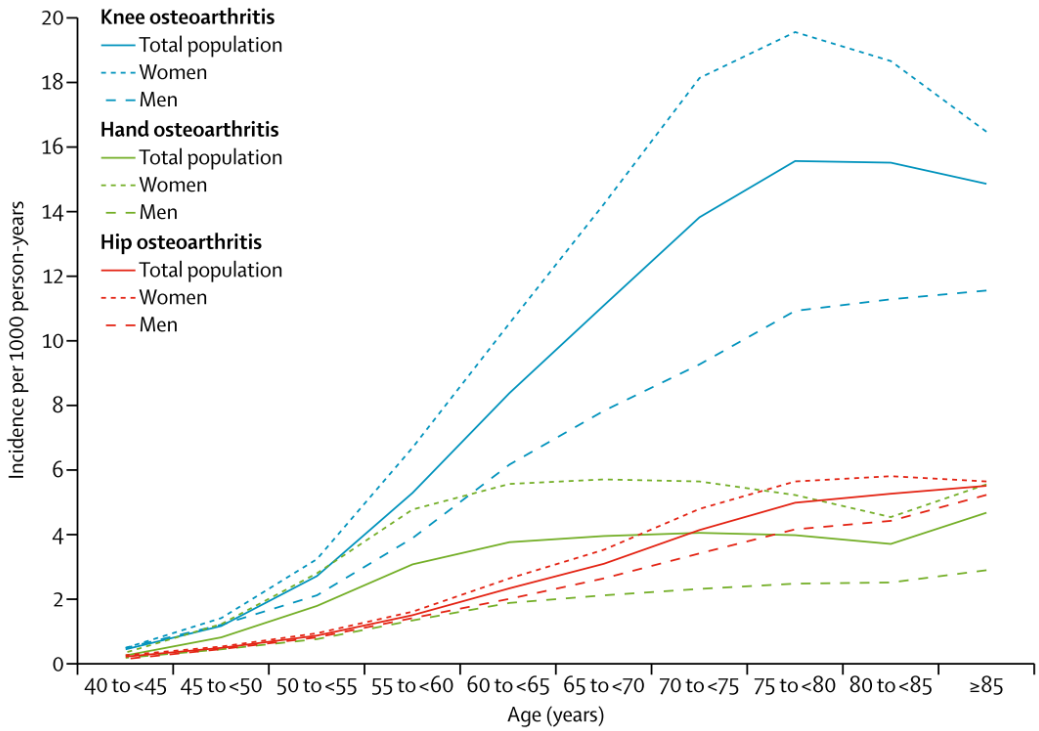


Figure 4. Comparative Incidence of Osteoarthritis by Age and Gender
Incident rates (per 100,000 person-years) of knee, hip, and hand osteoarthritis based on data representative of the general population in Spain, with parallel findings observed in the UK (20,26). It highlights the rapid increase in the incidence of osteoarthritis, particularly between the ages of 50 and 75, with a higher prevalence in women than men. Understanding the epidemiological patterns of osteoarthritis can aid in resource management, focusing research and developing health and social care policy.

1.5 Disease burden

Given the substantial progress in reducing mortality and extending life expectance over the past few decades, non-fatal diseases such as osteoarthritis are growing in importance as they prevent individuals from living in good health. Knee osteoarthritis accounts for 85% of the years lived with disability across all forms of osteoarthritis worldwide (2). Alongside diabetes, osteoarthritis was responsible for the most significant increase in years lived with disability, relative to the top 20 causes of disability, with the number of years lived with disability increasing by 106% between 1990 and 2015. Osteoarthritis was ranked seventh in 2005 and sixth in 2015 for its overall global prevalence compared to more than three months of other diseases.

1.6 Cost of disease

The medical cost for treating osteoarthritis is estimated to account for 1% and 2.5% of a country's gross domestic product (1). While hip and knee replacement surgeries represent most of the healthcare costs associated with osteoarthritis, the loss of economic production maybe three times higher than the direct medical costs (27). Data from the UK comparing the cost of hip and knee replacements between 2000 and 2010 found a 66% increase in cost

after adjusting for inflation (28). A more recent study from 2021 found that with the growth of the ageing population relative to the working population, costs associated with age-related diseases such as OA will pose a substantial challenge for the National Health Service (29).

1.7 Diagnosis

The diagnosis of knee osteoarthritis is made using clinical and radiographic features with some variation in definitions between the European League Against Rheumatism (EULAR) and the American College of Rheumatology (30,31). The EULAR evidence-based recommendations include three symptoms (persistent knee pain, limited morning stiffness and reduced function) and three signs (crepitus, restricted movement and bony enlargement) for diagnosis of knee osteoarthritis. The estimated probability of having radiographic knee osteoarthritis increases with the number of positive features to 99% when all six symptoms and signs are present, as shown in Figure 5.

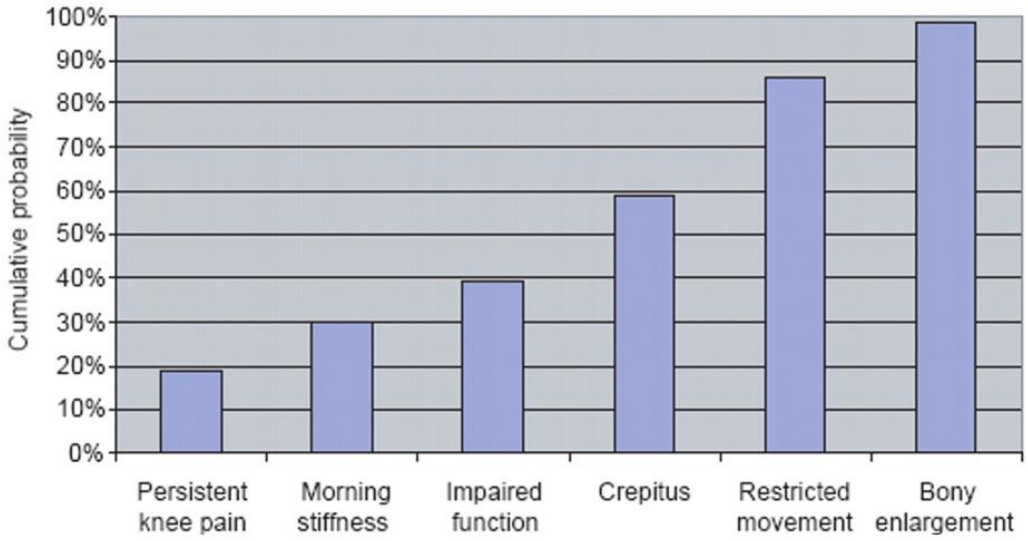


Figure 5: Diagnostic Probability Based on EULAR Criteria for Knee Osteoarthritis
This figure, adapted from Zhang et al. (31), illustrates how the likelihood of diagnosing radiographic knee osteoarthritis (as per Kellgren and Lawrence criteria, grade ≥ 2) escalates with an increasing number of positive features outlined by the European League Against Rheumatism (EULAR). The EULAR criteria encompass three symptoms (persistent knee pain, limited morning stiffness, reduced function) and three signs (crepitus, restricted movement, bony enlargement). Clinical evaluation remains the primary method for diagnosis of knee osteoarthritis, with a 99% probability of having radiographic knee osteoarthritis when all six EULAR criteria are present.

Plain radiographs are only required if other diagnoses are strongly suspected or the presentation is atypical (32). For research studies, radiographs can help differentiate symptomatic osteoarthritis from radiographic osteoarthritis, given that nearly half of patients with radiological features of osteoarthritis have no symptoms and vice versa (33).

The Osteoarthritis Initiative defined radiographic tibiofemoral knee osteoarthritis as tibiofemoral osteophytes (Kellgren and Lawrence ≥ 2) on a fixed flexion radiograph. In contrast, symptomatic tibiofemoral knee osteoarthritis included “pain, aching or stiffness in or around the knee on most days” for at least one month during the past 12 months (34).

1.8 Risk Factors

Age is the most decisive single risk factor for osteoarthritis and likely represents a cumulative exposure to various risk factors and biological changes that occur with ageing (35). Gender is also a decisive risk factor for osteoarthritis; women are both more likely to develop OA and have more severe symptoms (36). The onset of knee pain often indicates underlying osteoarthritis (11). Meta-analysis for risk factors for the onset of knee pain found previous knee injury (odds ratio 2.8), obesity (odds ratio 2.7), and female gender (odds ratio 1.7) were significant risk factors (37).

In knee osteoarthritis, an increasing degree of varus alignment was associated with the development of radiographic knee osteoarthritis, particularly for obese or overweight individuals (38,39). However, alignment was not related to the development of clinical knee osteoarthritis or joint space narrowing (39).

Osteoarthritis has a vital hereditary component; evidence from family and twin studies suggests the genetic element of osteoarthritis is between 40 and 80%, similar to other complex traits (12). In a large twin study, the overall heritability of knee OA surgery was 52%, indicating that half the variation in susceptibility to knee osteoarthritis surgery in the population is explained by genetic factors (12). Gender differences were observed with a higher heritability among women (heritability 80%) than men (heritability 39%).

A brief overview of the genetic risk factors for OA (40) identified linkages to OA on chromosomes 2q, 9q, 11q, and 16p, among others. Genes implicated in association studies include VDR, AGC1, IGF-1, ER alpha, TGF beta, CRTM (cartilage matrix protein), CRTL (cartilage link protein), and collagen II, IX, and XI.

A systematic review across different elite-level impact sports (41) found an increased risk of development of hip OA when compared to matched controls (odds ratio, 1.8-8.7). Similarly, participation in football (odds ratio 3.5), elite long-distance running (odds ratio 3.3) and competitive weight lifting (odds ratio 6.9) were at higher risk of developing knee osteoarthritis, with authors suggesting knee damage may cause the development of OA in later life (41). In the absence of joint injury, no evidence exists that sport contributes to knee injury (42). A cohort study of younger adults (43) showed common injury significantly increases the risk of subsequent knee osteoarthritis (odds ratio 5.7) and hip osteoarthritis (odds ratio 3.5).

The prevalence of osteoarthritis increased in occupations with high exposure to repetitive physical work, such as carpenters and miners (44). An occupational activity study (45) found the risk of developing knee osteoarthritis is increased in individuals whose work involves

more than 30 minutes a day of squatting (odds ratio 6.9), kneeling (odds ratio 3.4), or climbing more than ten flights of stairs (odds ratio 2.7).

1.9 Comorbidity

It has been hypothesised that metabolic disorders may increase the prevalence of OA by activating pro-inflammatory pathways that promote synovial inflammation, cartilage degeneration, and bone matrix remodelling (46). However, a simple association between metabolic disorders and symptoms cannot show the cause-and-effect direction.

Meta-analysis of the association of osteoarthritis with the risk of cardiovascular disease (n = 358944) found an increased risk of cardiovascular disease of 24% in patients with osteoarthritis compared with the general population (47). Additionally, osteoarthritis was associated with increased mortality due to cardiovascular disease (48).

As part of the spectrum of metabolic disorders, the prevalence of diabetes in patients with osteoarthritis is higher than that of the general population (odds ratio 1.41). However, after correction for body mass index, only 7 out of 12 studies found diabetes as an independent risk factor (49). The prevalence of osteoarthritis increases with age, and co-morbidities become very common (50). A meta-analysis of comorbidities in hip and knee osteoarthritis (51) suggests that patients with osteoarthritis with one or more comorbidities have faster progression of pain or joint function.

In a systematic review and meta-analysis of individuals with osteoarthritis (52), the prevalence of any comorbidity was 67% (95% CI 57 - 74%) in individuals with OA versus 56% (95% CI 44 - 68%) in individuals without OA. The prevalence ratio increased from 0.73 (95% CI 0.43–1.25) for one comorbidity to 1.58 (95% CI 1.03–2.42) for two and to 1.94 (95% CI 1.45–2.59) for ≥3 comorbidities. The most significant comorbidities associated with OA were stroke (PR 2.6 [95% CI 2.1 - 3.2]), peptic ulcer (PR 2.4 [95% CI 1.7–3.3]), and metabolic syndrome (PR 1.9 [95% CI 1.2–3.1]).

1.10 Pathology

Osteoarthritis is more than just a cartilage disease, as multiple structures are adversely affected, including subchondral bone, ligaments, synovium, capsule, periarticular muscle and the fat pads (53–56). As a joint disease, osteoarthritis pathogenesis involves metabolic, inflammatory and mechanical factors contributing to the degeneration and eventual failure of the synovial joint. Rather than a wear-and-tear disease, osteoarthritis is a dynamic process comprised of damage and repair mechanisms, with degeneration caused by an imbalance in these activities (57,58).

In normal ageing, articular cartilage remains intact but loses thickness and has a reduced glycosaminoglycan. With osteoarthritis, fraying and splitting of the cartilage surface occurs in focal areas and staining for glycosaminoglycan may be completely absent (59). Glycosaminoglycans resist compressive loading in cartilage (60) and are involved in many biological interactions (61,62). Initially, erosions are superficial but followed by fissures that expose the calcified cartilage zone.

Chondrocytes maintain the cartilage matrix under normal conditions of low turnover (63). Within damaged cartilage, repair mechanisms include the increase in hypertrophic chondrocytes that secrete products in a concentrated effort to maintain and sustain the cartilage. This can destabilise the balance of these products and increase cartilage matrix degrading products and other inflammatory mediators (64). Additionally, deregulated chondrocyte products act on the adjacent synovium and stimulate a pro-inflammatory response (63).

Synoviocytes on the inside lining of the synovium react to the molecules released by the degraded cartilage and produce pro-inflammatory mediators, which increase angiogenesis and induce a phenotypic shift in chondrocytes (65). A cycle follows in which the chondrocyte activity degrades the cartilage and increases synovial inflammation (56).

In osteoarthritis, bone turnover in the subchondral bone increases, and vascular infiltration occurs from the subchondral bone into the cartilage (66). Bone remodelling and repair are associated with the formation of subchondral bone marrow lesions, which are related to local cartilage damage (67). Osteophyte creation at the bone margins can occur as the endochondral ossification process is reactivated (65), which is affected by the loading of the joints and inflammatory factors (68). The pathogenic process is covered in detail in Figure 6.

Osteoarthritis is a heterogeneous disease with diverse risk factors that are thought to have a common end pathway (69). Rather than considering osteoarthritis a specific disease with an unambiguous cause, osteoarthritis may be regarded as a syndrome, a collection of complex symptoms indicating a condition (66). In this context, differing risk factors may trigger different disease pathways, ultimately ending in joint destruction. Understanding these mechanisms and the phenotypes may improve treatments, especially following early diagnosis.

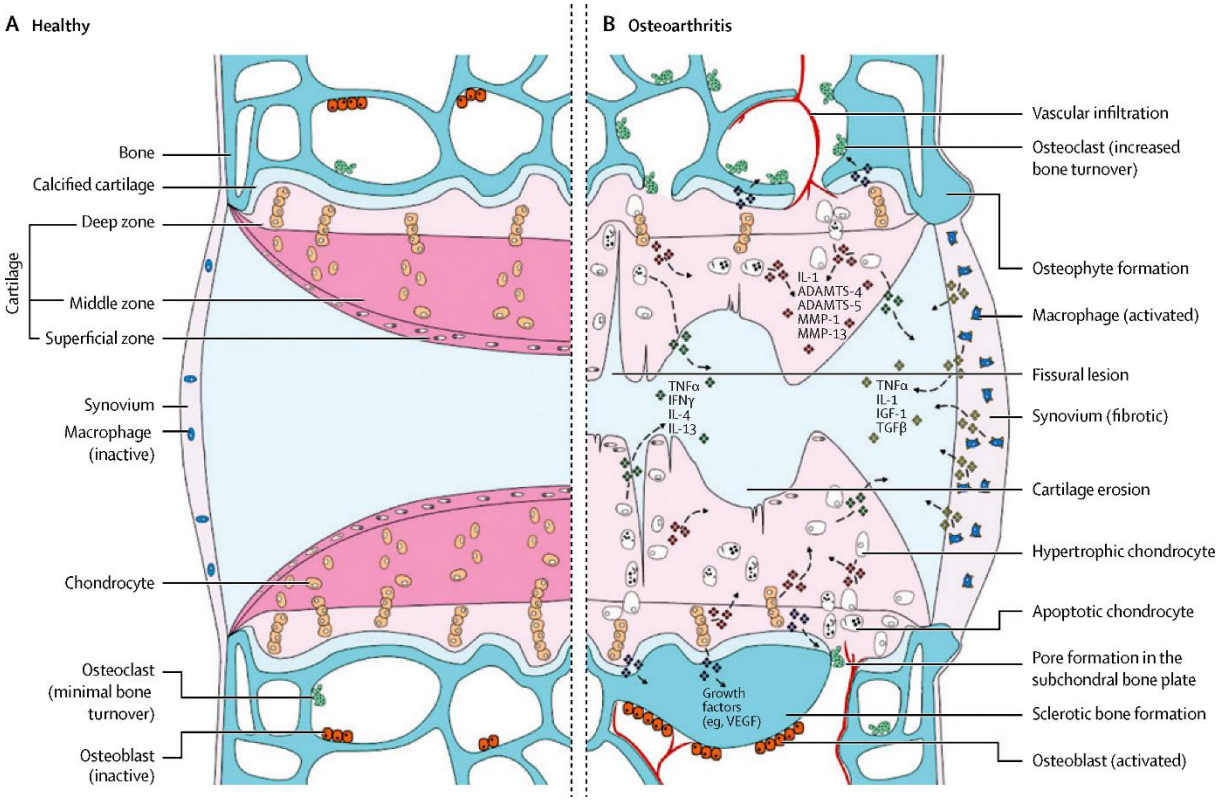


Figure 6: Pathological Evolution in Osteoarthritis Development
This figure presents a comparative analysis between healthy joint tissue (A) and osteoarthritic tissue (B), showcasing the key structural changes and biochemical signalling differences. It highlights the progression of osteoarthritis, beginning with the initial degradation of cartilage and reduction in glycosaminoglycans, to the deeper changes in subchondral bone and the formation of osteophytes. Abbreviations include: ADAMTS = a disintegrin and metalloproteinase with thrombospondin-like motifs. IL = interleukin. MMP = matrix metalloproteinase. TNF = tumour necrosis factor. IFN = interferon. IGF = insulin-like growth factor. TGF = transforming growth factor. VEGF = vascular endothelial growth factor. Reproduced from Glyn-Jones et al. (70).

1.11 Pain mechanisms

Patients with symptomatic osteoarthritis experience pain as the most disabling symptom, which accounts for significant health service use and drives clinical decision-making (12). OA has a range of symptoms; from Figure 5, with decreasing significance for diagnosis, these are persistent knee pain, morning stiffness, impaired function, crepitus, restricted movement and bony enlargement.

Osteoarthritis pain is typically weight-bearing and intermittent but can increase in frequency, duration and severity, leading patients to categorise their pain as unacceptable (71). Symptom flare-ups are a common feature of osteoarthritis pain where patients experience acute increased pain intensity above their baseline pain levels, which may be described as ‘sudden’ or ‘quick’ and lead to increased medication (72).

In their revised definition of pain, the International Association for the Study of Pain notes that pain and nociception are different phenomena. Pain is a subjective and personal experience influenced by biological, psychological and social factors. In contrast to pain, nociception refers to activity that occurs in the nervous system in response to a noxious stimulus (73).

Knee OA nociception can arise from the stimulation of peripheral nerve endings (74). In non-innervated tissues such as the cartilage, in most cases, damage does not cause nociception (75). Instead, increased vascular penetration and nerve growth in these tissues are sources of nociception in osteoarthritis (76).

Inflammation and ongoing tissue injury in the joint can increase the responsiveness of peripheral nerve endings, leading to peripheral sensitisation (77). Increased nociceptive inputs from the joint lead to complex changes in the central nervous system, known as central sensitisation (78). Peripheral sensitisation can contribute to developing and maintaining central sensitisation (79). In centrally sensitised pain, nociceptive neurons are hyperexcitable and amplify the nociceptive processing (80).

Pain may also arise from damage beyond the specific nociceptors. The International Association for the Study of Pain define neuropathic pain as "pain arising as a direct consequence of a lesion or disease affecting the somatosensory nervous system" (81). Rather than arising from tissue-damaging or potentially tissue-damaging stimuli, neuropathic pain is produced by damage to the peripheral or central nervous system (82).

A meta-analysis of 15 studies examined using quantitative sensory testing (QST) to measure pain thresholds in individuals with knee osteoarthritis with high and low symptom severity and individuals without knee osteoarthritis (83). This testing involves the assessment of sensitivity to noxious or innocuous stimuli using standardised mechanical, thermal and electrical test modalities (83). The analysis found a lower pain pressure threshold between knee osteoarthritis participants with high symptom severity versus those with low symptom severity (corrected mean effect size 0.51). A meta-analysis of 9 studies examined the use of self-reported questionnaires to determine neuropathic pain and estimated a 23% prevalence of neuropathic pain in patients with knee and hip osteoarthritis (84). The meta-analysis used a range of definitions for neuropathic pain, including PainDETECT, modified painDETECT, Leeds Assessment of Neuropathic Signs and Symptoms and Douleur Neuropathique (84).

In patients with chronic symptomatic knee osteoarthritis, signs of central sensitisation are common (46% prevalence) and more likely to occur in individuals with neuropathic-like symptoms (85). This study used quantitative sensory testing to define central sensitisation based on hyperalgesia to mechanical stimuli (pinprick or pressure), a shift in the stimulus-response curve, mechanical allodynia and enhanced temporal summation.

Pain pressure thresholds measured by quantitative sensory testing are associated with knee osteoarthritis pain (odds ratio 1.7 – 2.0) but not knee osteoarthritis duration or radiographic severity (86). Inflammation assessed as synovitis and effusion in knee osteoarthritis is associated with peripheral pain sensitisation; in contrast, bone marrow lesions do not

appear to contribute to pain sensitisation (87). Knee radiographs have been used in clinical studies, but nearly half of patients with radiological features of osteoarthritis have no symptoms and vice versa (33).

1.12 The search for disease-modifying therapeutics

Cartilage repair aims to improve the loss of cartilage structure and function. Several key anabolic and catabolic pathway enzymes are dysregulated in cartilage due to osteoarthritis, providing targets for new interventions (88). One proposed disease-modifying osteoarthritis drug is sprifermin, an anabolic protein truncated version of fibroblast growth factor 18 that induces cartilage matrix production and chondrocyte proliferation (89). Sprifermin has been associated with a dose-dependent reduction in femorotibial cartilage loss assessed at 12-month follow-up (90) and is the subject of several randomised control trials (91,92). Bone morphogenic proteins have also been proposed to promote cartilage regeneration (92–94) and have demonstrated osseous repair in numerous animal models (95). A particular challenge with cartilage repair therapeutics is evaluating the clinical importance of outcomes and the durability of any induced changes, given that cartilage damage is typical in asymptomatic osteoarthritis (33).

Subchondral bone plays a role in nutrient supply, shock absorption and support of overlying cartilage (96,97). Pre-clinical and clinical studies suggest that changes in subchondral bone may precede cartilage loss and may be a trigger for cartilage degeneration (98,99). Additionally, bone marrow lesions are associated with osteoarthritis pain (100). Bisphosphonates are a potential candidate therapy, as they slow bone turnover osteoclast activity in osteoporosis. However, their suitability for osteoarthritis remains uncertain (101). While they do not appear to defer radiographic progression or alleviate symptoms, bisphosphonates are potentially valuable for patient-specific subgroups such as those with high bone turnover (98). Several other drugs targeting bone cells are currently under investigation for treating osteoarthritis based on differing degenerative pathways (101–103).

Treatment of pain using paracetamol, NSAIDs and opioids has limited benefits and associated risks that make pain control challenging in osteoarthritis (104) (105). Alternative attempts include directly targeting neuronal structures to reduce pain. The use of anti-nerve growth factor antibodies has been studied. Still, clinical trials were halted due to concerns about faster OA progression and increased joint replacement surgeries (89). Another strategy involved the inhibiting nerve growth factor receptors, which has been successful in pre-clinical animal models (106,107) and is currently under investigation in early clinical trials (108–110).

1.13 Image Features

Given the common pathophysiology of all forms of osteoarthritis, many parallels exist in the radiographic features of the disease across different sites. However, the examination of the imaging features is contrained to knee osteoarthritis, given the focus on knee OA in this thesis.

Plain radiographs are not the primary method for clinical diagnosis of knee OA (Section 1.7); they are only required if another diagnosis is strongly suspected or the presentation is atypical (32). For research studies, it can be helpful to differentiate symptomatic osteoarthritis from radiographic osteoarthritis using radiographs(33). While there are several radiographic techniques for assessment of the osteoarthritic knee (111), all involve a standing stance to enable quantification of the joint space under weight-bearing conditions. A systematic literature review from EULAR recommends weight-bearing and patellofemoral views for the most reproducible measurement of OA features (32).

X-ray remains the most accessible imaging tool in evaluating the osteoarthritic joint (112). Radiographs are commonly used to report osteophyte formation and joint space narrowing, with the Kellgren-Lawrence grading scheme often used in research studies (113). The Kellgren-Lawrence scores vary from grade 0 (none) to grade 4 (severe) based on joint space narrowing, osteophytic changes, sclerosis and bone end deformity (114), as shown in Figure 7.

Magnetic resonance imaging provides excellent tissue contrast, capturing a wide array of pathological changes in osteoarthritis, although at increased cost and image acquisition time (112). Currently, the diagnosis of knee osteoarthritis is made using only clinical and radiographic features and does not include MRI (30,31).

Several semi-quantitative scoring systems exist for evaluating knee osteoarthritis, including the Whole-Organ Magnetic Resonance Imaging Score (116), the Boston Leeds Osteoarthritis Knee Score (117) and more recently the Magnetic Resonance Imaging Osteoarthritis Knee Score (MOAKS) (118). These scoring systems were developed for research to standardise the assessment of structural endpoints for clinical trials. These scoring systems included semi-quantitative cartilage evaluation signal and morphology, subarticular bone marrow abnormality, subarticular cysts, subarticular bone attrition, marginal osteophytes, medial and lateral meniscal integrity, anterior and posterior cruciate ligament integrity, medial and lateral collateral ligament integrity, synovitis, loose bodies and periarticular cysts/bursae.

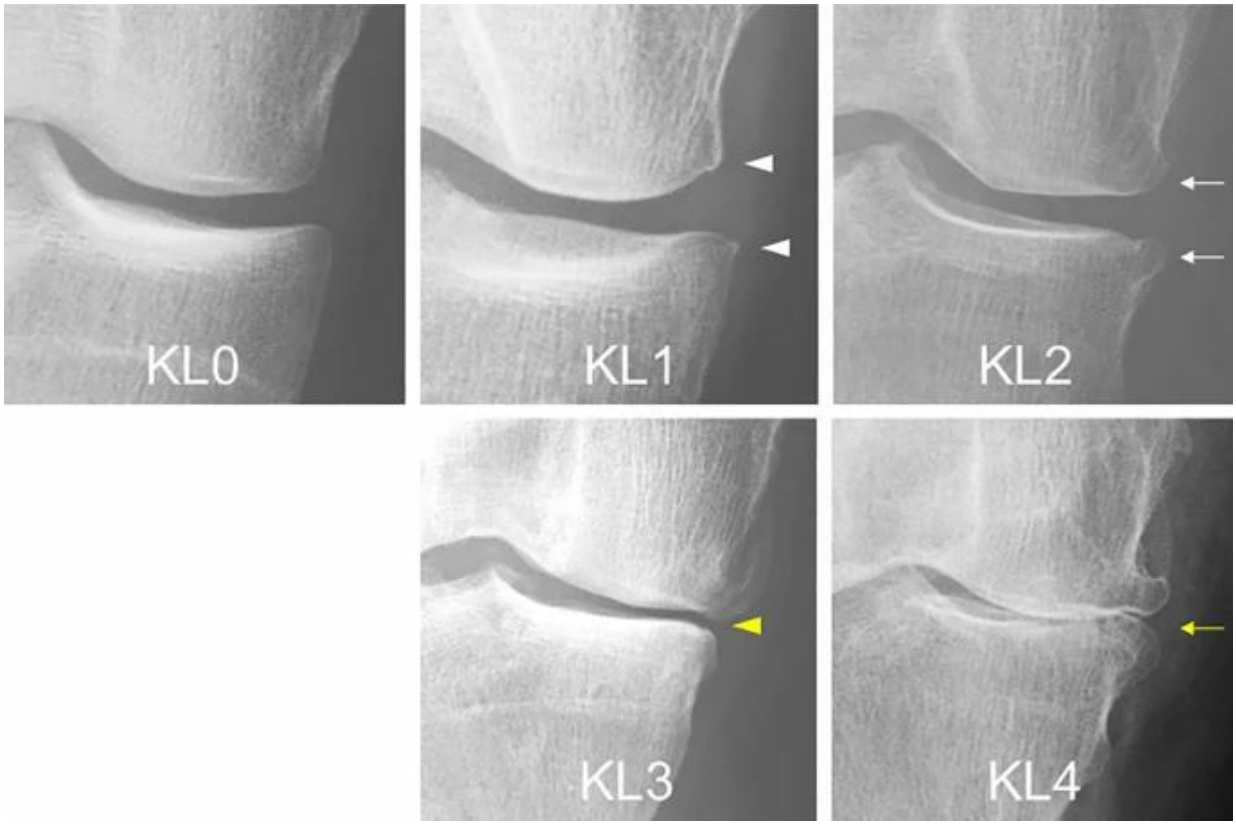


Figure 7. Progressive Stages of Knee Osteoarthritis in Kellgren-Lawrence Grading
This figure is reproduced from Sekyia et al. (115). Anterior-posterior radiographs show varying degrees of knee osteoarthritis as classified by the Kellgren-Lawrence grading system (113). The radiographs demonstrate a progression from grade 0 (no osteoarthritis) to grade 4 (severe osteoarthritis), illustrating key features such as joint space narrowing and bone end deformities. Radiographic grading provides a standardized assessment of joint damage and is primarily used in research as x-ray is usually only required when diagnosis is uncertain, for example to rule out rheumatoid arthritis.

1.13.1 Articular cartilage

The articular cartilage provides a smooth, lubricated surface for articulation and facilitates the transmission of loads across the joint (119). The articular cartilage in the knee comprises the patella, femoral, medial tibial cartilage and lateral tibial cartilage. MRI can evaluate cartilage integrity from minor focal defects to full-thickness cartilage loss (116).

Several techniques for imaging cartilage include proton density, T2-weighted, spoiled gradient echo and fast spin echo MRI (120). Detection of knee cartilage defects can be improved by using a three-dimensional MRI at 3.0T with multiplanar reformatting (120). Meta-analysis comparing the diagnostic performance of three-dimensional MRI for detecting cartilage defects found that 3.0T MRI instead of 1.5T MRI improved the diagnostic odds ratio and performed better than multiple 2D acquisitions (121).

Dual echo steady state (DESS) imaging has proven helpful for cartilage morphology imaging (122) acquiring two gradient echoes separated by a refocusing pulse and combining both as

a single image. However, early assessment of the DESS technique showed it did not improve performance compared to conventional proton density-weighted MRI (123). Three-dimensional fast spin echoes have demonstrated improved signal and contrast compared to other 3D sequences for imaging cartilage lesions (124) but show similar diagnostic performance for detecting ligament tears, meniscal tears, and bone marrow oedematous lesions within the knee joint (125).

1.13.2 Bone marrow lesions and cysts

Bone marrow lesions have been shown to progress the structural deterioration in knee osteoarthritis and contribute to the variance of malalignment (126,127). Bone marrow lesions on MRI are strongly associated with pain in knee osteoarthritis, with more extensive lesions being present almost exclusively in persons with knee pain (35.9% vs 2%) (100). Bone marrow lesions may contain a subchondral cystic component, which has been found to influence the rate of cartilage loss in knee osteoarthritis (117). Bone marrow lesions are characterised as ill-defined hyperintensities on T2-weighted, fat-suppressed MRI (118), an example shown in Figure 8.

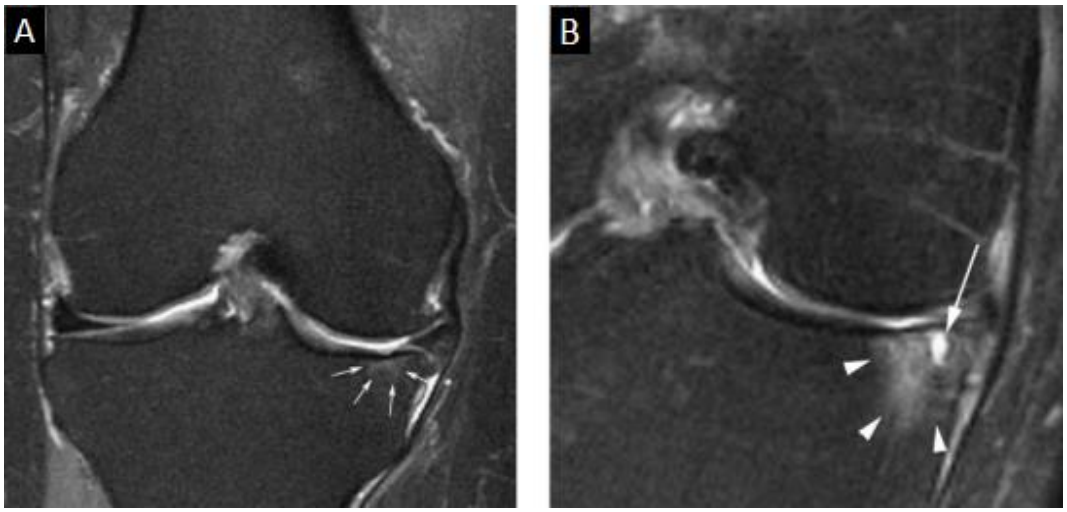


Figure 8: Bone marrow lesions

Subchondral bone marrow lesions are shown on a coronal T2-weighted, fat-suppressed MRI. Two types of lesions are depicted: (A) A small bone marrow lesion located in the central subregion of the medial tibia, indicative of early-stage structural deterioration. Such lesions are characterized by their subtle appearance and are often the initial sign of underlying joint pathology. (B) A medium-sized bone marrow lesion, more advanced in nature, exhibiting a combination of ill-defined hyperintense areas (marked by arrowheads) and a distinct subchondral cyst (indicated by an arrow). These lesions, as visualized on MRI, are essential for assessing the severity and progression of knee osteoarthritis, aiding in the formulation of targeted treatment strategies.

1.13.3 Osteophytes

Osteophytes most commonly appear at the margin of the joint and are osteocartilaginous protrusions, starting as outgrowths of cartilage and undergoing endochondral ossification (128). Osteophytes have been strongly associated with the progression of osteoarthritis and subchondral sclerosis; they are also associated with knee pain (129,130). Osteophytes can be viewed on T1-weighed, spoiled gradient echo and double echo steady state MRI (118) example is shown in Figure 9.

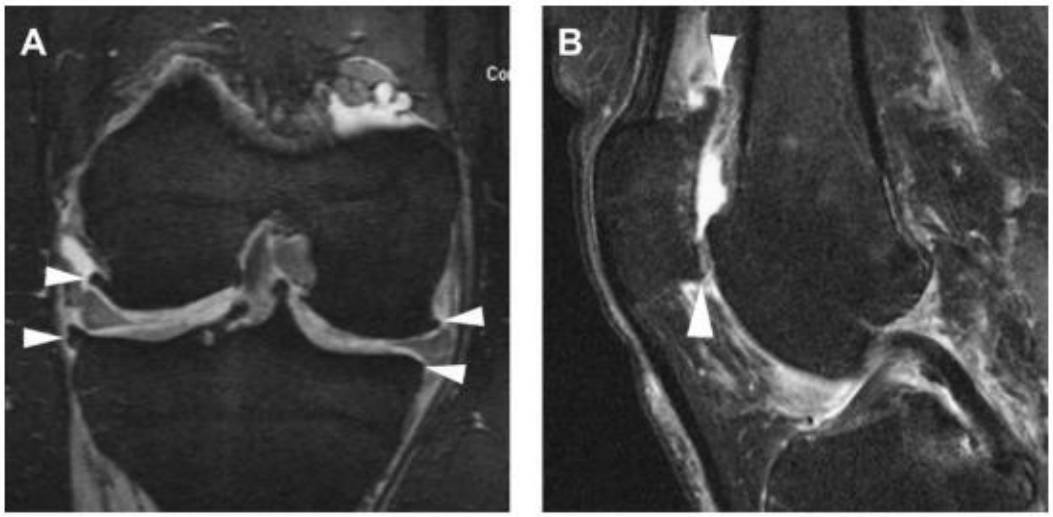


Figure 9: Magnetic Resonance Imaging of Osteophytes in Knee Osteoarthritis
The appearance of knee osteophytes is shown on a double echo steady state MRI, highlighting their common locations and morphological characteristics in the context of osteoarthritis. (A) The coronal plane image vividly displays osteophytes at the joint margins. These protrusions, characteristic of osteoarthritis, typically originate as cartilage outgrowths and undergo a process of endochondral ossification. (B) In the sagittal plane, osteophytes are observable at the superior and inferior poles of the patella. These protrusions are indicative of joint degeneration and are associated with the progression of osteoarthritis.

1.13.4 Hoffa synovitis

Hoffa synovitis is an inflammation of the infrapatellar fat pad and is often present in osteoarthritis; in some studies, this marker correlates with pain, although the literature is conflicting (131,132). It has been proposed as a source of knee pain in patients suffering from OA and from patellofemoral pain, which is often seen as a precursor to OA (132,133). It is well-reported in the literature that Hoffa's synovitis is a precursor to structural changes in OA (134–138). Hoffa synovitis is typically visualised in a mid-sagittal plane on either a T1-weighted, fat-saturated T2-weighted or proton-density weighted MRI (118,139) and an example is shown in Figure 10.

The best results were obtained using contrast, enabling the accurate differentiation of Hoffa synovitis from synovial membrane inflammation (140). Fat-saturated T2-weighted images do not distinguish inflamed synovium from fatty tissue. Giving an injection of gadolinium contrast agent into a vein, then imaging after a few minutes, allows for good characterization of synovium given the difference in perfusion between the synovium and the fatty tissue.

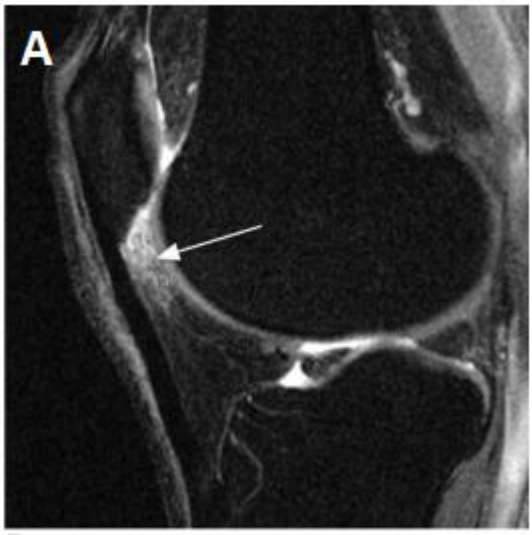


Figure 10: Magnetic Resonance Imaging of Hoffa Synovitis in knee Osteoarthritis
Mid-sagittal fat-suppressed proton density MRI scan reproduced from Romer et al. (141). The MRI shows Hoffa synovitis, a significant inflammatory condition associated with knee osteoarthritis. The image reveals oedema at the superior aspect of the infrapatellar fat pad, a key location of Hoffa synovitis association with patellofemoral impingement. This condition has been proposed as a source of knee pain not only in patients with osteoarthritis but also in those suffering from patellofemoral pain, which is frequently considered a precursor to osteoarthritis.

1.13.5 Synovial effusion

T2-weighed and proton density-weighted MRIs are typically used to visualise changes to the synovial membrane and joint cavity (118). Without gadolinium contrast, it is impossible to differentiate synovitis (synovial inflammation) from effusion (fluid equivalent signal within the joint cavity). Therefore, non-contrast imaging measures should be referred to as effusion-synovitis (141). Delayed contrast and fat-saturated T1 weighted imaging enable assessment of the synovial membrane, as shown in Figure 11. A semi-quantitative evaluation of the synovitis on contrast-enhanced images has been proposed, with worsening synovitis associated with increased pain in knee osteoarthritis (142).

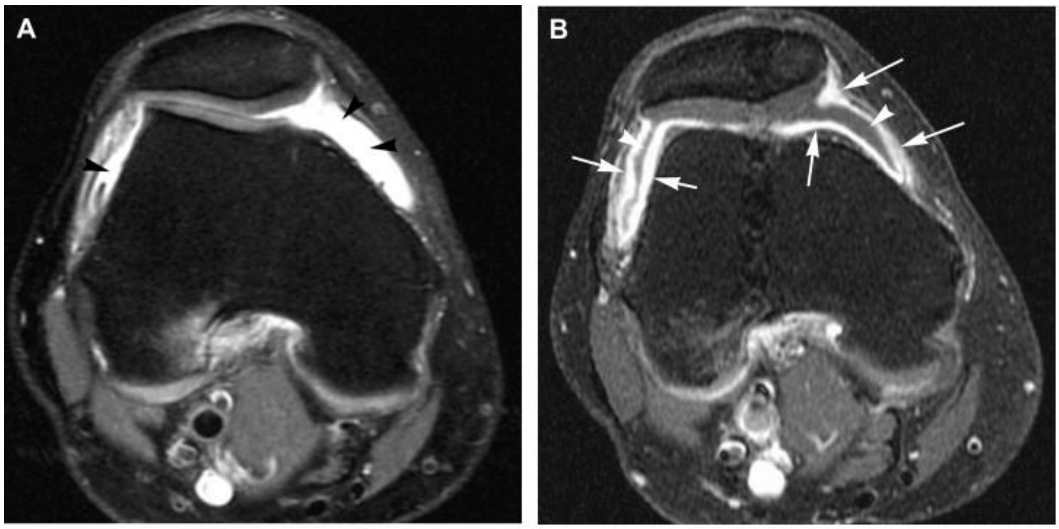


Figure 11: Assessment of Synovitis and Synovial-Effusion
Images reproduced from Romer et al. (141). Assessment of non-contrast-enhanced MRI (A) shows effusion-synovitis; without contrast, the synovial fluid and synovial membrane cannot be differentiated. 5-minutes after venous injection of gadolinium contrast fat-saturated T1-weighted MRI (B) shows synovial inflammation, depicted as the hyperintense lining of the synovial membrane. Synovitis is associated with knee pain and inflammatory changes in the joint.

1.13.6 Ligaments

Most ligament ruptures occur in the anterior cruciate ligament and are more common in knees with symptomatic knee osteoarthritis than those without osteoarthritis (143). While anterior cruciate ligament injury is associated with a previous knee injury, only 48% of individuals with complete anterior cruciate ligament tears reported a prior knee injury. Anterior cruciate ligament injury is a vital precursor for osteoarthritis (144); at 10 to 20 years after the diagnosis, on average, 50% of those with a diagnosed anterior cruciate ligament have osteoarthritis with associated pain and functional impairment.

Ligament injury is not associated with increased knee pain in patients with osteoarthritis (143). The anterior cruciate ligament can be assessed in axial, coronal and sagittal proton density imaging (with or without fat suppression) and fat-saturated T2-weighted imaging (118,145).

1.13.7 Meniscus

The meniscus provides several vital functions in the knee, including load bearing by enlarging the femur-tibia contact area, sock absorption and stability enhancement (146). Asymptomatic meniscal tears are common in middle-aged or older patients. They may lead to knee osteoarthritis, but knee osteoarthritis can also cause meniscal tears through the breakdown and weakening of the meniscal structure (147). Menisci can be imaged using T1-weighted, T2 weighted or fat-saturated proton density MRI and are typically assessed in coronal and sagittal planes; an example is shown in Figure 12. The three-dimensional fast spin echo sequence used in articular cartilage imaging can also image the menisci without reducing diagnostic performance (148).

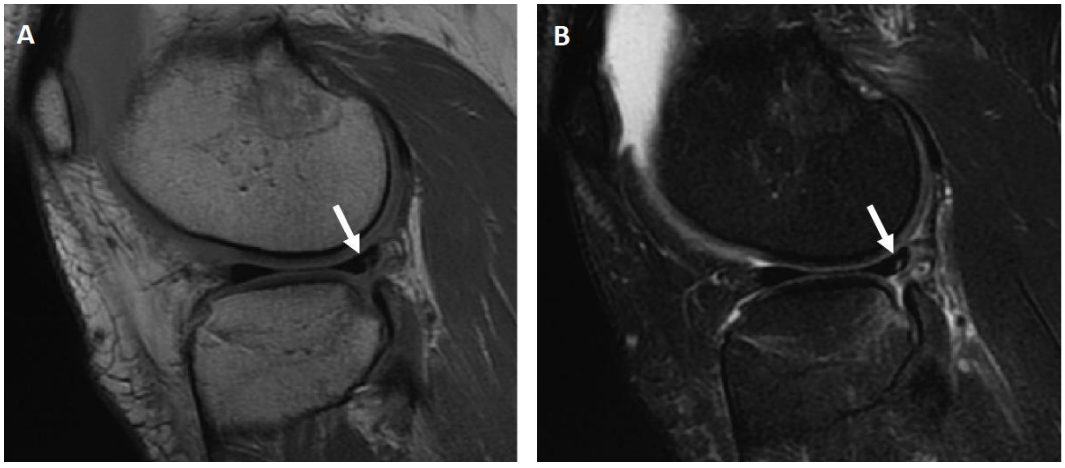


Figure 12: Magnetic Resonance Imaging of the Meniscus
Images are reproduced from Rosas et al. (80). The meniscus plays several crucial roles in knee mechanics, including load distribution by increasing the femur-tibia contact area, shock absorption, and enhancing joint stability (143). While asymptomatic meniscal tears are frequently observed in middle-aged or older individuals, these tears can be both a precursor to, and a consequence of, knee osteoarthritis. In this case a partial tear is shown on (A) a sagittal proton density MRI and (B) a sagittal fat-suppressed proton density T2-weighted MRI.

1.14 Disease management in knee osteoarthritis

1.14.1 Pain Medication

In a systematic review of osteoarthritis management guidelines (149) included paracetamol (first line) and NSAIDs (topical or oral, second line). A 2017 meta-analysis showed that overall, topical NSAIDs were superior to placebo for relieving pain (corrected mean effect -0.30) and joint function (fixed mean effect -0.35) in osteoarthritis (150). However, NSAIDs have been associated with a clinically relevant 50-100% increased risk of myocardium infarction or cardiovascular death compared with placebo (104). Another meta-analysis from 2017 (n = 58451) found that single-agent paracetamol was below the minimum clinically significant effect for pain reduction, given the effect size compared to placebo was less than 0.2 (104).

Duloxetine, a reuptake inhibitor of serotonin and norepinephrine, is effective for the treatment of depressive disorders, diabetic neuropathic pain, anxiety disorders and fibromyalgia (104). Meta-analysis of three randomised controlled trials (n = 1011) showed significant differences between patients taking duloxetine and those taking placebo, with a mean difference of -0.88 (151). While there was an increase in adverse events (relative risk = 2.2) in patients taking duloxetine compared to placebo, differences in serious adverse events were not statistically significant.

Tramadol, a synthetic, centrally acting analgesic agent and weak opioid, produces symptom relief and improves function when taken alone or combined with paracetamol, but these benefits are small (105). Given the limited benefits and associated risks of addiction and overdose, recommendations typically discourage the use of prescribing opioids in osteoarthritis (93).

Intra-articular corticosteroids were generally recommended for hip and knee OA in cases where oral or topical analgesics prove ineffective (149). A 2015 Cochrane review (41) identified 13 trials with ‘low’ levels of evidence for all outcomes and found a mean effect size of -0.41 at six weeks after inter-articular corticosteroid injection. This corresponds to a difference in pain scores of 1.0cm on a 10cm visual analogue scale. Additionally, while only of borderline significance (p = 0.08), mid-term pain was reduced in patients with inflammatory signs compared to those without. Meta-analysis using individual patient data from seven trials (n = 620) found patients with severe baseline pain have significantly more significant reductions in short-term pain compared to those with less severe pain at baseline when treated with intra-articular corticosteroids.

A systematic review (152) of the use of injectable hyaluronic acid in knee osteoarthritis showed less than half of the minimally significant difference for pain stiffness and function compared to placebo (saline). A follow-up meta-regression analysis in 2016 found that publication bias and the risk of selective outcome reporting suggest only minor clinical effects of hyaluronic acid compared to saline (153).

1.14.2 Non-pharmacological treatment

Exercise, self-management and weight loss if overweight or obese are recommended as first-line treatments (154). Network meta-analysis into the effects of exercise in hip and knee osteoarthritis compared to no exercises showed clinically meaningful benefits of exercise (155); strengthening, flexibility, aerobic and aquatic exercises were more effective than no exercise, with combined exercises having the most significant effects. The analysis found that the overall best exercise intervention for pain in lower limb osteoarthritis was aquatic strengthening plus aerobic flexibility exercise, with a reduced visual analogue scale of -1.7cm compared to no exercise control. In obese patients with knee osteoarthritis, a meta-analysis showed decreased pain and disability (standardised mean 0.20 and 0.23, respectively) at a weight reduction of 6kg (156).

Self-management programmes are formalised training, which includes medication compliance, pain management, coping strategies and exercise advice. There is evidence that self-management programmes reduce pain and improve function in knee and hip osteoarthritis (157–160). Self-management education programmes encourage people with chronic disease to take an active role in managing their condition. These programmes aim to improve patient outcomes by supporting, not replacing, medical care.

A Cochrane review of self-management education programmes for osteoarthritis (161) found self-management programmes may slightly improve pain, function and symptoms, although these benefits are of unlikely clinical importance. The review found these programmes were of similar effectiveness to attention control training. Rather than focus on patient education, attention management teaches techniques such as attention diversion, imagery and mindfulness exercises (161).

Despite several high-quality studies (162,163), there was no evidence for using lateral heel wedges to relieve pain or improve function in medical compartment knee osteoarthritis. Knee braces for patients with painful patellofemoral knee osteoarthritis can reduce visual analogue pain by -1.3cm (10cm scale) (162), which was below the minimal clinically significant improvement of pain, which is reported to be -1.5 for pain measured on the WOMAC Pain scale for the knee (164). For patients (121) with symptomatic medial knee osteoarthritis, the reduction in visual analogue pain was -2.6cm when using a knee brace (165). Using walking aids in patients with knee osteoarthritis reduced visible analogue pain by -2.1cm; however, 35% of participants refused to participate in the study because they did not want to use a cane. As studies into knee bracing and walking aids cannot control for placebo effects, the clinical significance of these effects remains controversial (166).

A Cochrane review looking at sham-controlled studies for the use of acupuncture in osteoarthritis (n=3498) showed minor improvements in pain (0.9 points on a 20-point scale) and function (2.7 points on a 68-point scale). However, these did not meet thresholds for clinical relevance, with much of the effect size being explained by inadequate sham controls (167). Although not currently recommended in primary knee OA treatment guidelines, a meta-analysis of low-level laser therapy (168) demonstrates a reduction in pain of -3.2cm (at 2 – 4 weeks post-treatment) measured on the visual analogue scale, compared to

placebo. The heterogeneity of reported methods and limited access to laser treatment should be addressed to improve the use of laser treatment for pain control (169).

1.14.3 Joint replacement

Joint replacement is recognised as a clinically relevant treatment for end-stage osteoarthritis, which substantially improves the quality of life and is highly cost-effective (170). A population-based cohort study from UK general practice estimated the lifetime risk of undergoing total joint replacement to be between 5% and10%, substantially less than the risk of developing symptomatic knee or hip osteoarthritis (171), with lifetime risks as high as 45% for developing knee osteoarthritis (172,173) and 25% for developing hip osteoarthritis (172). The mortality-adjusted lifetime risk at the age of 50 for total knee replacement was 10.8% for women and 8.1% for men; for hip replacement, it was 11.6% for women and 7.1% for men. Referrals for joint replacement surgery should be made before there are prolonged and established functional limitations and severe pain in patients with osteoarthritis who experience joint symptoms that have a substantial impact on their quality of life and are refractory to non-surgical treatment (112).

Continued pain after total joint replacement remains an issue, even in well-performed surgeries. Two years after total hip replacement, moderate to severe pain was reported by 8.1% of individuals, with a higher proportion of reported mild to severe pain in patients with higher BMI and depression (175). The range of values reported for unfavourable long-term pain outcomes after knee replacement varies, with about 10% to 34% of individuals reporting postsurgical pain (176,177). In the best-quality studies, a lower proportion (20%) of patients reported unfavourable pain outcomes after knee replacement (177). Persistent pain was most commonly described as aching, tender, and tiring, with only 6% of patients reporting pain that was neuropathic.

Most patients receive a total knee replacement, but approximately 8% of cases receive a partial knee replacement of the medial, lateral or patellofemoral compartment (178). The benefits of a partial knee replacement include a lower postoperative length of stay, fewer surgical complications, lower re-admission rates and higher cost-effectiveness (179). These benefits must be weighed against higher revision rates than total joint replacement reported in nearly all national registries (178).

Complications associated with joint replacement surgery include those familiar with other surgical procedures, such as infection at the surgical site and venous thromboembolism, and those specific to involving joints, such as damage to the neurovascular systems, infection of the prosthesis, fracture of the bone around the artificial joint, and eventual wear of the prosthesis requiring a revision (180).

1.14.4 Knee osteotomy

A Cochrane review concluded that Knee osteotomy reduces pain and improves knee function in patients with medial compartment osteoarthritis of the knee (181). A 2018 Meta-Analysis compared Knee osteotomy to unicompartmental knee replacement and found reduced post-operative pain, lower revision rates and fewer post-operative complications in patients undergoing unicompartmental knee replacement (182). This analysis also identified a superior range of motion in individuals experiencing knee osteotomy compared to unicompartmental knee replacement, suggesting knee osteotomy may be suitable for patients with high activity requirements.

1.14.5 Arthroscopic knee surgery

The American Academy of Orthopaedic Surgeons and the British Medical Journal make strong recommendations against the use of arthroscopic surgery in knee osteoarthritis (183,184). In patients with a degenerative medial meniscus tear, knee arthroscopy shows no improvement over exercise therapy (185). A previous meta-analysis of 7 randomised controlled trials (n = 805) found that arthroscopic surgery did not achieve a minimally significant difference for short or long-term knee pain (186). Arthroscopic knee surgery remains the most common elective orthopaedic procedure despite a lack of evidence showing substantial benefits compared to the potential harms, which includes a shortening of the time to joint replacement and serious adverse events in 0.56% of individuals (187–189).

1.15 Unmet needs

Given the lack of effective treatments for osteoarthritis, the Osteoarthritis Research Society International white paper submitted to the FDA for consideration of osteoarthritis as a severe disease (104) argues the urgent need for clinical studies of new and existing agents that might intervene in the pathophysiology and progression of the disease. The white paper argues the lack of valid and responsive structural endpoints for clinical trials has constrained the development of new therapeutic agents.

The heterogeneous disease progression of osteoarthritis provides additional challenges for the design of clinical trials. Some individuals will have stable disease over many years, while others experience rapidly increasing pain, disability or structural damage in a relatively short timeframe (190). The underlying mechanisms and risk factors for progression are of interest for better disease prognosis. Phenotyping of OA may improve the stratification of participants before enrolment in clinical trials, thereby improving the efficiency of these trials and preventing too many negative trials. Developing a clinical decision tool to predict progression may help build more tailored support for individuals.

Within the pathophysiology of osteoarthritis, research advances have focused on mechanisms of cartilage breakdown and chondrocyte biology. While attempts have been made to incorporate whole joint disease markers in clinical trials, these remain primarily

descriptive (116,191). A better understanding of the interactions between cartilage and the adjacent bone (forming of bone marrow lesions) and between the synovium and adipose tissue (exchange of pro and anti-inflammatory markers) could lead to new interventions (192). Further research is needed to understand the relationship between pain and structure, mainly how pain modifies osteoarthritis's disease course and relates to disease-specific mechanisms (193).

Current treatments for osteoarthritis symptoms often provide inadequate pain relief and leave many individuals with osteoarthritis dissatisfied (194). Current pharmacological therapies come with significant toxicities that limit their widespread usage (195) highlighting the need for effective non-pharmacological interventions. Treatments targeted to specific pain mechanisms may improve the management of osteoarthritis; this may be facilitated by identifying novel pain targets (196). For example, in patients where progression is driven by synovial inflammation, treatment aimed at this inflammation (e.g. methotrexate) should be evaluated, whereas in cases caused by increased bone turnover, treatment aimed at reducing this increased bone turnover (e.g. bisphosphonates) should be investigated (33).

1.16 Potential added value of advanced MRI.

1.16.1 Imaging

Magnetic resonance imaging has proven to be a vital tool in improving our understanding of osteoarthritis as it can visualise all potentially relevant osteoarthritis joint structures. This contrasts with plain radiography, where features such as effusion, meniscal tears, subchondral bone architectural changes and cartilage loss cannot be seen (117).

Compared to other modalities, increased cost and long imaging times have limited the widespread use of MRI for osteoarthritis in a clinical setting (197). For example, the research imaging protocol for the osteoarthritis initiative requires 75 minutes to image both knees (198). Advances in MR hardware and the increasing use of 3.0T MRI for musculoskeletal imaging have led to the creation of high-resolution three-dimensional 10-minute sequences and 5-minute rapid two-dimensional sequences for evaluation of the knee (125,199,200).

Compressed sensing is a reconstruction technique with the potential to accelerate MRI via accurate reconstructions of sparsely sampled data (201). Acceleration factors reported in the literature for knee cartilage evaluation can reach 8 to 10-fold when combined with parallel imaging (202,203). The acceleration of sequences using compressed sensing has enabled a significant 30 - 57% reduction in scan time compared to conventional clinical MRI (204,205).

Further gains may be realised, but optimisation of the compressed sensing reconstruction remains challenging, often requiring multiple iterations of image reconstruction, leading to long reconstruction times (206). Advances such as variation networks have proved popular for accelerating compressed sensing algorithms. These have demonstrated high computational performance, i.e., reconstruction time of 193 ms on a single graphics card (207). This method has been adopted by SIEMENS for their most recent AI based reconstruction, deep resolve boost (208).

Compositional MRI techniques have been developed to evaluate cartilage matrix quality to detect early changes at a stage where the osteoarthritis is potentially reversible (209). Assessment of these compositional markers could accelerate the development of new interventions aimed at halting progression. Compositional techniques include T2 mapping, T1 mapping in the rotating frame (T1ρ), delayed gadolinium-enhanced MRI of cartilage (dGEMRIC), sodium imaging and glycosaminoglycan chemical exchange saturation transfer imaging (gagCEST). The use of contrast and hardware limitations has meant that most studies have used T2 and T1ρ mapping for compositional assessment of cartilage.

In previous in vitro studies, attempts to correlate T2 relaxation times with biochemical changes in cartilage reported poor correlation with PG content (210,211). Instead, T2 is mainly affected by collagen orientation and quantity, and the water content (212). The lack of specificity to quantify proteoglycan loss and sensitivity of T2 maps to knee orientation in the external magnetic field impose some limits on this quantitative technique. However, T2 Mapping can still be used for cartilage assessment, especially in patients with radiographic diseases, as shown in Figure 13.

Traditionally, T2 mapping is performed via a fast-spin-echo where multiple echoes are collected every TR. For high-resolution MRI needed to visualise cartilage, the repeated refocusing pulses can increase scan duration due to limitations on power deposition (213). Novel techniques such as magnetic resonance fingerprinting and synthetic knee MRI have reduced image sequence by as much as 6-fold (214).

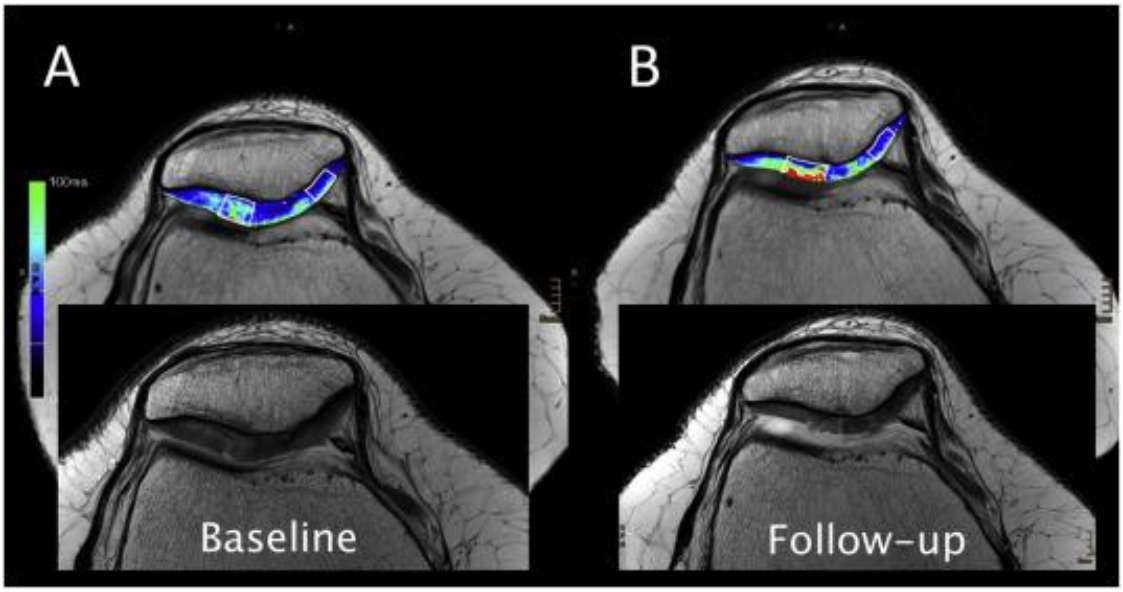


Figure 13: High-Resolution T2 Mapping of Knee Cartilage
True fast imaging with steady-state precession (True-FISP) sequence was used for high-resolution knee T2-mapping with a 0.5mm isotropic resolution acquired in 6:21 min. A small patellar-femoral cartilage lesion at baseline (A) is shown to increase in size at follow-up after 3.5 years (B). The T2 map shows a significant increase in T2 value across the lesion, whereas values from healthy cartilage remain stable. T2 mapping enables the assessment of cartilage

morphometry and longitudinal evaluation of cartilage degeneration in subjects with established disease.

T1ρ imaging is a “spin locked” technique where the magnetisation experiences a constant RF field followed by conventional image readout (215). T1ρ measures the spin-lattice relaxation in the rotating frame and changes in the cartilage extracellular matrix, specifically proteoglycan content (120,216). In cartilage, T2 imaging is also sensitive to proteoglycan content and the water fraction, providing information about early changes in osteoarthritis (120).

1.16.2 Artificial Intelligence

Healthcare systems now routinely generate data in massive quantities from sources such as 3D medical imaging, genomics, and electronic health records. This ‘big data’ far surpasses the ability of humans alone to interpret, leading to an increased reliance on machines. By leveraging the ability of computers to perform complex calculations far beyond the capacity of humans, it is argued that provided with the correct data, algorithms may be capable of solving many of the inefficiencies, diagnostic errors and complexity within healthcare (217).

The principles of AI were developed concurrent with the computer, with concepts laid out by Alan Turing over 80 years ago (218). Modern deep convolutional neural networks were firmly established in 2012 in computer vision when a deep CNN achieved unprecedented accuracy in image classification (219). Advances in deep learning were made possible due to the availability of large datasets for training, high-performance computers and improved algorithms (220).

The ability of deep convolutional neural networks to learn representations of data with multiple layers of abstraction enables this learning technique to be applied to a vast array of data-driven problems. Examples of modern AI include self-driving cars, the outperforming of human ability in multiple games such as Go and Texas Hold’em Poker, and the diagnosis of disease in radiology (221). In medicine, AI has spread far beyond visual pattern recognition, with applications in robotic surgery (222), cancer genomics (223), drug discovery (224) and wearable sensors (225).

Artificial intelligence and its application to MRI has gained much attention and can be leveraged to improve diagnosis, accelerate imaging and personalise treatments (226). Between 2016 and 2020, literature searches on ‘Artificial intelligence’ and ‘MRI’ increased 5-fold (227). A systematic review of 12 studies (228) determined comparable performance between machine learning models and clinicians in assessing musculoskeletal images. Deep learning-assisted tools for diagnosis have improved specificity and lowered misdiagnosis rates (229,230).

Multitask deep learning models have reliably assessed radiographic features of hip osteoarthritis (231), with performance relative to that of expert radiologists (area under the curve values 0.80 – 0.98). Similarly, deep learning identified hip fractures from radiographs with an accuracy of 94%, achieving expert-level classification (232). Attention maps derived from the network improved the performance of a human rater; an example is shown in Figure 14.

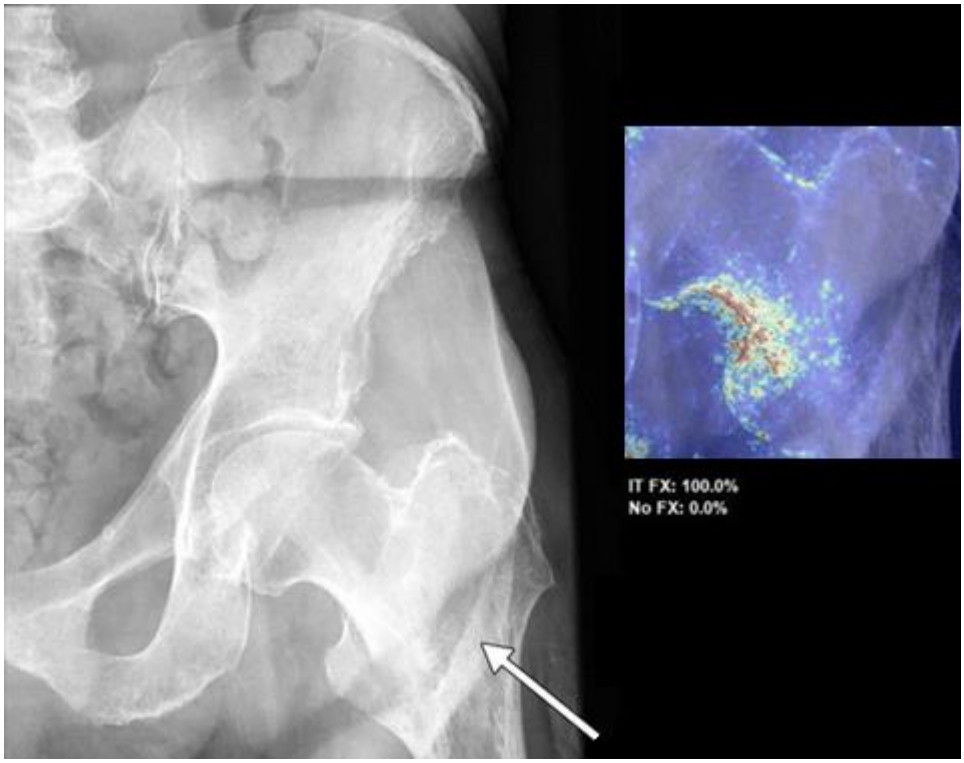


Figure 14: Advanced Deep Learning Applications in Hip Fracture Detection
This figure reproduced from Krouge et al. (227) exemplifies the advanced capabilities of deep learning in medical imaging, specifically in the detection of hip fractures from radiographs. Model-aided detection of hip fractures from radiographs can achieve expert level diagnosis, and aid in the reporting of fractures via attention maps which highlight areas where the model focuses when making decisions.

Neural networks have been used to generate synthetic images with CT contrast from MRI, enabling the evaluation of skeletal structures (albeit at lower resolution) without ionising radiation while simultaneously providing soft tissue contrasts in a single examination. Synthetic CT images generated to evaluate spinal pathologies showed moderate to excellent performance (233). For structural lesions, synthetic CT improved the detection of structural lesions in patients with suspected sacroiliitis with reliability comparable to that of CT (234) example is shown in Figure 15.

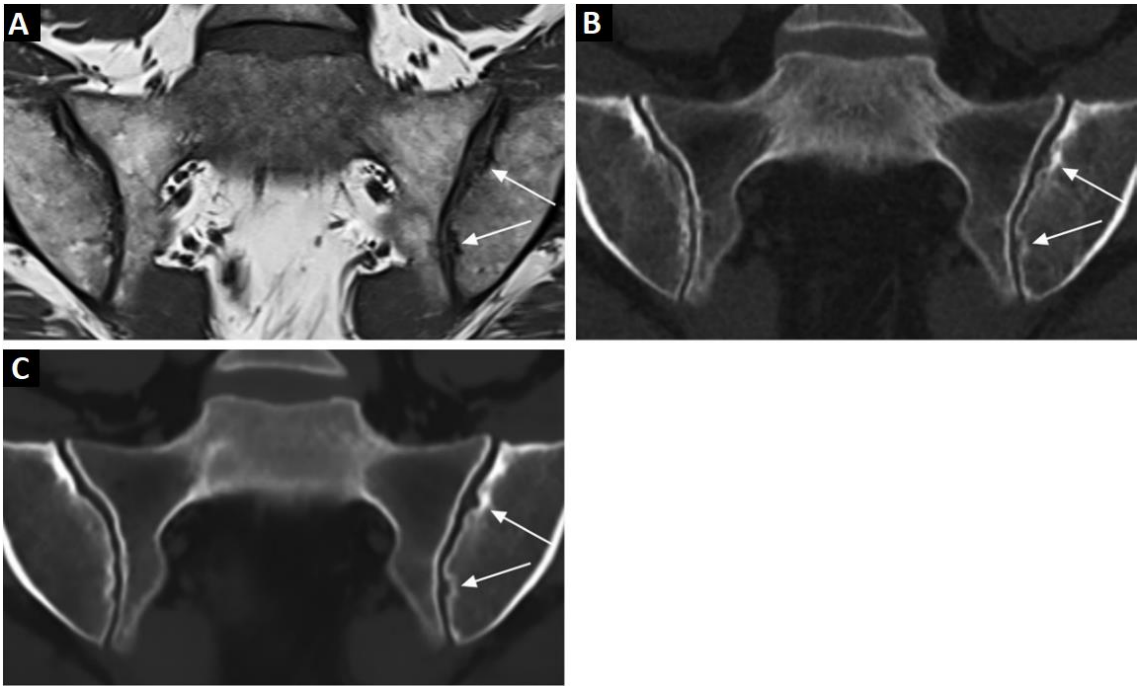


Figure 15: Advanced Imaging using MRI-based Synthetic CT
This figure, sourced from Lennart et al. (234) illustrates the innovative use of neural networks in medical imaging, particularly in the generation of synthetic CT images from MRI scans. The figure displays (C) a synthetic CT image, created using commercially available software from (A) a T1-weighted MRI scan. For comparison, (B) shows a reference CT scan. While the sacroiliac joint erosions are still clearly visible in the MRI, the enhanced appearance of bone in synthetic CT may reduce the need for additional imaging.

Deep learning image reconstruction has enabled short imaging times while maintaining diagnostic imaging quality. Recent developments suggest a potential paradigm shift in MRI, promoting a broader adoption of MRI and faster translation of research tools into the clinical space. Sequence-specific algorithms may develop ultra-fast imaging, for example, acceleration factors of 24-fold for three-dimensional double echo steady state imaging used in the osteoarthritis initiative protocol (235), an example of which is shown in Figure 16. Clinically, deep learning imaging reconstruction may enable up to 4-fold acceleration for a near-constant signal-to-noise ratio. At the same time, 43% reductions in scan time are achievable in practice while improving image resolution for conventional two-dimensional T1 weighted knee MRI (236,237).

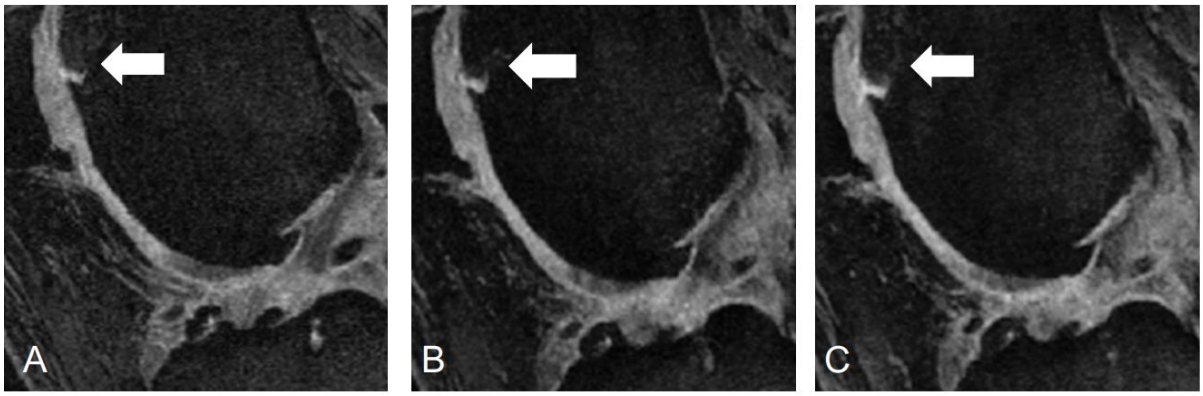


Figure 16: Ultra-Fast Deep Learning Image Reconstruction.
Image reproduced from Calivá et al. (235). Bespoke neural network-based image reconstruction shows a small bone marrow lesion (arrow) at (A) original resolution and at (B) 6-fold under-sampling and at (C) 12-fold under-sampling. The implementation of deep learning for image reconstruction has led to significantly shorter imaging times while preserving the quality necessary for diagnostic purposes. This advancement indicates a potential paradigm shift in MRI technology, fostering wider adoption of MRI in clinical settings and expediting the transition of research tools into practical clinical applications.

There has been some concern over the fidelity of Deep Learning (DL) algorithms in MRI. One example is an academic study on open-source DL algorithms, which demonstrated popular models can result in unstable image reconstruction with artifacts such as banding, blurring and structured noise (238). However, artefacts from manufacturer DL algorithms have not been reported to date. One feature that separates these commercial algorithms is training on unfiltered complex raw data. Later processes such as combining multi-channel data, creating magnitude images, gradient unwarping, and surface coil intensity correction change the noise profile and introduce spatial dependencies in image noise. The manufacturer's tight integration of their DL algorithms into the reconstruction process reduces artifacts and provides reproducible results (239).

2 Motivation, Aims and Overview of Thesis

This thesis relies on the excellent tissue contrast provided by MRI that gives extensive insights into osteoarthritis pathology. The focus on the knee joint is due to the high prevalence of knee osteoarthritis, leading to excess morbidity and mortality (2). The availability of large imaging datasets to support deep learning (240) has opened new avenues for diagnosing, managing and treating the disease. This thesis aims to support the development of methods for deep learning in osteoarthritis and apply these to advance our understanding of the disease. The thesis examines how imaging can help develop more tailored clinical trial designs and targeted treatments for a heterogeneous disease.

Segmentation is often the starting point for many quantitative assessments that have a role in assessing clinical trials and developing new methods for diagnosis and prediction in osteoarthritis. Due to the high time cost of manual segmentation, this has historically been limited to the assessment of clinical trials. However, advances in neural network segmentation have seen significant growth in these methods, supported by open-access datasets.

The use of segmentation in MRI has led to the rapid growth of outcome markers that can play an essential role in the stratification of individuals for clinical trials or as surrogate outcome markers to accelerate the discovery of new therapeutics. Mechanistic outcome markers may support tailored interventions, where subgroups with differing forms of OA may benefit from specific treatments.

Imaging in osteoarthritis has often focused on cartilage measures, which have been used to evaluate clinical trials (241,242). However, these measures only show a limited association with pain (243), which is the primary symptom of osteoarthritis. This dissonance may limit the use of cartilage-based markers to evaluate clinical trials adequately. Pain is the defining symptom of OA and the leading cause of individuals seeking help from their healthcare providers. New techniques, such as high throughput imaging features (radiomics), may support re-evaluating the association between cartilage and knee pain.

As a non-innervated structure, cartilage is likely to show secondary effects only when knee pain is examined. Adding structures directly responsible for knee pain may improve the ability of models to predict knee pain, thereby providing new biomarkers sensitive to knee pain. By combining features from different anatomical sites, such biomarkers may provide clinically meaningful information that can help untangle the heterogeneity within osteoarthritis.

The aims of the thesis are as follows:

1. Develop new segmentation methods for knee cartilage from MRI and examine how the automated segmentation-derived morphological and intensity descriptors are associated with descriptors of knee OA, including pain, stiffness and function.
2. Develop knee cartilage MRI as an endophenotypic assessment tool to predict which joint pathology will require future total knee replacement.' If successful, such a tool will have mechanistic potential to become a surrogate outcome marker for interventions to prevent or delay total knee replacement.
3. Examine the association between radiomic features of cartilage MRI and knee pain and test the ability of knee cartilage to predict pain in knee osteoarthritis.
4. Examine if the prediction of knee pain can be improved by adding new knee imaging targets beyond cartilage MRI.

In chapter three, the methods, concepts, and mathematical framework that inform the experimental chapters of the thesis are developed. This includes an introduction to deep learning and radiomic analysis. Deep learning is explored in supervised learning, where training data is labelled, and the mathematical framework for simple convolutional neural networks is introduced. The Radiomic analysis process includes segmentation, feature extraction, modelling and validation. In radiomics, the number of predictors is typically far greater than the number of observations, addressed via either regularisation or dimensionality reduction. The models explored are those used in the experimental chapters of the thesis and include Cox proportional hazard, multinominal regression and machine learning classification.

In chapter four, cartilage segmentation was developed using deep learning based on manual segmentation data made available by iMorphics (244). This chapter aimed to create a fully automated segmentation method that provided competitive results compared to existing cartilage segmentation methods. These segmentations were to be used in subsequent chapters for radiomic analysis. Therefore, segmentations were required to be anatomically plausible to enable meaningful radiomic feature extraction.

Using baseline knee MRIs, the cartilage segmentation method was deployed in the osteoarthritis initiative dataset (233). Before a radiomics analysis, two features were explored for their association with clinical scores. The first feature was cartilage volume, as its association with knee pain, stiffness and function has previously been studied (243). The second feature was the intensity range, as a proxy for signal intensity changes due to cartilage degradation. This comparison aimed to show that new biomarkers may quickly be developed and test the relative strength of association of the new biomarker with clinical symptoms.

In chapter five, the knee cartilage segmentation of the baseline data from the osteoarthritis initiative was used to predict total knee replacement within five years in patients with symptomatic knee osteoarthritis. This is based on the idea that timely identification of individuals at risk of developing severe symptoms is highly desirable for targeted

preventative interventions. Additionally, predicting total knee replacement (TKR) will enable optimised trial design through risk stratification of populations and mechanistic surrogate outcome markers based on the assumption that structural join pathology ultimately drives the need for TKR.

The model was developed using radiomic features from the Imaging Biomarker Standardization Initiative guidelines (245). This method sought to develop and internally validate a robust model using elastic net regression and to perform multivariate Cox regression to calculate individualised risk scores and proportional hazard scores for high and low-risk individuals.

Chapter six examines the association between knee cartilage MRI and osteoarthritis pain using new imaging descriptors (radiomics). The chapter assesses if knee cartilage radiomics may better identify suitable imaging biomarker candidates for selecting patients most likely to benefit from specific treatment options. The correlation between baseline knee cartilage MRI features, knee pain and established 9-year pain trajectories is investigated. Subsequently, the ability of cartilage radiomic features to predict knee pain and pain trajectories is explored using several different models for achieving this outcome (machine learning vs neural network classifiers).

Chapter seven uses an atlas-based approach to extend the existing cartilage-based radionics pain model to include other anatomical targets. As well as incorporating pain-sensitive regions of the knee, an atlas-based segmentation model is intended to smooth out the inter-subject variation due to knee size, thus improving the model's ability to detect changes based on the underlying pathology.

3 Methods

A range of different methods are used within this thesis. This chapter presents the framework and mathematical definitions to aid in understanding these concepts. A broader examination of deep learning and convolutional neural networks is given, as these are used in chapter four for segmenting knee articular cartilage and chapter five for classification. The feature extraction process (radionics) is described as forming the basis of the prediction models in chapters five and six. Atlas-based segmentation is also presented, as this is used in chapter seven. Other methods presented are predominantly related to feature reduction techniques and prediction (classification and regression).

3.1 Deep learning

Several definitions exist for Deep learning; at a conceptual level, deep learning is a process of learning the relation among two or more variables, the knowledge that governs the relation, and the knowledge that makes sense of the relation (246). Convolutional neural networks are intended to achieve deep learning within a given system. Neural networks are composed of multiple processing layers to learn representations of data with various levels of abstraction (247).

These methods have dramatically improved state of the art in image recognition (248–250)speech recognition (251–253) and have accelerated developments in domains such as drug discovery (254) and genomics (255). In radiology, these techniques are leveraged to extract information from, segment, diagnose or classify medical images. Deep learning has also been used for image reconstruction (256), generation of synthetic CT images (257), super-resolution (258) and image reconstruction (256).

The use of computer-assisted analysis is well-established in the field of medical image analysis. Typically, features are 'hand-crafted' and rely on domain-specific knowledge, such as the Fourier transform for conversion of frequency data to images in MRI or the characterisation of the pixels in the image for segmentation. In deep learning, the hierarchical feature representation is learned solely from data rather than pre-defined. In that way, deep learning is proving to be the state-of-the-art foundation in medical image analysis (259).

In applying machine learning for data analysis, meaningful feature extraction is central to accomplishing target tasks. Usually, experts design these features based on their knowledge of the target domains. This feature engineering extracts quantitative image markers such as image intensity, morphology and texture to represent the information within the image. While attempts to standardise feature extraction have been made (260), these features may not fully characterise image heterogeneity due to the disease process or differences in how the data was acquired.

Deep learning circumvents such obstacles by establishing the feature-engineering step as a learning step. Rather than designing features that give good predictive value, these are learnt by the deep learning model as part of the training process. The burden of feature-engineering shifts from human to computer-sided, thus enabling non-experts to use deep learning for their research, especially medical image analysis.

Deep learning is not a new concept; however, its unprecedented success is multi-factorial given more powerful computer processing units and graphical processing units, the availability of large datasets (i.e. big data), the development of learning algorithms and open source machine learning platforms (247).

3.1.1 Supervised learning

The most common form of learning is supervised learning, where training data is labelled. In the scenario where the aim is to identify planar X-ray images that have broken bones out of a larger set of planar X-rays. Firstly a large dataset of X-ray images containing fractured and non-fractured planar X-rays should be collected and labelled. During training, the network is shown an image and outputs a probability score, one for each category. The aim of the network is to have the output of the desired category have the highest score, with the other category having a low score. To achieve this an error function that calculates the difference (known as the error) between the output and desired scores is calculated. The internal parameters of the network are then updated to reduce the difference in the error function in each iteration. Several adjustable parameters, often called weights, define the network's input-to-output function. In a typical network, there may be many thousands or millions of these weights and a similarly large number of outputs.

To adjust the weights, the learning algorithm calculates a gradient vector that indicates the quantity the error would increase or decrease if a small amount increased the weights. The weight vector is adjusted to decrease the error, reducing the difference between the output and desired scores.

The error function averaged over all the training data can be considered a landscape with peaks and valleys in the high dimensional space of weight values. The negative gradient vector indicates the direction of the steepest descent in this landscape. After several iterations of revising the weights, the gradient vector descends closer to the minimum of this landscape, where the difference between output and desired scores is the lowest.

This process is known as stochastic gradient descent (SGD); the number of examples used to calculate the gradient vector before recalculating the weights is called the batch size. The process is ‘stochastic’ because each batch gives noisy estimates of the average gradient over all the training data. A complete pass through all the training data is known as an Epoch. The process is repeated for several epochs until the average of the error function stops decreasing or a desirable result is achieved.

3.1.2 Formal definition

Artificial Neural Networks are computational models inspired by the human brain. They are comprised of many connected nodes, with each node performing a mathematical operation. By connecting these nodes, very complex functions can be learned. A simple diagram of an artificial neural network is shown in Figure 17.

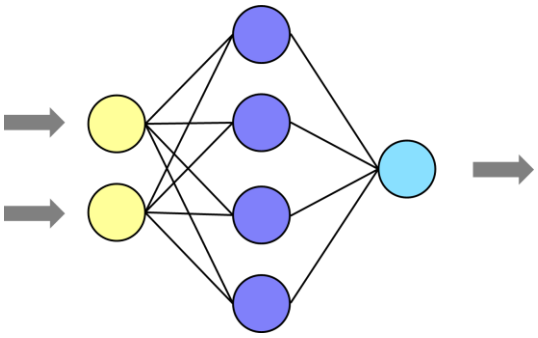


Figure 17: Artificial Neural Network Architecture

In this simple example of a neural network, the yellow column of circles represents the inputs, the middle purple column represents the computational units that act on the inputs, and the third column represents the network's output.

The mathematical operations performed by a node include applying weights and a bias to inputs followed by an activation function. Figure 18 shows a node with three inputs $\vec{x} = (x_1, x_2, x_3)$, one output y, bias b, and weight vector $\vec{w} = (w_1, w_2, w_3)$. These are passed to the activation function f, so the output of a node is given by (261):

$$y = f(\vec{w} \cdot \vec{x} + b) \qquad (1)$$

Each of the inputs $\vec{x} = (x_1, x_2, x_3)$ can be the outputs of another node or the initial input. Similarly, the output y can be the input to another node or the final output. Each component w_i of the weight vector corresponds to each component x_i of the input vector. Therefore, the summation of the product of the individual w_i, x_i components is equivalent to the dot product.

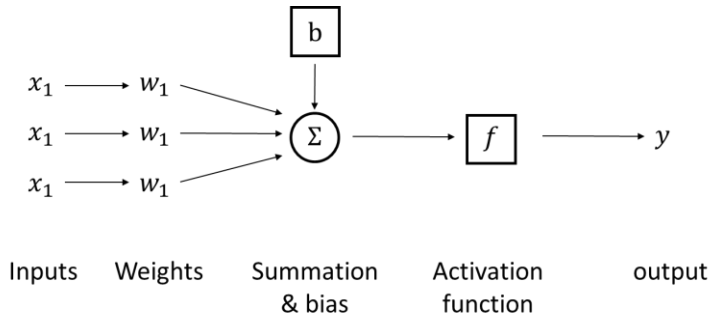


Figure 18: Node within a neural network
The neuron has three inputs $\vec{x} = (x_1, x_2, x_3)$, weight vector $\vec{w} = (w_1, w_2, w_3)$ and bias b followed by an activation function φ which determines how the weighted sum of the inputs is transformed into an output y.

In the human brain, neurons function by firing only when they receive enough input from other neurons they are connected to(262). In artificial neural networks, activation functions allow non-linear transformations between inputs to learn more complex functions; without activation functions, a neural network would behave as a simple linear regression model (262). Different activation functions have been proposed, but possibly the most common function is the rectified linear unit (Relu), which has the form (262):

$$\varphi(x) = \begin{cases} 0 \; for \; x \leq 0 \\ x \; for \; x > 0 \end{cases} \qquad (2)$$

To enable the learning of more complex functions, artificial neural networks often have many more nodes and can be arranged in highly complex connection patterns.

The neural network can calculate some function $f_\theta(\overrightarrow{x_i})$ for input-output pairs $(\overrightarrow{x_i}, \overrightarrow{y_i})$, together with the individual nodes' weight vectors and biases θ. To train the network requires the model's weight vectors and biases θ to be altered such that $f_\theta(\overrightarrow{x_i}) \approx \overrightarrow{y_i}$ for all i.

This can be done by choosing an error function E over the set of pairs $X = \{(\overrightarrow{x_1}, \overrightarrow{y_1}), \dots (\overrightarrow{x_n}, \overrightarrow{y_n})\}$ that can be minimised to re

duce the difference between predicted outputs $f_\theta(\overrightarrow{x_i})$ and the output targets $\overrightarrow{y_i}$. Common error functions are the mean squared error in the case of regression problems and the cross entropy in the case of classification problems. The mean squared error function is defined by (263):

$$E(X, \theta) = \frac{1}{2N} \sum_{i=1}^{N} \left(y_i - f_\theta(\overrightarrow{x_i})\right)^2 \qquad (3)$$

To train a neural network this function is minimised iteratively using gradient descent. Each iteration of gradient descent updates the weights and biases θ according to (259):

$$\theta^{t+1} = \theta^t - \alpha \frac{\partial E(X, \theta)}{\partial \theta} \tag{4}$$

where θ^t denotes the parameters of the neural network at iteration t in gradient descent. Training a neural network typically starts with a random initialization of the parameters θ followed by successive updates to those parameters based on gradient descent until the error function converges.

Since nodes within the network prior to the output have no target y_i, an energy function can't simply be defined that is specific to that node. Instead, any error function for that node will be dependent on the value of the parameters in the previous layer (since previous layers determine the input for that node) and following layers (since the output of that node will affect the computation of the energy function $E(X, \theta)$. Back propagation applies the chain rule to calculated the derivative $\frac{\partial E}{\partial \theta}$ allowing the weights and biases to be updated at every node in an efficient manner (263).

The purpose of the $1/2$ factor in the cost function in equation (3) becomes apparent in the gradient decent. When you take the derivative of the cost function with respect to the model parameters (like weights in a neural network), the 2 from the exponent in the squared term comes down and cancels out the $1/2$ in front. This results in a cleaner, simpler expression for the gradient.

3.1.2.1 Convolutional neural networks

Convolutional neural networks (ConvNets) are a popular choice for deep learning in medical imaging (243) as they preserve the structural relationship between neighbouring pixels or voxels (252). ConvNets are designed to process data that is composed of multiple arrays. For example, inputs can be two-dimensional for images or three-dimensional for volumetric data.

A typical ConvNets architecture is structured as a series of stages, as shown in Figure 19. The first few stages feature convolutional layers, rectified linear units and pooling layers. These are followed by fully connected layers and a softmax layer (output).

To describe these layers in more detail, the example of a CovNet with multiple 2D inputs of size $n_{in} \times n_{in} \times channels$ is used. The variable $channels$ allow a certain generalizability, for example $channels = 3$ is typical for RGB colour images while $channels = 1$ is appropriate for grayscale images(264).

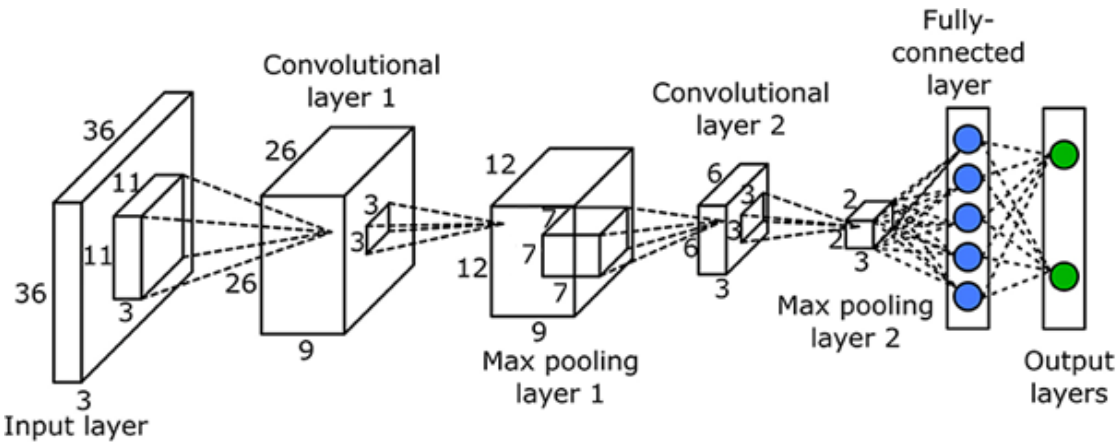


Figure 19: Convolutional neural network
Image reproduced from (262). A Convolutional neural network consists of multiple layers, such as convolutional layers, max-pooling and fully-connected layers. In this example, a small colour image is classified by the network. In the convolutional layers, filters detect specific features, such as edges, textures, or more complex patterns in higher layers. Pooling layers reduce the input volume's spatial dimensions (width and height) for the next convolutional layer. These decrease the computational load, memory usage, and the number of parameters, thereby reducing the risk of overfitting. At the end of the network, a fully connected layer connects to all activations in the previous layer. This layer helps classify the image based on the high-level features extracted by the convolutional and pooling layers.

3.1.2.2 Convolutional layer

A convolutional layer consists of nodes connecting to sub-regions of the input image or the previous layer's outputs (as seen in Figure 19). The objective of convolution operation is to extract the high-level features. In a convolutional layer, a set of 'k' filters (kernels) are applied to the input image. The value k defines the depth of the output volume. In Figure 19, $k = 9$ for convolutional layer 1 and $k = 3$ for convolutional layer 2; different filters may identify distinct image feature such as colour, texture, or edges.

Each filter is of size $f \times f \times channels$ in the convolutional layer. From Figure 19, the filters in convolutional layer 1 are of size $11 \times 11 \times 3$. Each filter is applied, then stepped across the input image, the number of pixels moved is referred to as the stride size s. When the stride is 1 then the filter is moved one pixel at a time. The stride defines the number of pixels moved at a time as the filter moves across the image.

Padding the input volume with zeros around the border helps control the size of the output volume The amount of zero padding, p is used to ensure the filter can be applied an integer number of times across the image.

The spatial size of the output volume is a function of the input volume size, the filter size, the stride and the amount of zero padding (262):

$$n_{out} = \left[\frac{n_{in} + 2p - f}{s}\right] + 1 \qquad (5)$$

These hyperparameters have mutual constraints due to their spatial arrangement such that n_{out} must be integer. For example, in Figure 19, convolution layer 1, for the k filters to 'fit' when applied to the input image then $p = 0$ and $s = 1$ gives $n_{out} = \frac{36+0-11}{1} + 1 = 26$.

3.1.2.3 Pooling layer

Pooling layers progressively reduce the spatial size of the representation to reduce the amount of parameters and, therefore, computation in the network. Pooling layers are commonly inserted in between successive convolutional layers. This downsampling also controls for overfitting when training the network. In Figure 19, max pooling layer 1, the filter size is $7 \times 7 \times 7$ and is applied with a stride $s = 1$ and padding $p = 0$ resulting in 92% of the activations being discarded.

3.1.2.4 Fully connected layer

Nodes in a fully connected layer have full connections to all the nodes in the previous layer (262). The full connection to the previous layer enables high-level reasoning in the neural network.

3.1.2.5 Output layer

In the output layer, a softmax function modifies the outputs as a vector of values that sum to one, so each output value can be interpreted as a probability. While several softmax activation functions exist, the most common form is (265)

$$f_i(x) = \frac{e^{x_i}}{\sum_j e^{x_i}} \qquad (6)$$

for each output i, with a total of j outputs.

3.2 Radiomics

Radiomics is the high-throughput mining of quantitative image features from radiographic images to improve diagnostic, prognostic and predictive accuracy, leading to more individualised treatments (266). The suffix *-omics* has been used in fields that generate complex data from single objects or samples, such as DNA (genomics), RNA (transcriptomics), proteins (proteomics) and metabolites (metabolomics) (267). The term radiomics is often used to describe the entire analysis, including image post-processing, data extraction and subsequent analysis, from the point the image has been acquired to a final imaged-based signature (268–270). It is proposed that the quantitative features extracted from modalities such as CT or MRI represent changes in the underlying tissue structure, biology or pathology that correlate with clinical and biological data from the individual (269).

Radiomics builds upon computer-aided diagnosis (260) and uses hundreds of high throughput image features and other pertinent data sources (including clinically obtained, treatment-related or genomic data) to produce robust, clinical decision support systems (271). Radiomics studies have previously been described as having five distinct phases: data selection, medical imaging, feature extraction, exploratory analysis, and modelling (272), as shown in Figure 20. There are many potential pitfalls in developing and deploying a complete radiomics workflow, including poor reporting due to technical complexity, overfitting of data, incomplete validation of findings, and unrecognised confounding variables in the databases used, mainly if data are derived retrospectively (269).

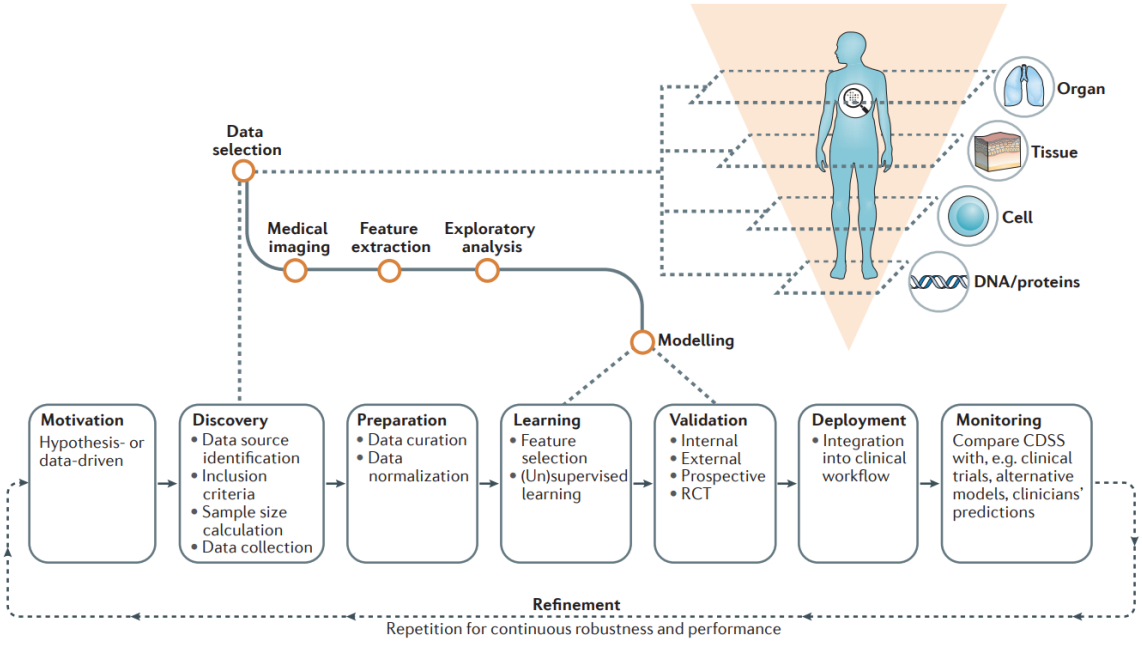


Figure 20: Radiomic study workflow

Radiomic analysis includes several stages, from initial conception to external validation, deployment and continual assessment for robust deployment as a clinical tool. There are many potential challenges when building models for clinical deployment, and following standardised frameworks may lead to more valuable models that are accurate and reproducible.

3.2.1 Segmentation

Segmentation is the process of creating automatic labels based on ground truth annotated data. The ground truth segmentations are typically image masks made by expert raters; these define the anatomy or pathology of interest in the image. Many techniques exist to automate image segmentation; simple thresholding has been used to segment musculoskeletal structures, including bone marrow lesions, knee effusion, bone and muscles (273–277). Conversely, deep learning has provided robust and scalable methods for segmenting a wide range of musculoskeletal targets (221), often outperforming previous attempts at segmentation (278,279).

3.2.1.1 Atlas-based segmentation

Atlas-based segmentation uses a predefined template in an atlas space to enable automated geographical (template-based) segmentation. In this method, the atlas comprises an image and associated segmentations transferred to a target image via registration. After registration of the template to the target image, the atlas labels are propagated to the target image. This segmentation method is routinely employed in morphological and morphometrical MRI studies of the brain (280). In a typical workflow, the template is registered to the atlas via an affine transform (translate, rotate, scale and shear) followed by a non-rigid registration based on mutual information (281).

In knee osteoarthritis, atlas-based methods have been used to segment femoral and tibial cartilage in pre-surgical cases, create maps of cartilage thickness and T2 relaxation time of the patellar, and perform knee pathology-pain mapping of bone marrow lesions (282–284). Atlas-based segmentation in knee joint structures is inherently challenging due to heterogeneity of image intensity, variability of joint structures, differences in position within the knee coil and image artefacts due to motion. Conversely, recent advances in convolutional neural networks have seen deep learning segmentation overtake atlas-based methods (285).

3.2.1.2 Deep learning segmentation

Various techniques exist for deep learning image segmentation, the most common of which are encoder-decoder networks and conditional generative adversarial networks (286–288). In convolutional encoder-decoder networks, the encoder part of the network consists of a contracting path that generates successively compressed feature maps. At the same time, the decoder contains consecutive up-sampling operations, which map the low-resolution encoder feature to full input resolution feature maps for pixel-wise classification (288). A conditional generative adversarial network builds on the encoder-decoder network. It receives either the predicted or ground truth segmentations and outputs a single scalar to discriminate if the input is fake or real. This adversarial role boosts the segmentation network and encourages long-range spatial contiguity (286).

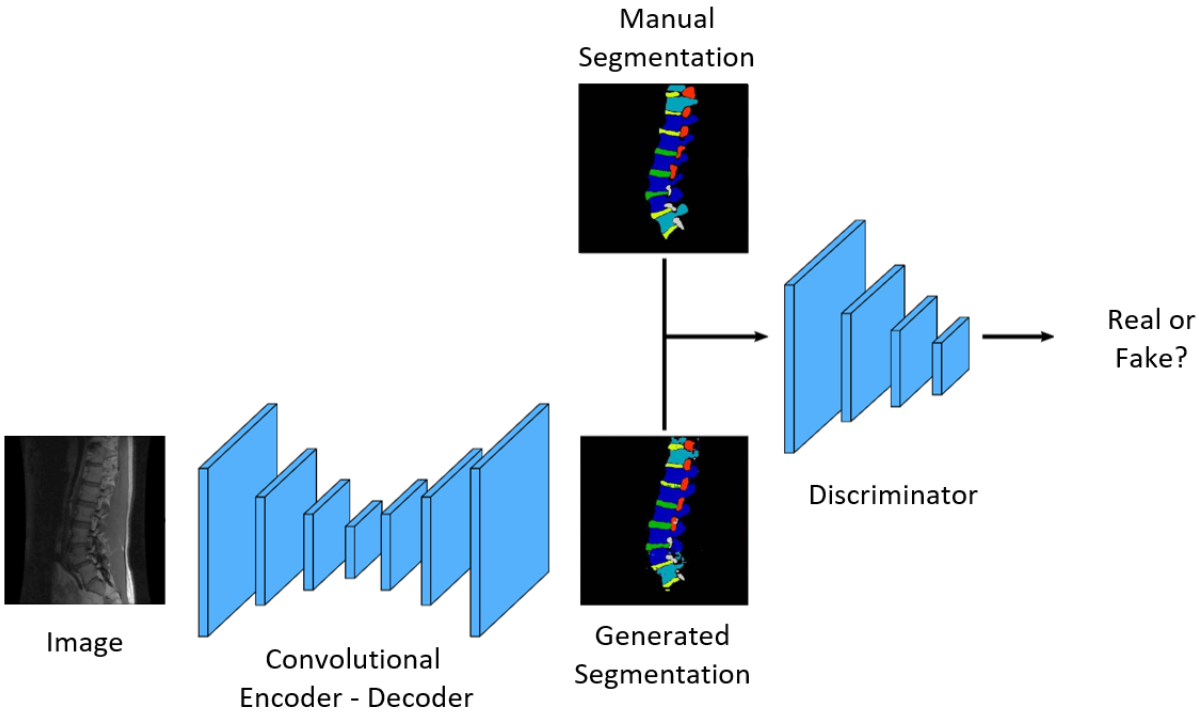


Figure 21: Convolutional generative adversarial network
The Illustration is based on the segmentation network from Han et al. (289). Image segmentation is performed by an encoder–decoder network, where each pixel is annotated as one of seven spine tissues. A generative convolutional network attempts to identify if its input is the manual or generated segmentation; this discriminative network encourages the segmentation network to produce correct predictions. The adversarial role of the discriminator is intended to boost spatial contiguity, leading to more accurate anatomical segmentations.

A limitation of deep learning imaging segmentation is the requirement of significant quantities of high-quality data for training, i.e. the Structural Knee Imaging 2010 Challenge: 250 cases, International Workshop on Osteoarthritis Imaging Challenge: 176 cases, and spinal segmentation in Figure 21: 253 cases (278,279,290). Hybrid approaches, utilising advanced post-processing techniques for label assignment, have shown promising results with as little as 20 cases, as shown in Figure 22 (291). The advantage of atlas-based registrations is that they can be implemented using only a single segmentation (the atlas label), making it possible to achieve fast segmentations with only limited manually annotated data.

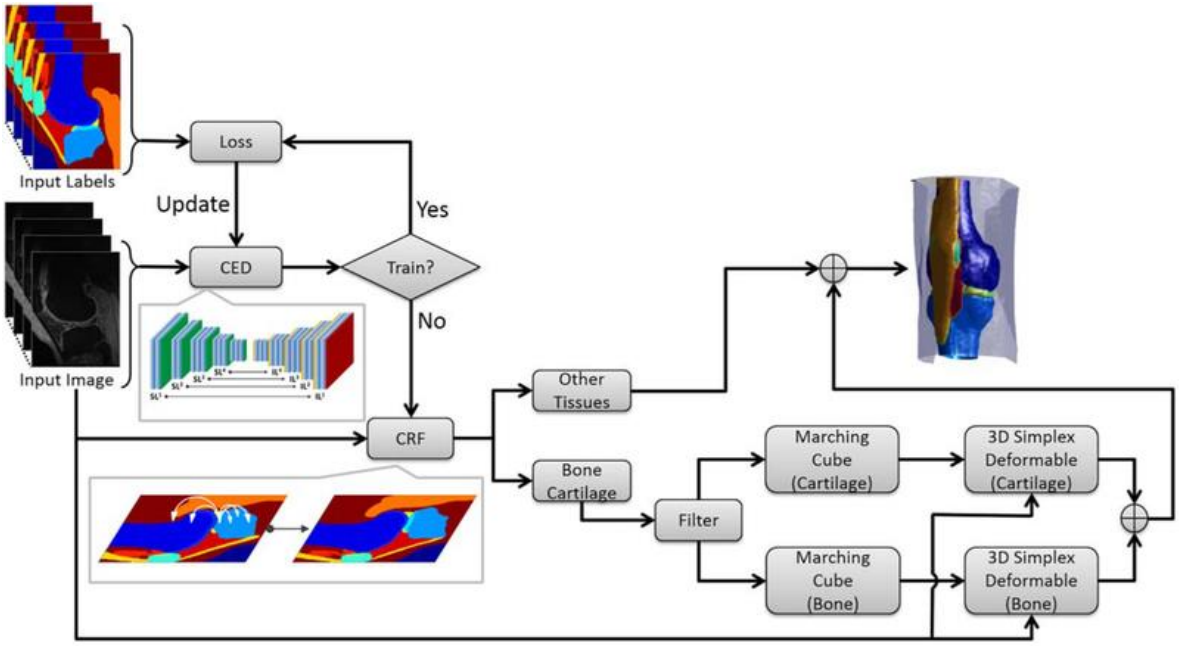


Figure 22: Deep learning-based knee joint structure segmentation
Image reproduced from (291). The backbone of the segmentation task is performed by a convolutional encoder-decoder (CED) deep learning network, which segments 12 different joint structures. A 3D conditional random field (CRF) was applied to regularise contextual relationships among voxels within the same tissue class and between other classes. A deformable model refines this output to preserve the overall shape and maintain smooth surfaces for joint structures. This type of end-to-end model can achieve fast and accurate segmentations of a wide range of tissues with only limited amount of manually annotated data.

3.2.1.3 Dice Coefficient

The Dice similarity coefficient (DSC) is the most used metric in evaluating the performance of different segmentation models. In image segmentation, the DSC is used as a direct spatial comparison between automatic and ground truth segmentations. The DSC is defined by (262)

$$\mathrm{DSC} = \frac{2(S_g \cap S_t)}{S_g + S_g} \qquad (7)$$

Where S_g is the ground truth segmentation and S_t is the test segmentation, $S_g \cap S_t$ is the intersection of these two regions, as shown in Figure 23. One limitation of the dice coefficient is that it does not account for partial volume effects, which can be significant for smaller structures, and does not measure boundary delineation errors, which may be necessary for certain tasks (292,293).

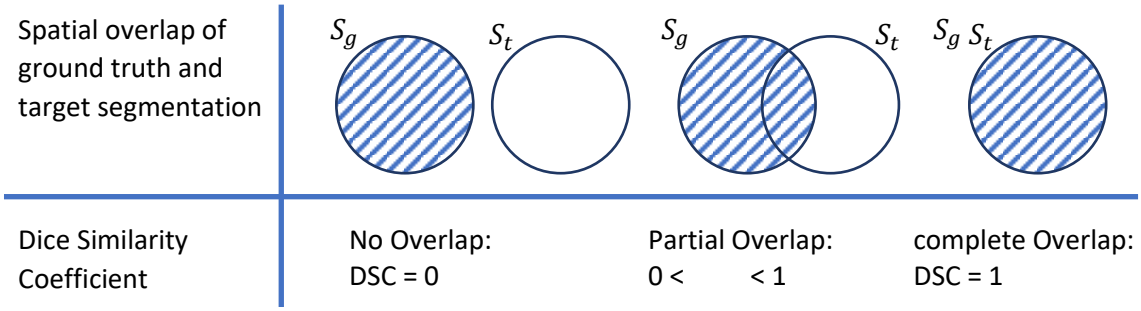


Figure 23: Visualization of the Dice Coefficient
The dice coefficient represents the spatial overlap between automatic and ground truth segmentations. A value of one indicates complete overlap, while a value of zero indicates no overlap. While the dice coefficient remains a popular metric it does not account for boundary delineation errors or partial volume effects.

3.2.2 Feature Extraction

Attempts to unify quantitative image features for radiomic analysis, such as the Image biomarker standardisation initiative (260) have been widely adopted (294–296). A systematic review found that the reproducibility of radiomic features was sensitive to image acquisition settings, image pre-processing, and software used to extract radiomic features (297), highlighting the need for proper external validation of study findings.

Quantitative radiomic features can be divided into several families, which include intensity-based statistical, intensity histogram-based, intensity-volume histogram-based, morphological features, local intensity, and texture matrix-based features (245). A segmentation mask is required to calculate features and identify voxels within a region of interest. The definition of the region of interest is a source of potential bias, while manual segmentation by an expert rater is considered the gold standard; due to time requirements, this is not scalable to larger data sets (298).

3.2.3 Modelling

This section explores the methods for modelling using high-dimensionality data and explore the models used within the thesis, including the Cox proportional hazard model.

Radiomic features are incorporated in a model for radiometric analysis; this can either be via regression (the output is a continuous variable) or classification (the output variable takes class labels). Radiomics datasets are typically sparse due to their high dimensionality, referred to a 'small n, large p' as radiomic analysis often deals with many more predictors (p) than observations(n). This presents a unique challenge, as the data samples needed to estimate an arbitrary function with a given level of accuracy grow exponentially with respect to the number of input variables in the function (299). This reduces the ability of conventional algorithms to find optimal solutions when modelling the data and risk overfitting (289).

Classically, there are two leading solutions to dealing with high-dimensional data. One is to classify them directly in high-dimensional space with regularisation methods, and the other is to reduce data dimension first and then organise them in feature space (300). It is also possible to use convolutional neural networks to model data without specifying a particular dimensionality reduction or regularisation technique (301) the network learns the mapping from the predictor variables and their outcomes.

3.2.3.1 Principle Component Analysis

Principal component analysis is a multivariate dimensionality reduction technique (302). It represents the data as a set of new orthogonal variables called main components (303). The first principal component is chosen to minimise the distance between the data and their projections onto the principal component. This is equivalent to maximising the variance of the projected points (σ^2). Subsequent components are selected using the same criteria and must be orthogonal to all previous principal components. This limits the number of principal components to equal to or less than the number of features or samples, whichever is smaller. A two-dimensional example is shown in Figure 24.

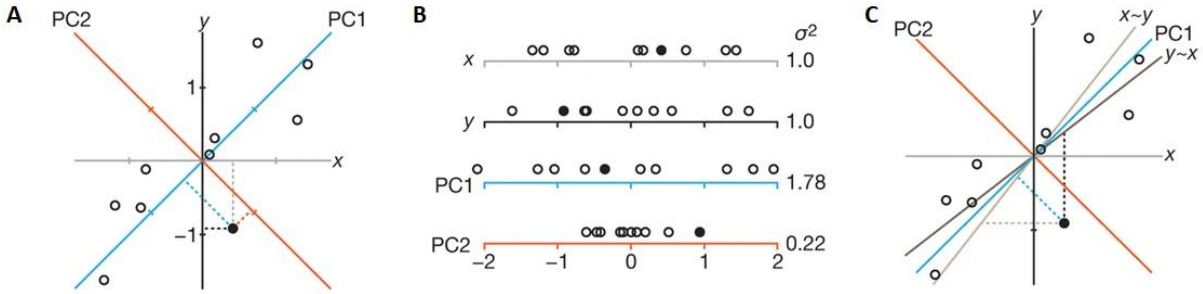


Figure 24: Principal Component Analysis

(A) Data sampling includes variable x and y, these are projected onto the first principal component (PC1) by minimising the distance between $[x, y]$ and the projection, the second projection (PC2) is orthogonal to the first principal component. (B) The variance σ^2 can vary between projections but is a maximum for the first principal component and reduces for subsequent components. (C) The PC1 is not the same as the regression of y vs. x (light grey) or x vs. y (dark grey) as principal component analysis minimises the distance between data points and the projection. In contrast, linear regression minimises the distance between the response variable and its predicted value. The figure is adapted from Lever et al. (304).

3.2.3.2 Regularization

The purpose of regularisation is to remove the influence of predictors that are not useful. Collinearity of predictors reduces the interpretability of model coefficients and inflates the variance of regression parameters, potentially leading to incorrectly identifying relevant predictors in a model (305). There are two commonly used regularisation techniques: lasso

regularisation removes unimportant coefficients by shrinking their coefficients to zero (299), while ridge regression reduces the size of coefficients with high covariance (304). Lasso regression is often more appealing due to its sparse data representation, improving data interpretability. However, lasso regularisation has several limitations; in the case where the number of predictors (p) is much bigger than the number of observations the model (n) lasso regularisation does not provide satisfactory variable selection. Additionally, for usual $n > p$ situations, if the predictors are highly colinear, then ridge regression dominates the predictive performance of the lasso (306).

Elastic net regularisation combines the lasso and ridge techniques and provides the means to overcome these limitations. The naïve description of elastic net regularisation presented here is adapted from the original paper by Zou et al. (306).

Consider a dataset with given p predictors and n observations to perform elastic net regularisation Let $\boldsymbol{y} = (y_1, \dots, y_n)^T$ be the response with predictors $\boldsymbol{x}_j = \left(x_{1j}, \dots, x_{nj}\right)^T$ where $j = 1, \dots, p$. For this dataset the model matrix takes the form:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{8}$$

After a translation and scale transform, responses can be centred with a standardized set of predictors

$$\sum_{i=1}^{n} y_i = 0\,, \quad \sum_{i=1}^{n} x_{ij} = 0, \quad \sum_{i=1}^{n} x_{ij}^2 = 1, \quad \text{for} \quad j = 1, 2\,, 3, \dots, p \tag{9}$$

In the case of linear regression, the response is predicted by

$$\hat{\boldsymbol{y}} = \hat{\beta}_0 + \boldsymbol{x_1}\,\hat{\beta}_1 + \cdots + \boldsymbol{x_p}\hat{\beta}_p \tag{10}$$

where a model fitting procedure produces the vector of coefficients $\widehat{\boldsymbol{\beta}} = (\hat{\beta}_1, \dots, \hat{\beta}_p)$. In the case of elastic net regularization, the following criterion is defined,

$$\mathrm{L}(\lambda_1, \lambda_2, \boldsymbol{\beta}) = |\boldsymbol{y} - \boldsymbol{X\beta}|^2 + \lambda_2|\boldsymbol{\beta}|^2 + \lambda_1|\boldsymbol{\beta}|_1 \tag{11}$$

For any fixed non-negative values λ_1 and λ_2, where

$$|\boldsymbol{\beta}|^2 = \sum_{j=1}^{p} \beta_j^2 \qquad \text{and} \qquad |\boldsymbol{\beta}|_1 = \sum_{j=1}^{p} |\beta_j| \tag{12}$$

The elastic net estimator, $\widehat{\boldsymbol{\beta}}$ should minimize the value:

$$\widehat{\boldsymbol{\beta}} = \min_{\beta_0,\beta} \{L(\lambda_1, \lambda_2, \boldsymbol{\beta})\} \tag{13}$$

if a constant is defined as $\alpha = \frac{\lambda_2}{\lambda_1+\lambda_2}$, then equation 13 is equivalent to the optimization problem:

$$\widehat{\boldsymbol{\beta}} = \min_{\beta_0,\beta} |\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}|^2 \text{ subject to } (1-\alpha)|\boldsymbol{\beta}|_1 + \alpha|\boldsymbol{\beta}|^2 \leq t \text{ for some } t \tag{14}$$

This function, $(1-\alpha)|\boldsymbol{\beta}|_1 + \alpha|\boldsymbol{\beta}|^2$ is the elastic net penalty and is a combination of lasso and ridge penalties. With a value of $\alpha = 1$ the problem becomes ridge regression without any lasso component, while as alpha shrinks towards zero, the elastic net penalty approaches lasso regression.

Solutions to the minimization problem have enabled elastic net regularization to be applied to both linear regression and general linear regression problems (307). Elastic net regularised models have successfully improved the identification of molecular markers in genomic data and the diagnosis and prognosis of cancer outcomes based on radiomic features (308–310). To minimize overfitting, the original datasets were split into training and testing datasets; within the training dataset, the parameters for the elastic net regularization were fine-tuned before being deployed in the test dataset.

3.2.3.3 Cox Proportional Hazards Model

Cox proportional hazard regression is a semiparametric method to qualify the effect of predictor variables on time-to-event estimates. In the Cox model, explanatory variables share a common baseline hazard function, $h_0(t)$, where variables are explained as a multiple of this baseline variable.

Consider a dataset with given p predictors and n subjects. For subjects $i = 1, \dots, n$, and predictors $\boldsymbol{x}_j = (x_{1j}, \dots, x_{nj})$ where $j = 1, \dots, p$, the model matrix has the form given in equation (15). The predictor variable $\boldsymbol{X_i} = (x_{i1}, \dots, x_{ip})$ for the i-th subject has a hazard rate at time t of:

$$\mathrm{h}(\boldsymbol{X_i}, t) = h_0(t) \exp\left[\sum_{j=1}^{p} x_{ij} b_j\right] \quad (15)$$

where the hazard rate is the instantaneous probability of an event occurring at time t, given that the event has not already occurred for that individual (311). From this definition, the hazard ratio can be defined as:

$$\mathrm{HR}(\boldsymbol{X_i}) = \frac{h(\boldsymbol{X_i}, t)}{h_0(t)} = \exp\left[\sum_{j=1}^{p} x_{ij} b_j\right] \quad (16)$$

By relating the hazard rate for individual predictor variables to the baseline function, the hazard ratio is time-independent. This proportional hazard assumption enables the calculation of survival rate (312), defined as the probability that the event of interest has not occurred by duration t,

$$S_{\boldsymbol{X_i}}(t) = S_0^{HR(\boldsymbol{X_i})} \quad (17)$$

Cox regression is a widely used technique to identify significant predictors, evaluate the effect of different treatments, and predict disease progression (313–315)Cox models often use iterative fitting, such as forward selection or backward elimination variable selection algorithms, to reduce the number of model variables (316).

3.2.4 Validation

Proper validation is a limitation of many existing studies, given the fact that radiomics is a young discipline. While positive results for internally validated studies may suggest the clinical utility of a given model, robust validation is required. This includes validation against a completely independent data set, preferably from another institution, and a prospective multicentre trial wherein the biomarker is one of the primary endpoints (317,318). Given the high threshold required to develop clinical radiomic tools, state investment, such as the £250 million NHS AI labs, is justified based on the belief that these tools will significantly benefit patients, clinicians, and the overall system (319).

One internal validation method used in this thesis is k-fold cross-validation (314); it is often implemented in machine learning and radiomic models (320–322). It involves partitioning the training data into k equal-sized subsamples, training the model on k-1 subsamples, and evaluating it on the remaining one. This process is repeated k times, with a different subsample used as the test set in each iteration. The final performance measure is the average of the performance measures obtained on each iteration.

An example of how k-fold cross-validation can be implemented:

1. Divide the training data into k equal-sized subsamples.
2. Train the model on k-1 subsamples and evaluate it on the remaining one.
3. Repeat this process k times, using a different subsample as the test set in each iteration.
4. Calculate the average performance measure across all k iterations.

K-fold cross-validation is a valuable technique for evaluating the performance of a machine learning model's performance because it estimates how well the model is likely to perform on unseen data. It also helps to reduce the variance in the performance measure, which can be helpful in cases where the model's performance can vary significantly depending on the training data given. One potential drawback of k-fold cross-validation is that it can be computationally expensive, especially if the model takes a long time to train and evaluate.

4 Ensemble Learning for Robust Knee Cartilage Segmentation

4.1 Introduction

Osteoarthritis (OA) is associated with an increasing socioeconomic impact owing to the ageing and increasingly obese population (41). Manual segmentation of cartilage tissue from MRI has shown sufficient sensitivity to detect changes, including early progression (323,324). Also, greater rates of cartilage loss are associated with frequent knee pain, irrespective of adjustment or stratification for radiographic disease (325). However, morphological measurements do not assess cartilage quality and are complicated by the biphasic nature of cartilage swelling in early OA prior to thinning and loss (326,327). To address this, quantitative changes in MRI have focused on the isotropy of the collagen network (e.g., T2 mapping) or tissue composition (e.g. gagCEST or dGEMERIC) (328).

Morphology and quantitative cartilage measures have been associated with knee pain (243,329). Given the central position of cartilage in drug development, pathological investigation and the aetiopathogenesis of OA, it is critical to understand which cartilage measurements are more strongly related to pain and function. This can be addressed by comparing results from knee cartilage segmentations. However, manual cartilage segmentation is time-consuming and susceptible to inter-rater variability (330) traditional automated registration and shape-based classification (331–335) have weaker performance metrics when compared to deep learning approaches (336–340).

Convolutional Neural Networks (CNNs) provide an efficient and precise alternative for segmenting many anatomical structures in medical images. (341–343). In particular, U-Nets, which comprise a contraction path (encoder) to capture context followed by an expansion path (decoder) to enable localisation and shortcuts between selected layers, have shown improved segmentation accuracy (344).

Two-dimensional and more resource-expensive, fully volumetric 3D U-Nets have been proposed for multi-label knee cartilage segmentation (337–339). U-Net-derived architectures have shown promising results when trained and evaluated on small datasets with as few as 20 subjects without augmentation for segmenting knee joint structures (331). Further refinements have been made using conditional random fields (22) and statistical shape modelling (SSM) and applied to meniscal cartilage (340), femoral cartilage (333) and tibial cartilage (333,334). 2D U-Nets may struggle to differentiate medial from lateral cartilage. The SSM methods referenced may overcome this, but using 3D U-Nets and bespoke registration templates is more computationally expensive.

In this chapter, ensemble learning is used to increase the accuracy of the deep learning method while keeping computational resources limited by training a 2D U-Net with only two labels (cartilage vs background) to offset the high inter-rater variability in manual annotations (278,335), and a multi-label U-Net to identify cartilage compartments. Ensemble learning for segmentation has demonstrated robust performance by aggregating predictions (345,346). Using a hierarchical approach cartilage is segmented from the background before classifying specific cartilage components to reduce over-fitting and boost segmentation results. Using simple post-processing, anatomically valid segmentations that contain only a

single region of interest per cartilage label are developed. This ensemble learning approach performed better than or better than existing methods that provide labels for all cartilage compartments.

The segmentation model is applied to MR images from the OAI. From these segmentations radiomic features are extracted, including knee cartilage volume and intensity range. These features were then associated with clinical outcomes. As DESS MRI show a strong association with T2 values (347,348), the range of intensity values served as a proxy for qMRI changes and examine the association between intensity features and WOMAC measures. While the strength of correlations remained weak, the range of intensity features showed a stronger correlation compared to cartilage volume in the femoral, medial tibial, and lateral meniscal cartilage.

4.2 Materials and Methods

4.2.1 Study Dataset

To develop the automated cartilage segmentation, freely available data from the Osteoarthritis Initiative (OAI) was used, with cartilage annotations produced by iMorphics (244). Ethical approval was obtained for collecting this information about the subjects from the OAI. The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000. Informed consent was obtained from study participants (349).

The dataset from iMorphics used for this retrospective analysis consisted of MRI scans acquired with a 3D- Double Echo Steady-State (DESS) sequence from 88 subjects with osteoarthritis; imaging was obtained at two time points, doubling the number of images available for training. Images from all subjects were included with manual segmentations for six cartilage compartments. The manual segmentations were performed by a trained musculoskeletal radiologist and reviewed by a technical specialist at iMorphics (350). The segmentations passed the iMorphics cartilage segmentation training protocol, requiring an intra-observer coefficient of variation lower than 3% on paired test images (350). The subjects consisted of 45 male and 43 female participants aged 61.2 ± 10.0 years and BMI of $31.3 \pm 4.6 kg/m^2$ (mean $\pm$ standard deviation). Based on the Kellgren-Lawrence (KL) scores for the classification of knee OA (113), baseline scores were KL-1 2%, KL-2 34%, KL-3 59% and KL-4 5%. MRI volumes had a 140mm field of view, 0.7mm slice thickness and a matrix of 384×384 with 160 slices. Manual annotations were provided for the patellar, femoral, lateral tibial, medial tibial, lateral meniscal and medial meniscal cartilage.

4.2.2 Data Preparation

N4 Bias field correction was used to remove low-frequency intensity non-uniformity in the MRI data (351). Each MRI volume was intensity normalised by subtracting the mean voxel intensity and dividing by the standard deviation. MRI data was randomly split 70/20/10% for training, validation, and testing. As the dataset consisted of baseline and 12-month MRI, this partitioning of the data was limited such that the same person would not appear in both the training and test sets. For the training data, 2D in-plane slices (348×384 matrix)

were shuffled and randomly combined in batches for training. To train the 2-label U-Net (2l model) for cartilage segmentation, cartilage compartments were combined such that there were only two labels: cartilage and background. To train the 7-label U-Net (7l model) the fully annotated dataset was used.

4.2.3 Model architecture and training

Using the 2D U-Net architecture shown in Figure 25, the 2-label model was trained for cartilage segmentation from background and a 7-label model for segmenting each cartilage compartment. Convolutional layer inputs were zero-padded to preserve the height and width of the outputs from these layers. The network was trained using the Adam optimiser (352) with a learning rate of 1×10^{-4}; a weighted categorical cross-entropy loss function was used to account for imbalances in the number of training examples for each cartilage label. The U-Net was trained for 300 epochs on the training dataset. At each epoch, the Dice coefficients were averaged over all labels for the entire validation dataset to give a multi-label Dice score. The model weights from the epoch with the highest multi-label Dice coefficient were selected for the combined model.

The predictions from both models were used to create a combined (2l7l) model with pixel voting; the 2-label model was used to determine if a voxel contained cartilage, and if so, the 7-label model was used to determine the cartilage compartment (ignoring the background class). All code was implemented in Python 3.7.3 utilising Tensorflow 1.13.1 (*https://www.tensorflow.org/*) and Keras 2.2.4 (*https://keras.io/*) packages. All training and evaluation were conducted on a dedicated server hosting a 64-bit CentOS 7 operating system. Computing hardware included 80 Intel Xeon E5-2698 v4 CPUs @ 2.20GHz, 512GB DDR4RAM, and two Nvidia Tesla K10 graphic cards. These resources were made available by the University of Nottingham's Precision Imaging Beacon (353).

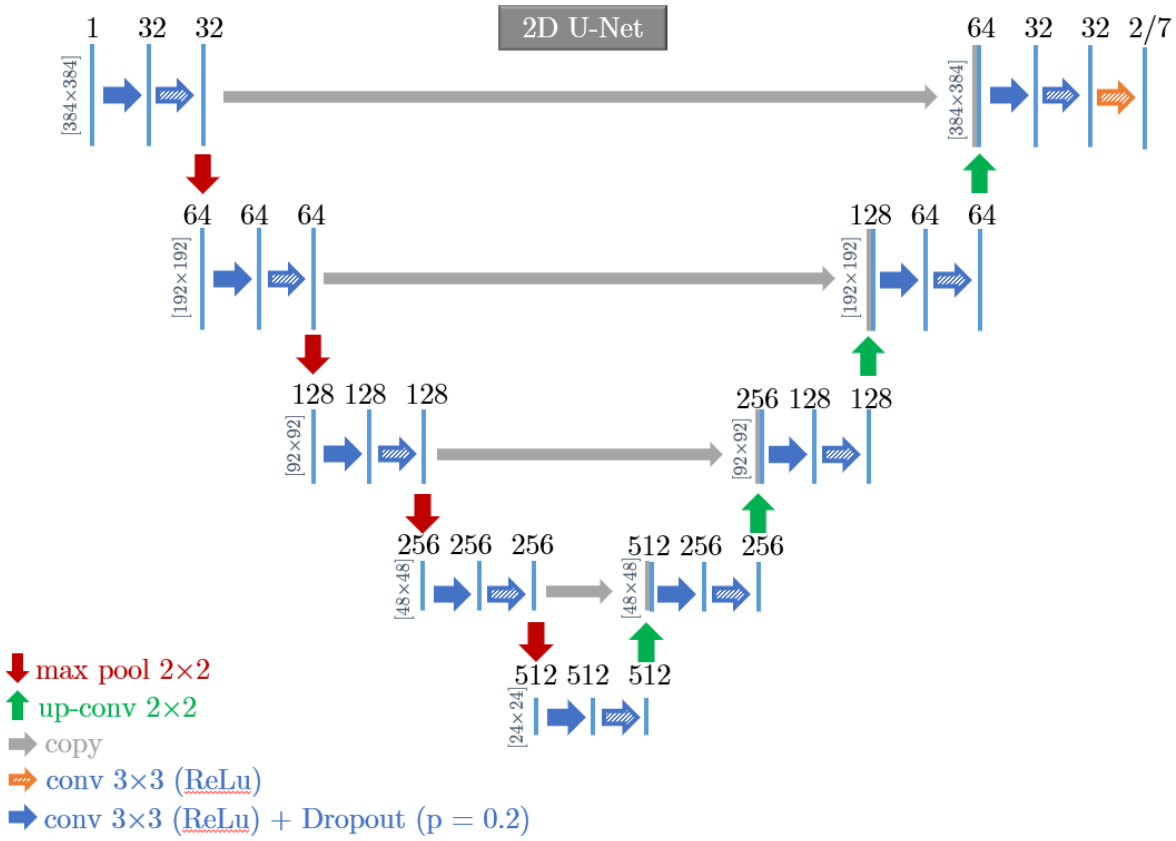


Figure 25: U-net architecture for cartilage segmentation
This figure illustrates the U-net architecture explicitly designed for cartilage segmentation from MRI images. The network is 2D, with inputs and outputs the same as the in-slice matrix size of the DESS MRI. The multi-channel feature maps, represented by blue boxes, vary in depth, with the number of channels indicated at the top of each box and the spatial dimensions provided on the left edge at each depth level. The grey boxes signify copied feature maps, illustrating the network's mechanism for combining high-level features, with more detail and low-level features with more generalisation through skip connections. Arrows depict various operations, including convolution and pooling, with the first 3x3 convolution in the expansion path utilising a stride of 2 to reduce feature map size. The final output layer, shown as a softmax layer, adapts the number of channels to accommodate a 2-label or a 7-label cartilage model, which are recombined later to improve the sensitivity of the segmentation tasks.

4.2.4 Post-processing

To refine the outputs of the combined network (7l2l model), rule-based post-processing was applied to the segmentation maps provided by the model (Figure 26). The aim was to create anatomically plausible segmentations by removing or relabelling connected components to have a single region of interest per cartilage type to improve the quality of any subsequent morphometric analysis.

For each cartilage compartment, the connected components were relabelled based on the labels of neighbouring voxels, where the most significant connected component kept its

original labels assigned by the U-Net. 3D connected components with a connectivity type 26 were used such that voxels were connected if their faces, edges, or corners touched. The remaining connected components were relabelled based on the most frequent label of neighbouring voxels.

1. Identify all connected components for a given cartilage subtype.
2. For each component (excluding the largest), extract the neighbouring voxels in contact with the perimeter of the object.
3. Relabel the object based on the most frequent neighbouring voxel.
4. Repeat steps 1-3 for all cartilage subtypes.

Figure 26: Post-processing of segmented keen cartilage

The algorithm details the processing for segmentation outputs from the combined U-net model (7l2l). The aim was to produce anatomically plausible segmentations by ensuring a single connected component for each cartilage type, thus enhancing the segmentation quality for subsequent morphometric analysis. This involves relabelling or removing extraneous connected components, using 3D connected component analysis with a connectivity criterion that considers voxels connected if they share faces, edges, or corners. The largest connected components retain their original labels assigned by the U-Net. In contrast, smaller ones are relabelled based on the most frequent label among neighbouring voxels, thereby reducing the number of connected components to one per cartilage type and improving anatomical accuracy.

4.2.5 Model evaluation

The evaluation was performed with a 3-fold cross-validation approach in which 10% of the data was held out for testing. A similar approach to previous studies was used, which used varying folds ranging from 2 to 5 folds (338,339,354). For each fold, the remaining 70% was used to train the model and 20% to validate each epoch to prevent overfitting. The resulting segmentations from the test sets were pooled when evaluating the segmentation model. The accuracy of tissue segmentation was calculated using the Dice similarity coefficient for each cartilage compartment. It was defined as $DSC = \frac{2|S \cap R|}{|S|+|R|}$ where S denotes the automatically segmented voxels and R denotes the reference manually annotated voxels. Dice coefficients range between 0 and 1 for each cartilage compartment, with 1 representing a perfect overlap and 0 meaning no overlap of the segmentations. Volumetric measures (scatter and Bland-Altman) were used to compare the corrected 7l2l model and manual segmentations.

4.2.6 Association with clinical outcomes

Using this model, 6 cartilage types were automatically segmented from the 3D double echo steady state scans of 4784 participants from the OAI (age 61.1±9.2, BMI 28.6±4.9, 57% female, KL-0: 37% KL-1: 20%, KL-2: 27%, KL-3: 13% and KL-4: 3%, WOMAC Scores, Pain: 2.4±6.6, range 0 – 20, Function 16.1±21.9, range 0 - 68, Stiffness 1.5±3.3, range 0 - 8).

Where available, both left and right knee MRIs were included (n = 9383), as WOMAC scores were provided for each knee. The association between cartilage volume, intensity range and WOMAC scores were tested using Pearson's correlation coefficients; the Bonferroni adjustment was used to correct for multiple comparisons.

The intensity range (maximum grey level value minus the minimum grey level value) was used as a proxy for qMRI values; healthy cartilage has a wide range of intensity values as T2 is much shorter in deep cartilage zones compared to the superficial zones (355). While data was acquired from four separate sites, identical scanners and protocols were used to minimize inter-site variability (356,357). As OA progresses, superficial cartilage is lost, and the range of T2 values reduces, resulting in a narrower range of intensity values. The correlation coefficient using Fisher's r to z transform was used to determine if cartilage volume or intensity range is more strongly associated with WOMAC pain, function and stiffness scores.

4.3 Results

The overall training time for each fold was approximately 48 hours given the computing hardware used. The segmentation of all six regions of interest took 13.6s in total for each image volume, with a mean computing time of 3.6s for applying the combined U-Net and 10s for the post-processing.

Figure 27 shows the Dice coefficients for each segmented cartilage structure. The Dice coefficient for each label was evaluated over the test sets obtained from the 3-fold cross-validation to prevent small or large cartilage regions from a given subject from biasing the Dice score for a given cartilage compartment. The combined 7l2l model achieved higher Dice scores across all compartments than the initial 7l model. While the improvements in Dice from post-processing were minor (corrected 7l2l model), the number of connected components was reduced from 21 per cartilage compartment to just one connected component per cartilage compartment. The visual assessment showed that the refinement yields more anatomically plausible segmentations, which may improve the quality of radiomics features from these segmentations.

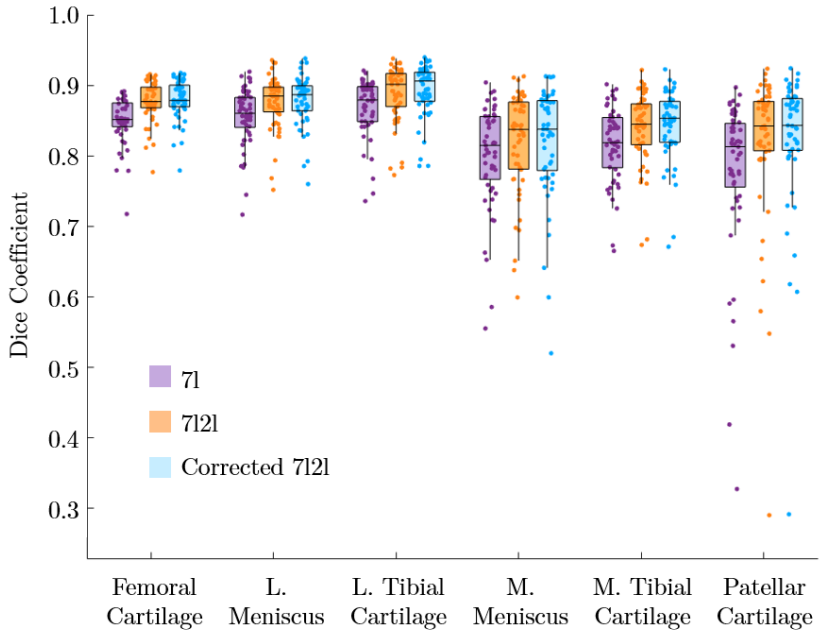


Figure 27: Dice coefficient for cartilage segmentation
Box and jitter plot of Dice coefficient values for different cartilage types segmented in the pooled test sets. The Dice coefficient indicates higher values and lower variance for femoral, lateral meniscal, and lateral tibial cartilage, in contrast to the patellar, medial meniscal, and medial tibial cartilage. The comparison includes the original 7-label U-Net, the combined model (7l2l), and the corrected 7l2l model. This figure underscores the improved segmentation results when combining the original 7-label with a discriminant 2-label model. Also, the corrected 7l2l model has nearly identical Dice scores, showing that these refinements have marginal effects on the volume overlap of the automated and ground truth segmentations.

The Dice coefficient was used to compare these results to other automated methods for segmentation using the iMorphics dataset (Table 2). The approach take in this chapter reached state-of-the-art segmentation accuracy compared to other deep learning-based methods (333,334,337–340,354) and outperformed previous methods that provided automated segmentations for all cartilage labels. In the recent review of methods submitted for the International Workshop on Osteoarthritis Imaging (IWOAI) Knee MRI segmentation challenge, methods had Dice coefficients of between 0.81 - 0.90. They used either 2D or 2D U-Nets or V-nets (279). Dice coefficients (mean and 95% confidence intervals) for the method used in this chapter were: femoral 0.88 ± 0.03, medial tibial 0.84 ± 0.08, lateral tibial 0.88 ± 0.04, patellar 0.85 ± 0.10, medial meniscal 0.85 ± 0.05 and lateral meniscal 0.90 ± 0.04.

1. Author	Femoral Cartilage	Tibial Cartilage *(Medial Lateral)*	Patellar Cartilage	Meniscal Cartilage *(Medial Lateral)*
Norman et al. 2018 (337)	0.87(0.03)	0.78(0.03) 0.80 (0.04)	0.77(0.09)	0.73(0.05) 0.81 (0.03)
Raj et al. 2018 (338)	0.85(0.01)	0.81(0.02) 0.86(0.01)	0.78(0.04)	0.80(0.01) 0.85(0.01)
Tack et al. 2018* (354)	-	-	-	0.84(0.06) 0.89 (0.02)
Ambellan et al. * 2019 (333)	0.89(0.02)	0.86(0.05) 0.90 (0.02)	-	-
Tack et al 2019. * (334)	-	0.88(0.04) 0.91(0.02)	-	-
Desari et al. 2019 (340)	0.88(0.03)	-	-	-
Panfilov et al. 2019 (339)	0.91(0.02)	0.90(0.03) †	0.87(0.05)	0.86(0.03) †
IWOAI Challenge (279)	0.91(0.05)	0.90 (0.6) †	0.88(0.09)	0.87 (0.05)†
Peake et al. (this study)	0.88(0.03)	0.84(0.08) 0.88(0.04)	0.85(0.10)	0.85(0.05) 0.90(0.04)

Note. *Incorporates 2D U-Nets, statistical shape modelling and 3D U-Nets, †lateral and medial components were combined.

Table 2: Dice coefficients for cartilage segmentation
Published Dice coefficients for automatic cartilage segmentation methods using the iMorphics dataset. Dice coefficients, expressed as mean values with their corresponding standard deviations in parentheses, offer insights into the segmentation performance of various studies. The table includes results from both individual research efforts and the top submission from the International Workshop on Osteoarthritis Imaging (IWOAI) Challenge dataset, serving as a benchmark for comparison. The inclusion of these diverse sources provides a broad perspective on the current state-of-the-art in cartilage segmentation accuracy, notably top performing models have similar Dice coefficients.

Figure 28 shows sagittal examples of tissue segmentation performed on the 3D DESS images in the test sets. The results from the 7-label U-Net demonstrate good agreement with the overall contours of the manual segmentation. In the example, the annotation for the lateral meniscus is missing in the manual annotations. The intermediary (uncorrected) 7l2l model recovers some of this information but does not correctly label all the voxels; the corrected 7c2c model recovers the proper label.

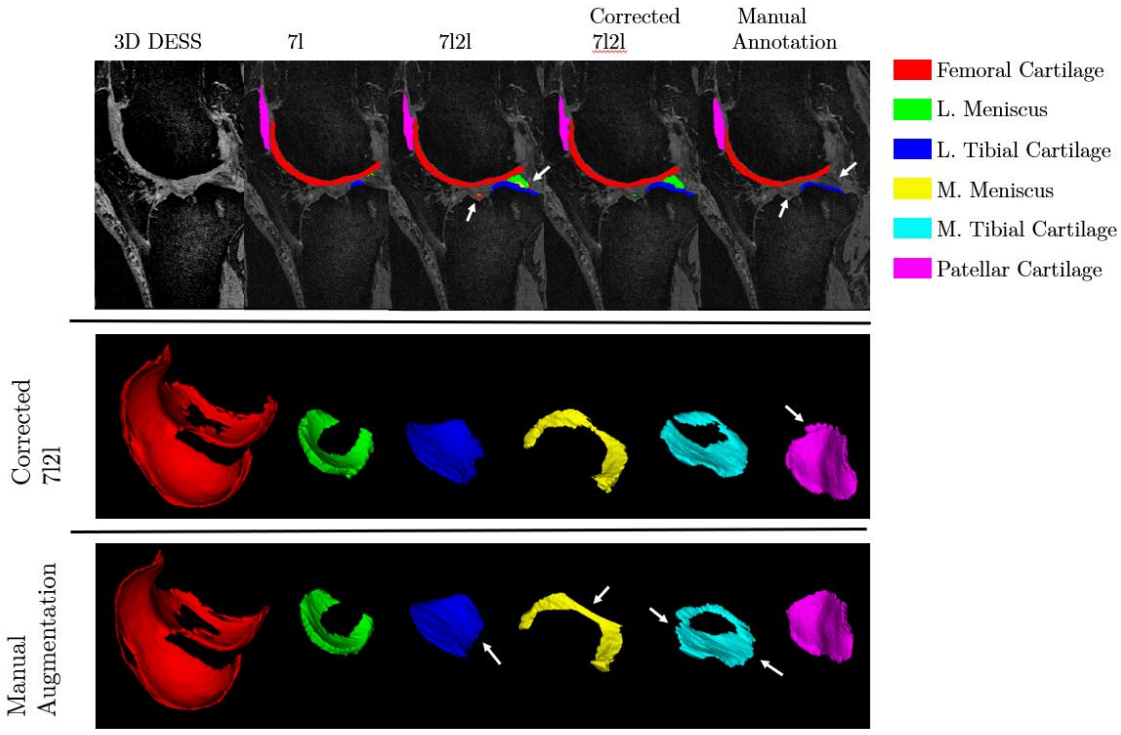


Figure 28 Knee cartilage segmentation from MRI
knee cartilage segmentations were performed on 3D Double Echo Steady State (DESS) MRI images, highlighting the segmentations by the original 7-label U-Net, the intermediary 7l2l model, and the corrected 7l2l model against manual annotations. The white arrows indicate areas where the corrected model has recovered missing annotations, such as those for the lateral meniscus. The model achieves smoother border profiles in 3D renderings but is also susceptible to incorrectly annotating synovial fluid in areas with poor tissue contrast.

Also shown in Figure 28 are the 3D rendered volumes for the corrected 7l2l model and manual segmentation. In the presence of advanced tissue degeneration, there is less agreement between the automatically segmented structures and the manual annotations. This is partly due to the increased subjectivity of segmentation in these regions, as adjacent structures have comparable signal contrasts. In addition, cropping and irregular boundaries are visible in the manual annotations, while the automated segmentations have smoother boundaries. However, automated segmentations occasionally spill over into adjacent structures, as seen in the patellar.

Figure 29 shows scatter and Bland-Altman plots for the corrected 7l2l model when applied to the pooled test sets. The test data set volume for manual and automatic segmentations shows a strong linear relationship across all cartilage compartments. For the pooled test sets, the average volume correlation across cartilage compartments was 0.988, with an average absolute mean difference across compartments of 0.43cm^3. The regression on the scatterplot shows a significant trend towards larger predicted volumes (gradient 1.034, 95% confidence interval 1.003 – 1.067). On the Bland Altman plot, the root mean squared error (RMSE) appears stable across the range of volumes (-9.3%). While the standard deviation is normally distributed (Kolmogorov–Smirnov test, $p < 0.0001$), the percentage difference between actual and predicted volume tends to decrease for larger volumes.

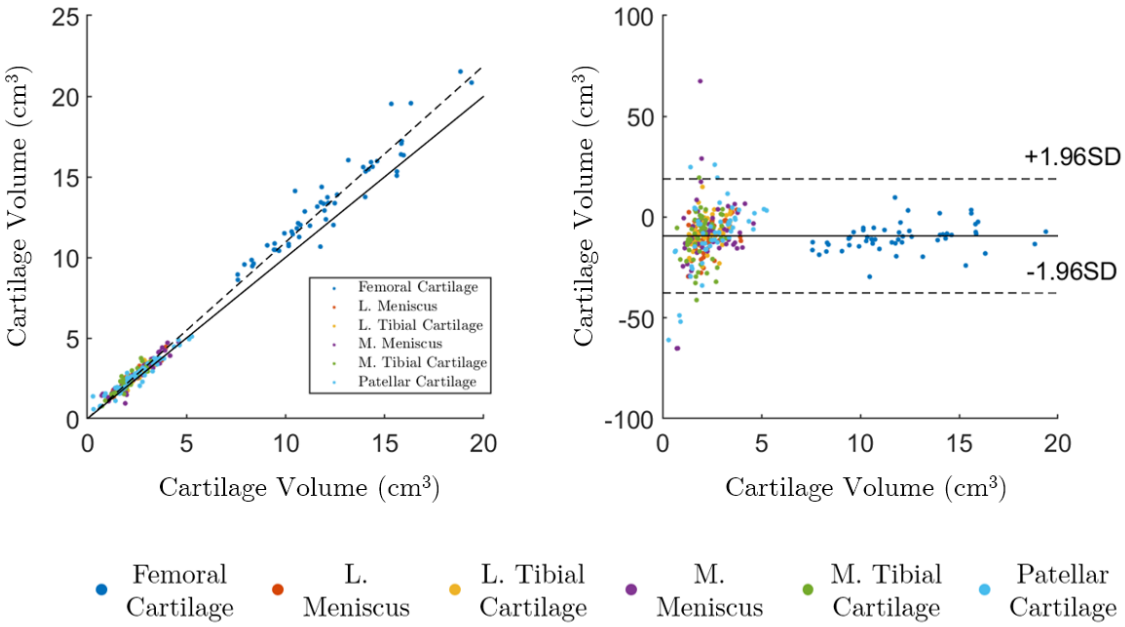


Figure 29: Volumetric analysis of knee cartilage segmentations
Scatterplots and a Bland-Altman plot comparing volumetric measurements from the corrected 7l2l model to manual annotations, broken down by cartilage type. The linear regression and the limits for the Bland-Altman plot were calculated using the pooled test data. The strong linear relationship depicted in the scatterplots and the Bland-Altman plot's indication of consistent error across the volume range demonstrates the model's accuracy in estimating cartilage volumes. Notably, the model shows a slight trend towards predicting larger volumes, as evidenced by the regression analysis, but maintains a high volume correlation across all cartilage compartments.

The correlation between WOMAC scores and the cartilage volume and intensity range are shown in Table 3. The mean cartilage volumes were (SD): femoral cartilage 9.1cm^3 (2.3), lateral meniscus 1.6cm^3 (0.5), lateral tibial cartilage 1.7cm^3 (0.5), medial meniscus 1.7cm^3 (0.5), medial tibial cartilage 1.5cm^3 (0.4) and patella cartilage 2.0cm^3 (0.7). In Table 3, all associations were significant at $p<0.0001$ except the femoral cartilage volume to WOMAC Pain ($p = 0.45$), Function ($p = 0.03$) and Stiffness ($p = 0.40$). The strength of association with WOMAC scores in the femoral, medial tibial, and lateral meniscal cartilage was greater for the intensity rage than the cartilage volume ($p<0.01$). There was no significant difference between the remaining three compartments.

		Femoral Cartilage	**Tibial Cartilage** *(Medial Lateral)*		**Patellar Cartilage**	**Meniscal Cartilage** *(Medial Lateral)*	
WOMAC Pain	Volume	-0.01	-0.07	-0.10	-0.10	-0.06	-0.04
	Range	-0.13	-0.10	-0.10	-0.12	-0.08	-0.09
P-value:		<0.0001	0.01	0.77	0.10	0.38	0.0006
WOMAC Function	Volume	-0.02	-0.08	-0.11	-0.13	-0.08	-0.05
	Range	-0.15	-0.12	-0.12	-0.14	-0.10	-0.11
P-value:		<0.0001	0.004	0.75	0.78	0.19	<0.0001
WOMAC Stiffness	Volume	-0.01	-0.07	-0.10	-0.12	-0.07	-0.04
	Range	-0.13	-0.11	-0.10	-0.12	-0.08	-0.10
P-value:		<0.0001	0.003	0.99	0.72	0.24	0.0001

Table 3: Associate of cartilage features with symptoms

detailed analysis of the correlation between cartilage volume and intensity range with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scores, which assess pain, function, and stiffness in osteoarthritis patients. P-values are given for a Fisher's r-to-z comparison of the cartilage volume and intensity range. Notably, all associations presented in this table are significant ($p<0.0001$), with the exception of the femoral cartilage volume's correlation with WOMAC Pain, Function, and Stiffness scores, where the p-values suggest a lack of significant correlation. While these p-values are significant, they do not represent strong correlations, indicating cartilage morphology may only be weakly associated with clinical symptoms. Additionally, volume measurements appear less important than signal intensity when assessing knee cartilage.

Figure 30 demonstrates the differences between segmentation and intensity values; before cartilage denudation, signal intensity changes occur relating to cartilage lesions and thinning.

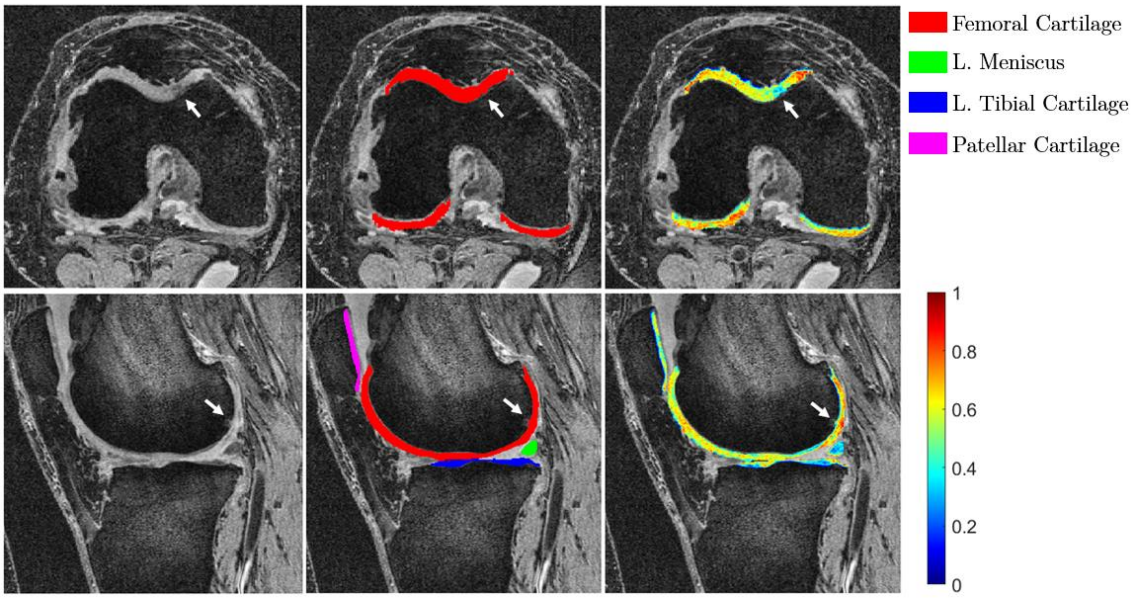


Figure 30: Femoral cartilage lesion on double echo steady state MRI
The lateral knee compartment shows automated segmentations (centre) and signal intensity variations within the segmented cartilage (right). **Top:** intercondylar femoral cartilage lesion causes loss of signal. **Bottom:** hyperintense cartilage appears adjacent to a small bone marrow lesion. Signal intensity changes that precede morphological changes and, therefore, may be more sensitive to clinical outcomes. These changes demonstrate the potential of automated segmentation in detecting and characterising early morphological changes that may precede clinical symptoms, with the benefit of potential early intervention.

4.4 Discussion

In this chapter, the aim was to increase the accuracy of automated knee cartilage segmentation by combining predictions from a 2-label (cartilage vs background) U-Net with a 7-label U-Net to determine the type of cartilage. The dice scores are higher than previous models trained and tested on all six cartilage compartments. Dice coefficients are slightly lower than the results published by Panfilov et al. (334) and the top model from the IWOAI Knee MRI Segmentation Challenge (279). However, both methods fused the tibial and meniscus cartilage labels' small medial and lateral compartments, resulting in only four cartilage types. Interestingly, the top four competing methods showed no significant differences in DICE sore in the top 4 submissions, suggesting equivalent performance amongst the best models. When comparing DICE scores, care must be taken due to differences in the training, validation, and test data splits.

Dice across all cartilage labels increased when comparing the combined model to the 7-label model. In addition, the post-processing is very simple. It has a similar effect to SSM in removing small isolated regions in the segmentation mask and assigning the correct labels to different cartilage compartments.

Studies that have used SSM to provide segmentations for specific cartilage targets have also produced high Dice scores (333,334,354). However, this multistep process requires 2D and 3D U-Nets in addition to prior shape information for each target; this method has also failed to adequately segment cases that exhibit full-thickness cartilage denudation (354). Furthermore, SSM methods are more computationally expensive; based on estimates, using a single computational node segmentation of a single knee joint (cartilage and bone) took around 10 mins (333). Despite differences in hardware, U-Nets are computationally more efficient; the combined model takes 13.6s to segment a single MRI volume, making it more feasible to segment entire databases such as the OAI, which contains 50000 or more MRIs. Also, compared to 3D U-Nets, 2D networks are generally more efficient to train with fewer parameters.

Dice coefficients of medial cartilage was lower than that of their lateral counterparts, a finding reproduced in other studies (279,333,334,337,338). The prevalence of radiographic medial compartment OA may be 5–10 times higher than in the lateral compartment (358). The lower Dice coefficients may result from greater uncertainty in the manual segmentations in medial regions due to advanced cartilage degeneration. Notably, for manual segmentations where the volume is less than 50% of the mean for a given label (n = 14), the average Dice coefficient is 0.68, compared to the remainder of the data (n = 304) with a DICE coefficient of 0.87.

The Dice coefficient for inter-rater agreement was 0.79 in a similar dataset (278), while the model achieved a Dice coefficient ≥ 0.85 across all cartilage compartments, suggesting automated segmentation is at least as good as manually generated annotations. When comparing the corrected 2l7l U-net results to those presented by Normal et al. (332), the average volume correlation is higher (0.988 compared to 0.938), and the average absolute mean difference across compartments is lower (compared to 0.510cm^3).

The U-net architecture adopted in this chapter is 2D rather than 3D and does not account for spatial information in the through-slice direction. This reduces the constrain to match the shape of the manual annotations in this direction, which exhibits inter-slice discontinuity artefacts caused by slice-wise delineation, resulting in the recovery of missing labels in the manual annotations and smoother, more realistic boundaries around automatically generated segmentations. Compared to the 2D U-net by Panfilov et al. (339), the Dice coefficients are slightly lower. However, the model can differentiate medial from lateral components, which may be of particular importance given the role of laterality in knee osteoarthritis.

U-Nets tend to overestimate total volume segmentation in the OAI dataset; in the study by Dam et al. (353), this was 14% for medial tibial and femoral compartments and 12% across all compartments measured by Normal et al. (8). The corrected 2l7l U-net overestimated the volume by 9% across compartments. This result is similar to inter-observer reproducibility

measures by Siorpaes et al. (359) who calculated an RMSE of 9.9% for medial meniscal cartilage using DESS MRI volumes from the OAI.

In this chapter, there are no quantifiable measures to determine the causes of the differences between manual annotations and automatic segmentations. Further work could address this limitation using higher-resolution MRI, utilising either higher field strengths or post-mortem imaging. The model is limited by its lack of generalizability to clinical imaging protocols as training data consists entirely of DESS MRI but could easily extended to multi-channel inputs with different contrasts. Improvements could include augmenting training data from clinically acquired scans or using generative adversarial networks to create images with synthetic DESS contrast.

A further limitation was the inter-site variability, which was not explored in this thesis. While the original OAI study aimed to reduce differences by running matching protocols on identical hardware (same scanner make and model), differences may still arise due to hardware (i.e. deviations in shim, gradient performance, RF noise, or software version) or site (patient demographics, adherence to study protocol, radiographer and support staff training). Data harmonisation may improve outcomes by reducing the variation between images across sites.

This chapter demonstrates the deployment of the model in the OAI dataset to investigate the association of clinical symptoms with cartilage volume and the range of intensity values. Volumetric results showed an inverse correlation between cartilage volume and WOMAC scores similar to previously published values (243). Comparisons showed that the range of intensity values was more strongly associated with WOMAC pain, stiffness and functional scores in three of the six compartments, with no difference in the lateral tibial cartilage, medial meniscus and patella cartilage. This finding suggests that intensity-derived features indicative of compositional alteration may be more strongly associated with clinical symptoms than volume and, therefore, a more suitable biomarker for clinical trials.

Overall, the association between clinical symptoms, as measured by WOMAC questionnaires, and cartilage features remained weak. The highest coefficient of determination (R^2) was 0.022 for the relationship between WOMAC scores and any cartilage regions, with a maximum of 0.010 for the relationship between WOMAC pain and any cartilage regions. In comparison, previously published work by Hunter et al. (243) found an R^2 value of 0.053 for the association of patellar cartilage and knee pain. Differences may be accounted for due to the tighter study population of this previous study, with 133 postmenopausal females, compared to a broader, multi-site cohort in this study. In chapter seven, using a knee atlas to register MRIs improves the R^2 for cartilage volume associations with knee pain to 0.044, potentially correcting for differences in knee cartilage volumes due to sex.

Inclusion of sociodemographic factors may lead to a stronger correlation, the study by Hunter et Al (243) found an inter-quartile regression of $R^2 = 0.40$ between patellar volume and knee pain when adjusting for BMI. Expanding on the work in this section, chapter six investigates pain prediction in more detail, using both sociodemographic factors and high throughput radiomic features. In contrast, chapter seven includes image registration to a knee atlas and the expansion of radiomic targets for predicting knee pain.

4.5 Conclusion

This chapter used an ensemble learning approach based on a 2-label 2D U-Net to differentiate cartilage from the background, followed by a multi-label U-Net to distinguish between different cartilage types. The combined model offers rapid and accurate cartilage segmentation with favourable performance compared to previous multi-compartment segmentation algorithms tested using the OAI/iMorphics dataset. Visually, the automated segmentations appear smoother with anatomically more realistic boundaries than manual annotations. Application to over 9000 knee MRI scans upheld previously published weak association between cartilage volume and clinical scores, while comparison showed the range of intensity values may be more predictive of pain, function and stiffness in OA.

5. Prediction of joint replacement in knee osteoarthritis

5.1 Introduction

A more expansive examination of Osteoarthritis as a disease is given in chapter one. In this introduction, a summary of the salient points is provided for context. OA of the knee is the most common joint disorder, with an estimated prevalence of 21% in men and 27% in women. It is set to increase due to increasing age, obesity and physical inactivity (22) with considerable health and socio-economic burden (41). Pain and restricted movement are the main symptoms of knee OA, but individual trajectories and responses to treatment are highly varied. There is a recognised need for tailored interventions to address specific needs (360) and ineffective care of patients (361). Joint replacement surgery is a cost-effective treatment for end-stage osteoarthritis (170) with a typical response rate of 80%.

Timely identification of individuals at risk of developing severe symptoms is highly desirable for targeted preventative interventions, such as lifestyle changes (exercise, weight loss), education, self-management, and walking aids that may delay or avert the need for surgery (362). Importantly, accurate risk prediction of total knee replacement (TKR) will enable optimised trial design through risk stratification of populations and mechanistic surrogate outcome markers, which in turn will accelerate the discovery of effective treatments in knee OA. Using a strong prediction model at pre-enrolment alone affords higher power for more efficient clinical trials in a high-risk target population.

Joint replacement surgery is a common primary outcome in clinical trials for disease-modifying osteoarthritis drugs (358), but disparities in access to healthcare (e.g., ethnicity, gender, socioeconomic status) confound it (359). Alternative outcome measures are actively sought (363), and a knee pathology-based risk model seems well-suited to address these concerns.

A large body of evidence describes the longitudinal multicompartment and tissue-specific OA changes on knee MRI (12,364). Cartilage volume and thickness changes are associated with knee pain, OA progression and joint replacement (243,325,365). Manual or semi-automated segmentation limits scalability for population studies, which is now feasible using deep learning techniques (336–340,366). Previous work on predicting TKR is shown in Table 4; methods include clinical-demographics models (367), the use of cartilage thickness (368), and convolutional neural networks trained directly on the knee MRI (369).

In this chapter, radiomics features are extracted from a fully automated cartilage MRI segmentation pipeline to predict both time to joint replacement and TKR at five years in participants with symptomatic knee OA. Knee cartilage MRI and high-throughput quantitative image analysis (245,370) enabled an interpretable and quantifiable knee cartilage signature to be developed. Time-to-event risk modelling proved more informative than case-based classification, which has yet to be shown in knee OA. Multivariate Cox regression was used to calculate individualised risk and compared high and low-risk populations using Kaplan-Meier curves and proportional hazard estimates. Finally, the radiomic signature-based model was tested in an independent incident cohort.

	Follow-up	Predictors	C-Statistic (95% CI)
Yu et al. (371)	6 years	24 Clinico-demographic features	0.79 (0.78 - 0.79)
Kwoh et al. (368)	7 years	2-year thickness change in tibiofemoral joint.	0.62 (0.55 – 0.86)
Topaldi et al. (369)	5 years	Deep learning on MRI plus 27 Clinico-demographic features	KL 0,1: 0.94 (0.91 – 0.97) KL 2,3: 0.83 (0.81 – 0.85) KL 4: 0.73 (0.69 – 0.77)

Table 4: Prediction of total knee replacement
Comparative analysis of models for predicting total knee replacement in knee osteoarthritis, highlighting the diverse methodologies and datasets, such as the UK Clinical Practice Research Datalink (371) and the Osteoarthritis Initiative (369). This table underscores the varied predictive performances, particularly noting the diminished efficacy of the Topaldi et al. model in cases of advanced radiographic knee OA.

5.2 Materials and methods

5.2.1 Study Participants and clinical-demographic features

This study is a retrospective analysis of the progression and incident subcohorts of the Osteoarthritis Initiative (OAI). This includes subjects recruited between 2006 and 2015 with baseline symptomatic tibial-femoral knee OA, defined as frequent knee symptoms in the past 12 months and tibiofemoral osteophytes (349).

Participants with pre-existing knee replacements were excluded (N = 20). Subjects with symptomatic knee OA and baseline MRI were identified from their first study visit (N = 1370). To develop the radiomics signature, Knee MRI cases (including bilateral knee osteoarthritis) with outcome data at 5 years (N = 1502) is used. To evaluate the radiomics signature, an additional 274 participants are included, totalling 361 cases of knee OA with right-censored data (27 died, and 247 withdrew from the study).

Clinical and demographic features were identified from existing literature on the risk factors for total knee replacement surgery. The OAI database was parsed, and 27 corresponding items were selected. These are shown in Table 5. Missing data was imputed using k-nearest neighbours; KOOS Sport/rec questionnaires had the highest level of missing data at 23%, followed by isomeric leg strength at 12%, and remaining items had missing data at levels of ≤ 3%.

As a generalisation test, the incident subcohort of the OAI was used to evaluate model performance. Subjects in the incident cohort have no symptomatic tibial-femoral OA in either knee but are at increased risk for symptomatic knee OA. The same exclusion criteria used for the progression subcohort were used; a total of 6628 cases were included, with 1038 censored cases. Clinical and demographic data for the Incident cohort are compared to the progression cohort in Table 5.

	Progression Cohort		Incident Cohort	
	Participants without knee replacement at 5-years	Participants with knee replacement at 5-years	Participants without knee replacement at 5-years	Participants with knee replacement at 5-years
No. of participants	1218	152	3376	154
No. of knees	1692	171	5590	184
No. censored cases	361	0	1038	0
Age (years)	61.1 (9.1)	64.2 (8.4)	61.0 (9.2)	64.7 (8.3)
Income				
<10k	66 (5.4)	3 (2.0)	262 (4.0)	4 (2.2)
$10k – $25k	147 (12.1)	13 (8.6)	702 (10.6)	18 (9.8)
$25k – 50k	399 (27.8)	38 (25)	1607 (24.5)	51 (27.7)
$50k - $100k	379 (31.1)	58 (38.2)	2355 (35.5)	59 (32.1)
>$100k	287 (23.6)	40 (26.3)	1644 (25.1)	52 (28.3)
Consider total knee replacement in the next 3 years				
No	1191 (97.8)	126 (82.9)	6596 (99.5)	184 (0)
Yes, single knee	18 (1.5)	16 (10.5)	16 (0.2)	0 (0)
Yes, both knees	9 (0.7)	10 (6.6)	16 (0.2)	0 (0)
Knee Osteoarthritis Outcome Score				
Quality of life subscale (100 – 0: No problems to Extreme problems)	73.8 (20.4)	58.3 (21.2)	73.7 (20.4)	65.1 (25.1)
Sport/Rec: Difficulty kneeling, last 7 days (0 – 4 No problems to Extreme problems)	1.99 (1.11)	2.20 (1.07)	1.01 (1.04)	1.53 (1.10)
Sport/Rec: Difficulty squatting, last 7 days (0 – 4 No problems to Extreme problems)	1.02 (1.05)	1.49 (1.04)	0.97 (1.01)	1.51 (1.08)
Sport/Rec: Difficulty twisting/pivoting on injured knee† (0 – 4 No problems to Extreme problems)	0.97 (1.01)	1.51 (1.03)	0.78 (0.98)	1.16 (1.10)
WOMAC Pain severity in the last 30 days† (0 – 10 No pain to extreme pain)	4.65 (2.34)	5.40 (2.32)	1.64 (2.66)	2.7 (2.6)
Seen physician for arthritis in previous year	204 (3.1)	11 (15.1)	212 (3.2)	16 (8.7)

Physical and Mental Health Composite Scores				
SF-12 mental score (0 – 100: No problems to Extreme problems)	53.7 (0.8)	57.1 (7.3)	53.7 (7.7)	55.2 (7.4)
Knee valgus alignment (Angle in degrees) †	-0.61 (4.1)	-0.12 (4.9)	-0.60 (3.46)	-0.56 (4.17)

Table 5: Progression cohort of the Osteoarthritis Initiative
Baseline clinical and demographic factors for total knee replacement in the Progression Cohort of the Osteoarthritis Initiative were selected based on existing studies. Note. Data are presented as the mean (SD) or number (%). Measurement units or possible ranges for variables are given in parentheses. Abbreviations: OA Osteoarthritis, BMI Body mass index, PASE Physical Activity Scale for the Elderly, WOMAC Western Ontario and McMaster Universities Osteoarthritis Index, NSAID non-steroidal anti-inflammatory drug. †The mean (SD) values are per case of knee OA rather than per individual.

5.2.2 MRI Measures

Subjects underwent knee MRI at one of four separate sites (The Ohio State University, Columbus, OH; University of Maryland, School of Medicine, Baltimore, MD; University of Pittsburgh, Pittsburgh, PA; and Memorial Hospital of Rhode Island, Pawtucket, RI) to accommodate the large number of subjects to be examined over the course of the study (357). The imaging protocol was standardised across sites to minimise acquisition variability (198). The MRI sequence used for this study was a Double Echo Steady State (DESS) acquisition with a 140mm field of view, 0.7mm slice thickness, and a matrix of 384×384 with 160 slices and an acquisition time of 10.6 mins.

5.2.3 MRI Image segmentation

MRI from the first study visit of the progression cohort was downloaded from the National Institute of Mental Health Data Archive (372). Cartilage segmentations were generated for all participants (N = 1370) with symptomatic knee OA. Following the inclusion of bilateral cases, in total, 1863 knee MRIs were segmented using fully automated methods (366) by combining predictions from two models, a U-Net for the segmentation of two labels (cartilage vs background) and a multi-label U-Net for specific cartilage compartments. Data for training and evaluation comprised 176 knee DESS MRI cases from the OAI, with manual segmentation of the patellar, femoral, lateral tibial, medial tibial, lateral meniscal and medial meniscal cartilage (279). N4 Bias field correction was applied to the MRI as a pre-processing step to remove low-frequency intensity non-uniformities (351). This was based on a previously reported method (366) that achieved an excellent Dice coefficient for all cartilage locations (range 0.84 – 0.90).

5.2.4 MRI Features

Cartilage MRI features were extracted using the Standardized Environment for Radiomics Analysis package developed for MATLAB (373). Features were generated following Imaging Biomarker Standardization Initiative guidelines (245), a collaboration intended to improve the reproducibility and validity of radiomic features. 487 standardised features were extracted for each cartilage compartment, including 79 first-order features (morphology, statistical, histogram and intensity-histogram), 272 higher-order 2D features, and 136 3D features.

5.2.5 Radiomics model

To generate the radiomics signature, only knee MRI cases were included with outcome data at 5 years (N = 1502). Knee MRI radiomic features were used as the input feature space (P = 2922 features per knee) to predict knee replacement surgery at 5 years, leading to a large P – small N setting. A machine learning pipeline in MATLAB was built consisting of feature normalisation (z-score), feature reduction based on multicollinearity of predictors, pre-selection of features, selecting the best K feature with the strongest relationship to the target variable and an Elastic Net regression model (a regularised regression that linearly combines L1 and L2-norms). To reduce the multicollinearity (M), features were compared using Pearson's r, and for correlated features above a threshold, the lowest-ranked feature was removed. The elastic net was chosen as it allows the optimisation of sparsity as a hypermeter, so no assumptions of the sparsity of the discriminative ground truth were required. The free hypermeters of the machine learning pipeline were the threshold for the multicollinearity of features (M), the number of pre-selected features (K), the regularisation coefficient (λ) and the weight of regularisation (α). Hyperparameters were optimised with a grid search and binary cross-entropy loss. Grid search values for M were [0.3, 0.6, 0.9, 1.0], K varied from 10 to 100 in intervals of 5, and α values were [0.001, 0.01, 0.05, 0.1, 0.5]. For the regularisation parameter λ, 10 values were used, varying on an exponential scale between 1e-5 and 0.1. Hyperparameter optimisation was performed using 10-fold cross-validation (internal validation). Cross-validation incorporated the complete machine learning pipeline to avoid data leakage between training and test samples.

The cartilage signature was evaluated on the out-of-fold samples. Receiver Operator Characteristic (ROC) curves were created. The C-index was calculated with 95% confidence intervals (CI) obtained by bootstrap with 100 replicas. Kaplan-Meier survival curves were created for high and low-risk groups, with 95% confidence intervals calculated using Greenwood's formula (374). These groups were separated using a threshold value of the cartilage signature that maximised sensitivity and specificity. Cox proportional hazard ratios were also calculated using these high and low-risk groups but with the inclusion of participants with right censored data.

5.2.6 Clinical and demographic model

The clinical-demographic model was constructed with data from Table 5 as covariates (excluding KL score); multivariate Cox analysis was performed and features were removed in a "backwards selection" manner, based on a significance level of $p > 0.05$. A total of 12 non-

imaging features survive the backwards selection process. Individualised risk estimates from the cox regression were calculated using the following formula (375):

$$\widehat{H(t)} = 1 - [S_0(t)]^{\exp\left(\sum_{i=1}^{k} \beta_i X_i - \sum_{i=1}^{k} \bar{X}_i\right)} \quad (18)$$

Where $\widehat{H(t)}$ is the risk of joint replacement for an individual, $S_0(t)$ is the baseline survival rate at follow-up time t, where $t = 5$ years, β_i is the regression coefficient, X_i is the value of the ith risk factor, $\bar{X}_i$ is the corresponding mean, and k denotes the number of risk factors. The analysis of the radiomics model was repeated for the clinical and demographic model using risk estimated calculated using Equation 18.

5.2.7 Combined model

To evaluate if the cartilage signature is of benefit, the radiomics signature was added to the trained clinical and demographic model and continued the multivariate Cox analysis with the removal of features using backwards selection (removing feature with p>0.05). Values for the radiomics signature were obtained from the internal validation (i.e. performance on the left-out participant data). Censored cases were not included in the training of the elastic net model (due to the absence of outcome data). The radiomics signature was calculated as the mean score over the cross-validation folds for these cases. Both un-censored and censored cases were included in the multivariate Cox regression.

The analysis of the radiomics model was repeated for the combined model using risk estimates from the Cox regression calculated using Eq. 1. The clinical and demographic model and combined model were compared using a non-parametric comparison of the area under the ROC curves (376).

5.2.8 Testing in incident cases

Using data from the OAI incident Cohort, the radiomics model was evaluated, clinical and demographic model, and combined model. This provides some generalisation of the radiomics model, as no retraining of the model is performed, and the incident cohort represents individuals only at risk of developing OA and, therefore, at an earlier disease stage.

For the radiomics model, MRI cartilage segmentations were generated using the same fully automated methods used for the progression cohort. The radiomics model was tested on the Incident Cohort with outcome data at 5 years (N = 5650). The feature normalisation and elastic net regression model developed in the progression cohort were used when testing the radiomics model on the incident cohort. For cases in the incident cohort, the radiomics signature was calculated as the mean score over the cross-validation models created for the progression cohort.

Testing of the clinical and demographic model and the combined model was performed using the Cox regression models developed in the progression cohort. To avoid bias when

testing on the incident cohort, the new Cox proportional hazard ratios were calculated with the threshold values determined in the progression cohort dataset.

For testing the radiomics model, it was retrained on the entire training dataset using the optimal hypermeters for the elastic net regression. In addition to evaluating the C-statistic measures, the Cox proportional hazard ratios were calculated using the threshold value of the cartilage signature determined in the training dataset.

5.3 Results

5.3.1 Radiomics model

Optimal hyperparameters for the elastic net regression (calculated on the in-sample folds) were M = 0.3, K = 25, α = 0.1 and γ = 1e-5. Statistical test results from the internal validation (calculated on the out-of-sample folds) of the radiomics model are provided in Table 6. The ROC curves for the radiomics model are shown in Figure 31, while the Kaplan-Meier survival curves are shown in Figure 32. The features that contribute to the radiomics model are provided in Figure 34.

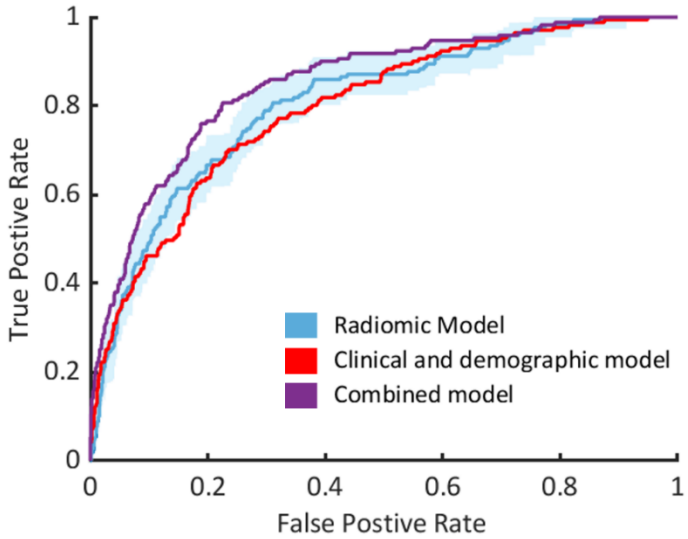


Figure 31: Prediction of total joint replacement within 5 years
The Receiver Operator Characteristic curves show three distinct models: radiomics, clinical and demographic, and a combined approach, highlighting their differential predictive capabilities for knee replacement within five years. The 95% confidence intervals are shown for the radiomics model in light blue. The AUC values are 0.81 (95% CI 0.77 – 0.84) for the radiomics model, 0.79 (95%CI, 0.75 – 0.83) for the clinical and demographic model and 0.85 (95%CI, 0.82 – 0.88) for the combined model. A combined model provides a superior prediction of total knee replacement but at the cost of added complexity with additional non-imaging factors.

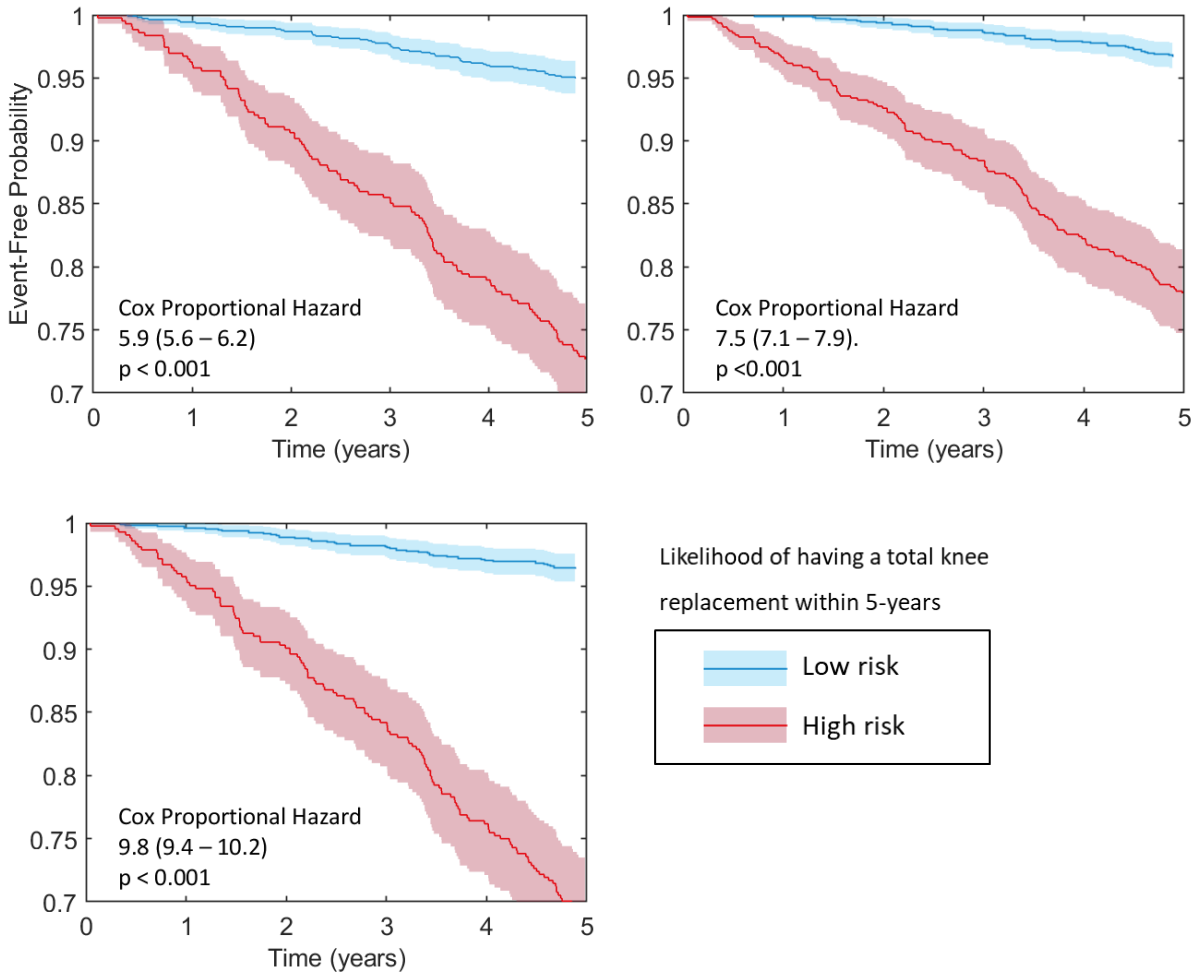


Figure 32: Longitudinal estimate of time to joint replacement
This figure illustrates the Kaplan-Meier survival analysis through high and low-risk stratifications derived from the clinical and demographic (top left), radiomics (top right), and combined (bottom) models. The figure conveys the difference in the probability of remaining joint replacement-free over a five-year trajectory for high and low-risk predicted groups. The proportional hazard is greater for the combined model, demonstrating better subject stratification.

The cartilage segmentations of three representative cases is shown in Figure 33 to visualise how changes in the knee MRI impact risk scores. The cases were chosen by normalising the variables in the clinical and demographic model (including BMI) and selecting cases closest to the mean values. The range of values for the chosen participants is Age (57 – 68), BMI (28.1 – 29.7), income ($25k –$50k), considering TKR (No), KOOS: quality of life (50 – 56.3), KOOS: Pain (66.7 – 77.8), WOMAC: pain (4), seen physician for arthritis in the previous year (No), Knee valgus alignment (0°), KOOS: difficulty squatting, last 7 days (2), KOOS: Sport/Rec: difficulty squatting, last 7 days (2), WOMAC: Pain severity in the last 30 days (3 - 6), SF12: Mental Health (51.6 – 61.2). In Figure 33, knee pathology significantly affects cartilage segmentation; full thickness tibiofemoral cartilage loss in cases 2 and 3 reduce the segmentation accuracy of the meniscus. The cases also demonstrate how the inclusion of clinical and demographic features in the combined model may modify risk scores from the

cartilage. For example, the very low score predicted by the clinical and demographic model reduces the higher score from the radiomics model.

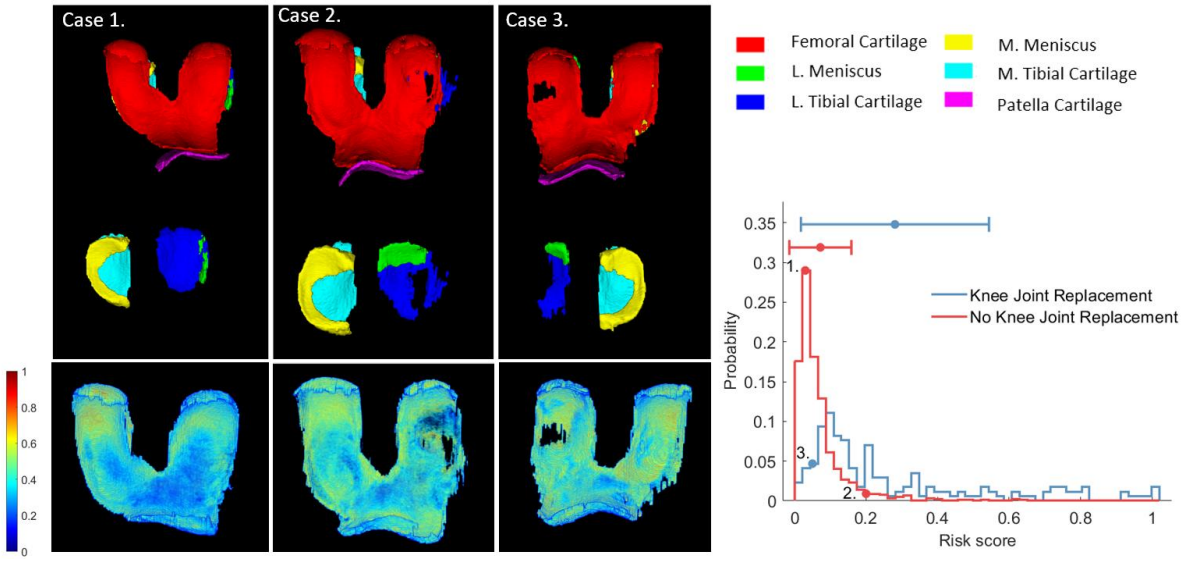
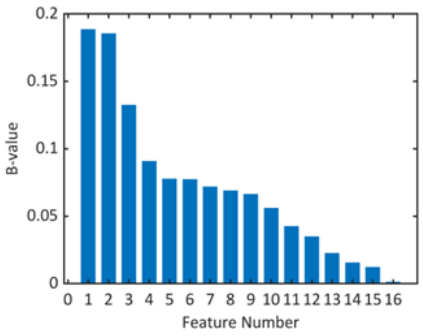


Figure 33: Total knee replacement scores

Automated cartilage segmentations (Left) for three individuals with clinical and demographic features close to the mean values across the cohort; 3D distal view through the knee joint are shown. The signal intensity variation is shown for the femoral cartilage (Bottom Left). A histogram plot (Right) for the combined model (Radiomics plus clinical and demographic features) shows the distribution of risk scores based on whether individuals undergo knee replacement within 5 years. The risk scores for the three cases are labelled; case 2 underwent knee replacement at 3.5 years. For cases 1-3, the radiomics scores were 0.05, 0.19 and 0.30, while the clinical and demographic scores were 0.07, 0.03 and 0.12, respectively.



Legend: Biomarker Characteristics

Feature	Location	Family	Imaging Biomarker
1	Lateral Meniscus	Morphology	Volume density (MVEE)
2	Femoral Cartilage	Co-occurrence matrix (2D, slice-merged)	Correlation
3	Femoral Cartilage	Co-occurrence matrix (2D, averaged)	Correlation
4	Lateral Meniscus	Co-occurrence matrix (2D, slice-merged)	Autocorrelation
5	Lateral Meniscus	Co-occurrence matrix (2.5D, merged)	Autocorrelation
6	Lateral Tibial Cartilage	Neighbourhood grey tone difference matrix (3D)	Coarseness
7	Lateral Meniscus	Co-occurrence matrix (2.5D, direction-merged)	Autocorrelation
8	Lateral Meniscus	Morphology	Area density (AEE)
9	Medial Meniscus	Intensity volume histogram	Intensity at 90% volume
10	Medial Meniscus	Neighbouring grey level dependence matrix (2D)	Low dependence high grey level emphasis
11	Medial Meniscus	Size zone matrix (3D)	Grey level non uniformity normalised
12	Medial Tibial Cartilage	Co-occurrence matrix (2D, averaged)	Joint maximum
13	Medial Meniscus	Size zone matrix (2.5D)	Small zone high grey level emphasis
14	Femoral Cartilage	Statistics	Skewness
15	Medial Meniscus	Co-occurrence matrix (2.5D, merged)	Cluster shade
16	Medial Meniscus	Intensity volume histogram	Intensity at 10% volume

Figure 34: **Feature significance of radiomic model**

This figure expositions the influential radiomic features within the predictive model, as quantified by the averaged coefficients from elastic net regression across validation folds. It identifies key imaging biomarkers critical for forecasting knee replacement and underscores how most of the predictive power is driven by a few features.

5.3.2 Clinical and demographic model

The clinical and demographic model comprised 12 features (Age, Income, considering total knee replacement in the next 3 years, KOOS: Pain subscale, KOOS: Quality of life subscale, KOOS: Difficulty squatting in last 7 days, KOOS: Difficulty twisting/pivoting on injured knee in last 7 days, WOMAC: Pain severity in the last 30 days, WOMAC: Pain subscale, Seen physician for arthritis in previous year, SF-12 mental score, Knee valgus alignment). The ROC curves are shown in Figure 31, while the Kaplan-Meier survival curves are shown in Figure 32. The C-index and Cox proportional hazard statistics are provided in Table 6.

.

	Model	C-Statistic (95% CI)	Proportional hazard ratio. (95% CI, p-value)
Progression Cohort	Clinical and Demographic	0.79 (0.75 – 0.83)	5.9 (5.6 – 6.2, p< .001)
	Radiomics	0.81 (0.77 – 0.84)	7.5 (7.1 – 7.9, p< .001)
	Combined	0.85 (0.82 – 0.88)	9.8 (9.4 – 10.2, p< .001)
Incident Cohort	Clinical and Demographic	0.75 (0.69 – 0.78)	4.8 (4.3 – 5.3, p< .001)
	Radiomics	0.82 (0.75 -0.86)	9.6 (9.2 – 10.1, p< .001)
	Combined	0.75 (0.71 – 0.80)	4.8 (4.3 – 5.2, p< .001)

Table 6: Prediction of total knee replacement.
Results provide a comprehensive evaluation of predictive models for total knee replacement over 5 years, using the progression cohort of Osteoarthritis for model development and the Osteoarthritis Initiative (OAI) incident cohort as a validation set. Interestingly, the radiomic model performs best in the incident cohort, indicating that an imaging approach may translate better into other cohorts.

5.3.3 Combined model

The combined model comprised 9 features, including the cartilage signature, and had the highest C-index and Cox proportional hazard in the training dataset (Table 6). Features removed from the clinical and demographic model include KOOS: Pain subscale, WOMAC: Pain severity in the last 30 days, WOMAC: Pain subscale, and Knee valgus alignment. The ROC curves are shown in Figure 31, while the Kaplan-Meier survival curves are shown in Figure 32. Statistical comparison of the ROC curves showed the combined model had a small but significant AUC increase of 0.06 (95%CI, 0.03 – 0.09, P< .001) compared to the clinical and demographic models.

5.3.4 Testing in incident cases

Table 6 provides the C-statistics and Cox proportional hazards for the models evaluated on the Incident Cohort. The Radiation model had the highest C-statistic and Cox proportional hazard ratio.

5.4 Discussion

The study aimed to derive a radiomics-driven model based on automatically segmented knee cartilage MRI to predict the risk of future joint replacement in subjects with knee osteoarthritis. The performance of a clinical and demographic model is compared with and without including the knee cartilage imaging signature. In the progression cohort, the C-index of the clinical and demographic model (0.79, 0.75 – 0.83) was significantly lower than the C-index of the combined model (0.85, 0.82 – 0.88), with a difference of 0.05 (0.02 – 0.08, $p < 0.01$). The Cox proportional hazard ratio for the combined models showed the greatest differential between high and low-risk groups (HR 9.8, 9.4 – 10.2, $p < .001$).

A cartilage signature is derived to predict knee joint replacement; this model uses image features extracted from knee MRI that relate to the morphology, statistical, intensity and higher-order features. The study demonstrates the added value of cartilage changes to inform the likelihood of undergoing joint replacement within 5 years, and the exemplar cases highlight how high or low clinical and demographic risk modifies this radiomics score. The predictability of TKR from the cartilage signature provides further evidence for latent joint features driving the need for TKR, which is noteworthy given the limited association of cartilage change with knee OA pain (377).

The breakdown of which features contribute to the radiomics model also provides insight into which cartilage features are most associated with the decision to undergo TKR. Of the 16 total features used across the 10-folds of the radiomics model, the menisci contributed 11 features, the femoral cartilage contributed 3, and the tibial cartilage contributed 2. The patella cartilage was absent across any of the ten folds of the radiomics model.

The decision to undertake joint replacement is complex and lacks clear objective criteria; the most successful model was expected to combine the cartilage signature with clinical and demographic data. The inclusion of knee cartilage MRI improved the prediction of a clinical and demographic model based on a comparison of the ROC curves. Moreover, Cox regression allows analysis of time-to-event data and accounts for censored cases. Comparing Cox proportional hazard ratios between models showed high-risk individuals, as defined in the combined clinical and radiomics, are nearly 10-fold more likely than low-risk participants to undergo joint replacement in the combined model compared to about 6-fold risk difference in the clinical and demographic model. This reflects a substantial improvement in risk prediction that may improve early identification of individuals at high risk of joint replacement with potential benefit for future clinical trials.

Such prediction models may enhance trial design by enriching study populations with high-risk participants. In addition, as the model is tailored to knee cartilage, the radiomics signature may have value as surrogate outcome markers in disease-modifying drugs and treatments (378), warranting further assessment.

Results from the clinical and demographic model (C-index 0.79, 0.75 – 0.83) closely match existing results (371). The radiomics model (C-index 0.81, 0.77 – 0.84) improves on findings derived from 2-year changes in cartilage thickness (368), which showed only modest performance (C-statistic of 0.55 – 0.68).

The radiomics model also compares favourably with the performance of a deep learning model based on knee MRI (369) which showed poorer performance in advanced knee OA: C-index of 0.76 (0.74 – 0.78) for moderate knee OA (KL 2-3), dropping to 0.56 (0.52 – 0.60) in

severe knee OA (KL 4). Additionally, testing on the Incident Cohort of the OAI showed that the Radiomics model was least affected by the change in the target population, representing the benefit of high-throughput quantitative image analysis for the characterisation of knee cartilage.

Limitations of this study include reliance on acquiring a DESS MRI sequence for image segmentation. The model may only be deployed in trials with this imaging sequence as part of the MRI protocol. External validation is required before adoption in future clinical trials. Currently the sequence acquisition time is 10.2 minutes, acceleration techniques may make a radiomics-based risk assessment tool for TKR more clinically feasible (379). Prospective and longitudinal validation of the cartilage signature is needed for risk stratification in future trials and clinical care. Additionally, the performance of the model is likely to depend on the severity of Knee OA and clinico-demographic features, which vary between populations.

5.5 Conclusion

In this study, benefit of cartilage MRI for predicting joint replacement at 5 years is demonstrated. The cartilage signature reflects interpretable changes in the knee MRI and provides a novel biomarker for specific joint pathology, driving the need for joint replacement. When added to a clinical and demographic model, the cartilage signature increased the Cox proportional hazard by 80%, highlighting the potential for individualised risk assessment and stratified clinical trials.

6. Pain Phenotyping with Cartilage

6.1 Introduction

Pain is the primary symptom and descriptor of osteoarthritis. In patients with symptomatic OA, pain is described as the most disabling symptom, which accounts for significant health service use and drives clinical decision-making (12). Predictive biomarkers that describe knee pain may have a wide range of applications. For example, predicting clinically significant increases in pain could improve the likelihood of successful treatment during the early stages of the disease. At this point, chronic nervous system sensitisation of pain may not be established, and changes may still prevent the progression of the disease (380).

Inclusion of predictors for pain progressors could improve clinical trial design, enabling stratification of patients for more sensitive and less costly randomised clinical trials. Localisation of peripheral pain drives may facilitate more personalised treatments. For example, targeted nerve ablation (381), compartmental knee replacement (382), preservation or resection of the infrapatellar fat pad during total knee replacement surgery (383), subchondroplasty for painful BMLs (384), and corticosteroid (385) injections for patients likely to experience pain progression.

Previously, significant investment has been focused on cartilage damage and repair, with a range of knee cartilage outcome markers used in clinical trials (236,237). These outcome measures have only shown a limited association with knee pain (238) and are often blamed for the high number of unsuccessful randomised control trials in OA.

Traditional biomarkers include volume based on segmented tissues. In an early study looking at the correlation of cartilage volume with WOMAC pain (243) only the patellar cartilage volume showed a significant correlation ($r = -0.21$). In contrast, tibial and femoral cartilage volumes did not correlate with pain. In chapter four, correlations between WOMAC pain and cartilage volumes were generally weak. Alongside quantitative measures such as cartilage thickness, radiologist scoring of cartilage defects (MOAKS) has also been used to examine the effect of cartilage loss on knee pain (377).

Cartilage thickness loss was only associated with a small increase in WOMAC Pain. There remains a radio-clinical paradox between reported knee pain and imaging metrics, which may be addressed by the use of advanced methods, such as high-throughput image radiomics. In chapter four of this thesis, cartilage volume had a maximum correlation coefficient of -0.10 with WOMAC pain across all participants in the OAI Study. However, this is only one of many possible descriptors when considering a radiomics approach.

While OA research has demonstrated high levels of heterogeneity and the involvement of many of the tissues in the knee (380), significant research efforts are still focused on using cartilage as a primary descriptor in OA (56). Using novel knee MRI cartilage assessments, i.e., radiomics, more suitable cartilage imaging biomarker candidates may yet be identified.

This chapter analyses the relationship between baseline cartilage radiomics, knee pain, and long-term knee pain trajectories previously identified by patient-centred latent class

analysis. It explores the ability of cartilage radiomic features to predict knee pain and pain trajectories to assess the suitability of cartilage radiomics for predicting knee pain.

6.2Method

6.2.1 Study participants

Data for this analysis is from the Osteoarthritis Initiative (240), using data collected at baseline. The entire cohort of the OAI and a progression subcohort is used. The entire cohort is used to assess the correlation of radiomic cartilage features with baseline knee pain and knee pain trajectories. The progression subcohort was used to develop pain prediction models as this cohort is more similar to those enrolled in OA clinical trials and patients presenting at clinics. All participants were included in the correlation analysis. However, for the model development, knees that had previously undergone total knee replacement surgery were excluded (n = 20).

The cohort consisted of 4796 participants, with 9385 knees available with suitable MRI data. The progression cohort consisted of 1370 participants, with 1863 knees available with suitable MRI data. Table 7 summarizes the clinical characteristics of these two groups.

	Entire Cohort	**Progression Cohort**
No. of participants	4796	1370
No. of knees	9385	1863
Kellgren and Lawrence index		
0	3466 (36.9)	136 (7.3)
1	1907 (20.3)	228 (12.2)
2	2501 (26.6)	740 (39.7)
3	1223 (13.0)	557 (29.9)
4	286 (3.0)	202 (10.8)
Age (years) †	61.1 (9.2)	61.3 (9.0)
Sex		
Male	3898 (41.5)	780 (41.9)
Female	5485 (58.5)	1083 (58.1)
Obesity (BMI)	28.6 (4.8)	30.4 (5.0)
WOMAC pain† (0 – 20: No problems to Extreme problems)	2.4 (3.2)	5.2 (4.0)

Table 7: Clinical and Demographic features of the Osteoarthritis Initiative
A comprehensive overview of the baseline clinical and demographic characteristics of participants from the Osteoarthritis Initiative. While demographic factors are similar for the cohorts, the progression cohort has higher levels of knee OA, demonstrated by the increased KL index scores. Data are presented as the mean (SD) or number (%). Measurement units or possible ranges for variables are given in parentheses. Abbreviations: BMI Body mass Index, WOMAC Western Ontario and McMaster University Osteoarthritis Index. †The mean (SD) values are per case of knee OA rather than per individual.

6.2.2 Outcome measures

Two outcome measures were used for knee pain. Firstly, as a cross-sectional measure, baseline pain measure was used. Secondly, knee pain trajectories were used based on previously published group-based trajectory modelling using data from the OAI (386). The model used latent class growth analysis to identify left and right knee pain trajectories based on annual repeated pain assessments over a 9-year follow-up. For both measures, the WOMAC subscale was used; it has a range of 0 – 20, where 0 represents no pain and 20 extreme pain (387).

For baseline pain scores, the classes provided by the 5-point Likert scale used in the WOMAC questionnaire are used: None (0), Mild (1), Moderate (2), Severe (3), and Extreme (4). Pain scores were grouped as follows: Mild (0-3), Moderate (4-6), Severe (7-9), Extreme (≥10). This was a pragmatic approach to maintaining a reasonable proportion of subjects in each group based on the distribution of pain scores (Figure 35).

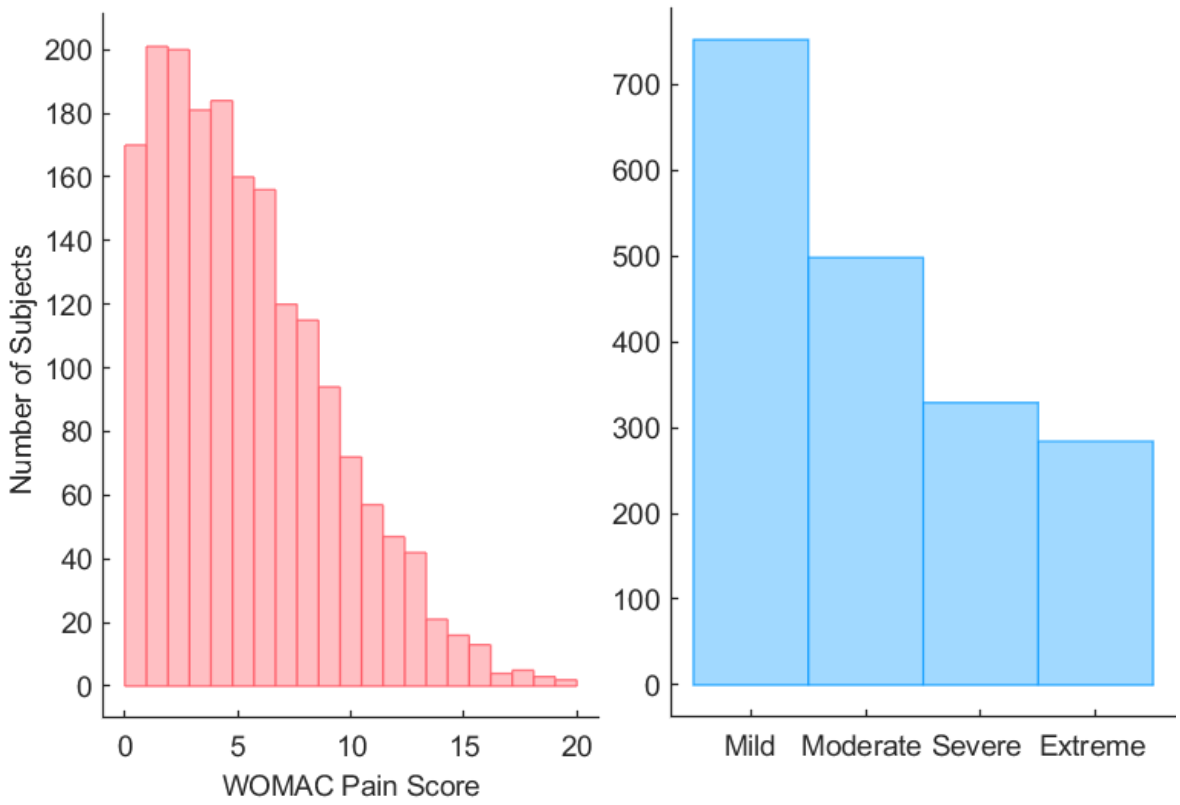


Figure 35: Pain scores in progression cohort of the Osteoarthritis Initiative
The distribution of WOMAC pain scores among the OAI progression cohort of the Osteoarthritis Initiative, categorised into four distinct groups based on severity. Left: Native distribution; Right: grouped pain scores: mild (0-3), Moderate (4-6), severe (7-9), and Extreme (≥10). This approach maintains a reasonable proportion of subjects in each group for classification.

For knee pain trajectories, there were six previously identified pain trajectory groups for each knee (Figure 36); the first group (reference) presented no pain phenotype, the second had low-fluctuating pain phenotype, the third mild-increasing, the fourth and fifth had moderate-treatment-sensitive pain, and the sixth group was severe-treatment-insensitive pain phenotype. The reliability of group-based trajectories indicated by posterior probabilities was very good, ranging from 0.80 to 0.90.

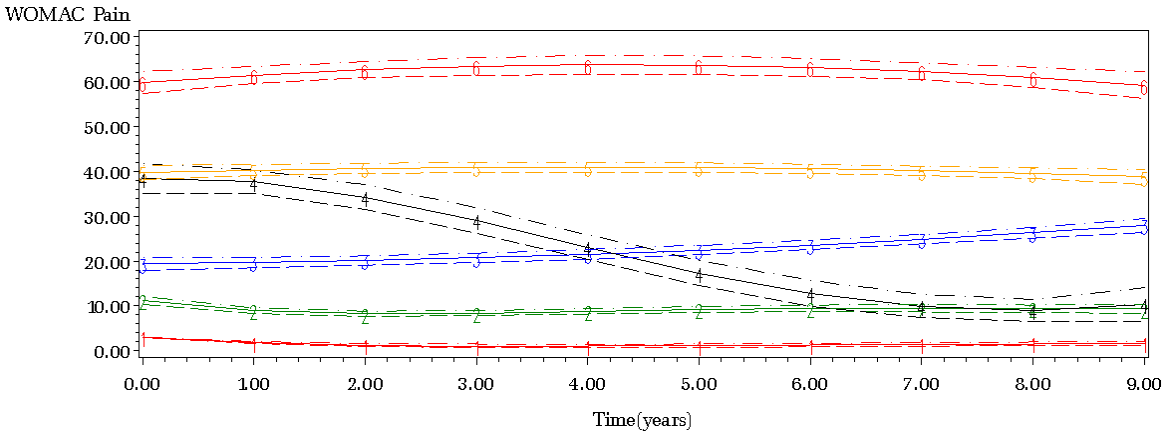


Figure 36: Knee Pain trajectories based on latent class analysis

The six knee pain trajectory groups identified through latent class analysis, showcasing the dynamic nature of knee pain over time. These were derived from the Osteoarthritis Initiative using group-based trajectory modelling and 9-year annual pain assessment by WOMAC questionnaire (score 0-100, i.e., no pain to the worst imaginable). Group 1 (red) no pain phenotype, group 2 (green) low-fluctuating pain, group 3 (blue) mild-increasing pain, groups 4 and 5 (black and yellow) present moderate-treatment-sensitive pain, and group 6 (amber) severe-treatment-insensitive pain phenotypes. Reproduced from Radojčić *et al*. 2020 (386).

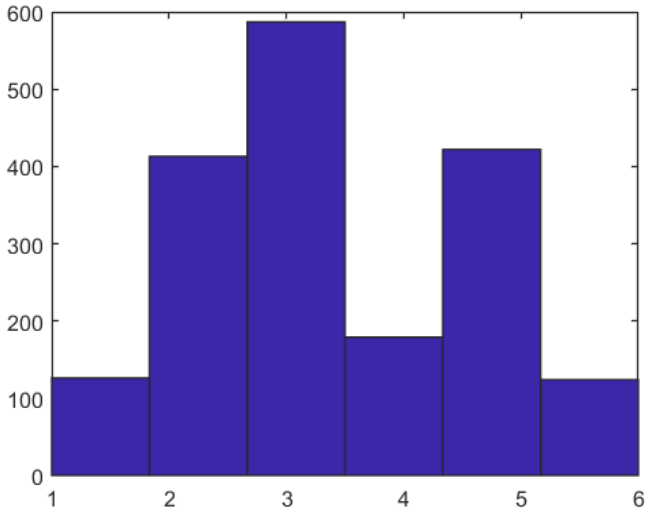


Figure 37: Pain Trajectories in Progression Cohort of the Osteoarthritis Initiative
The distribution of knee pain trajectories within the progression cohort of the Osteoarthritis Initiative. The groups 1 to 6 range from a 'no pain phenotype' to a 'severe-treatment-insensitive pain phenotype'. This breakdown provides a better understanding of the number of people in each pain trajectory group. The lowest and highest pain groups have fewer participants, with the majority in the central low to moderate groups.

6.2.3 Radiomic features

Automated knee cartilage segmentation is performed using an ensemble of convolutional neural U-nets, as described in chapter four. Radiomic features were extracted for each cartilage label, including lateral and medial tibial cartilage and meniscus, patella cartilage and femoral cartilage as described in chapter five. A total of 487 radiomic features from the Image Biomarker Standardization Imitative (IBSI) (260) were extracted for each cartilage label, resulting in a total of 2922 radiomic features per knee. Features underwent z-score normalisation (subtracting the mean and dividing by the standard deviation).

6.2.4 Correlation analysis

The entire OAI cohort was used to evaluate the correlation between radiomic features and baseline and longitudinal knee pain. This was a preliminary investigation before modelling knee pain/pain trajectories.

As part of an exploratory analysis of these features, two features are examined in more detail. Cartilage volume was selected as the first biomarker to investigate, as it has previously been used to assess knee cartilage in OA (243). The biomarker with the highest overall r-value is identified when looking for associations between knee cartilage, pain and pain phenotype.

The Pearson's Correlation Coefficient is calculated for each anatomical location between cartilage variables and baseline WOMAC knee pain and pain trajectories. Any significant

differences in correlation coefficients for both radiomic biomarkers are tested using Fisher r-to-z transformations. Further, the differences in cartilage radiomics biomarkers between pain trajectory groups is examined using analysis of variance and Tukey-Kramer multi-comparison correction.

To provide context for cartilage volume's performance as a pain/pain trajectory biomarker, the cumulative distribution functions of the mean absolute r-value across anatomical locations was plotted. This was used to calculate the performance centile (%) of cartilage volume to show how well it performs as a biomarker compared to other IBSI features.

6.2.5 Pain prediction models

The progression cohort of the OAI was used to develop pain prediction models. Separate pain prediction models are generated for baseline knee pain and knee pain trajectory. Two types of classification models are generated for each outcome: traditional machine learning with principle component analysis and a neural network approach. For each model, one-versus-many AUC values were calculated.

As the radiomic features from the IBSI were developed for high-throughput analysis, principal component analysis (PCA) can be used to develop a suitable model to predict knee pain and pain trajectories. Using PCA features, two linear machine learning classifiers were trained: linear discriminant analysis and linear support vector machine.

Additionally, a shallow neural network was trained for classification using the radiomics features as predictors. The network is designed with a single hidden layer with 10 activation nodes. To improve the convergence of the shallow network, a hyperbolic tangent sigmoid activation function is used (388). The network is trained with a cross-entropy loss function and gradient descent with a momentum optimisation training function. A softmax activation function is used in the last layer to perform classification. The network is trained via k-fold cross-validation with k equals 10 within each fold. The split between training, validation, and testing was 70%, 20%, and 10%, respectively.

6.3 Results

6.3.1 Correlation analysis

The performance of the cartilage volume against other radiomic features is shown via cumulative distribution plots for knee pain and pain trajectory (Figure 38). When evaluating correlation coefficients, 29 features have identical values and are excluded from the analysis. Cartilage volume was in the 43.5th centile for correlations with knee pain and 55.8th centile for correlations with knee pain trajectory. 3D run length variance was in the 99.6th centile for correlation with knee pain and 100th centile for correlation with pain trajectory.

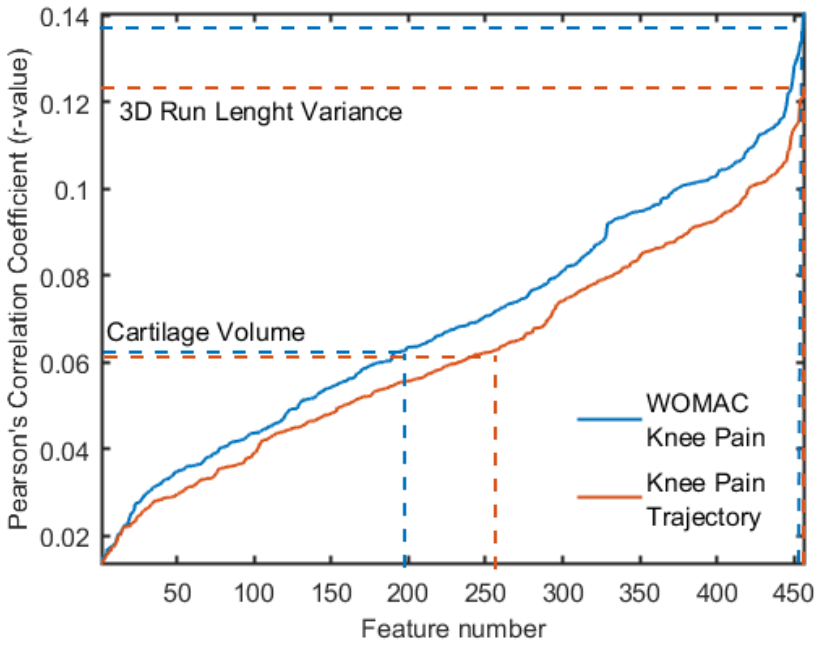


Figure 38: Feature correlation with knee pain and knee pain trajectory
This is a critical visualisation through cumulative distribution plots of the correlation coefficients between radiomic features and knee pain scores and trajectories. These distributions can be used to assess the relative performance of various cartilage radiomic features in their correlation with the severity and progression of knee pain. The Correlation coefficients shown are the mean absolute r-values averaged across anatomical locations. The dashed line shows the performance of cartilage volume at the lower end of the distribution and 3D run length variance at the upper end of the distribution. Certain features, like 3D run length variance, may serve as robust biomarkers for knee osteoarthritis pain with better correlation to outcomes compared with traditional biomarkers such as cartilage volume.

A weak correlation was found between all 3D run length variance variables with baseline knee pain and 9-year pain trajectories. Significant correlations existed for cartilage volumes, except for femoral cartilage (Table 8).

Correlation coefficients of baseline pain and pain trajectories with cartilage 3D run length variance of the femur and lateral meniscus, pain trajectory, and medial meniscus significantly differed from the correlation coefficients of the same location cartilage volume and pain variables. Significant differences were observed in cartilage volume and 3D run length variances between no pain phenotype and different knee pain trajectories, with the lateral tibia and patella variables showing the most consistent differences (Figure 39 and Figure 40).

Anatomical location	*Femur*	*Patella*	*Medial Tibia*	*Lateral Tibia*	*Medial Meniscus*	*Lateral Meniscus*
Baseline Pain						
Cartilage Volume						
r-value	-0.01	**-0.10**	**-0.07**	**-0.10**	**-0.06**	**-0.04**
p-value	0.52	<0.01	<0.01	<0.01	<0.01	<0.01
3D Run Length Variance						
r-value	**-0.12**	**-0.10**	**-0.11**	**-0.14**	**-0.12**	**-0.14**
p-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Significant difference	**Yes**	No	No	No	No	**Yes**
p-value	**(<0.01)**	(0.83)	(0.22)	(0.19)	(0.07)	**(<0.01)**
Pain trajectory						
Cartilage Volume						
r-value	0.01	**-0.13**	**-0.05**	**-0.09**	**-0.07**	**-0.04**
p-value	0.45	<0.01	<0.01	<0.01	<0.01	<0.01
3D Run Length Variance						
r-value	**-0.11**	**-0.15**	**-0.10**	**-0.15**	**-0.14**	**-0.16**
p-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Significant difference	**Yes**	No	No	No	**Yes**	**Yes**
p-value	**(<0.01)**	(0.48)	(0.18)	(0.10)	**(0.02)**	**(<0.01)**

r-value – Pearson's correlation coefficient.
Significant differences between correlation coefficients were calculated using Fisher's r-to-z transformation.
Bold results are significant.

Table 8: correlation of cartilage features with knee pain and knee pain trajectory
This table analyses the relationship between cartilage features, specifically cartilage volume and 3D run length variance, and their correlation with baseline knee pain and knee pain trajectories. Pearson's correlation coefficients and corresponding p-values show that most features correlate weakly with knee outcomes. According to Fisher's test, 3D run length variance appears to be more significant than cartilage volume in nearly half the cartilage regions.

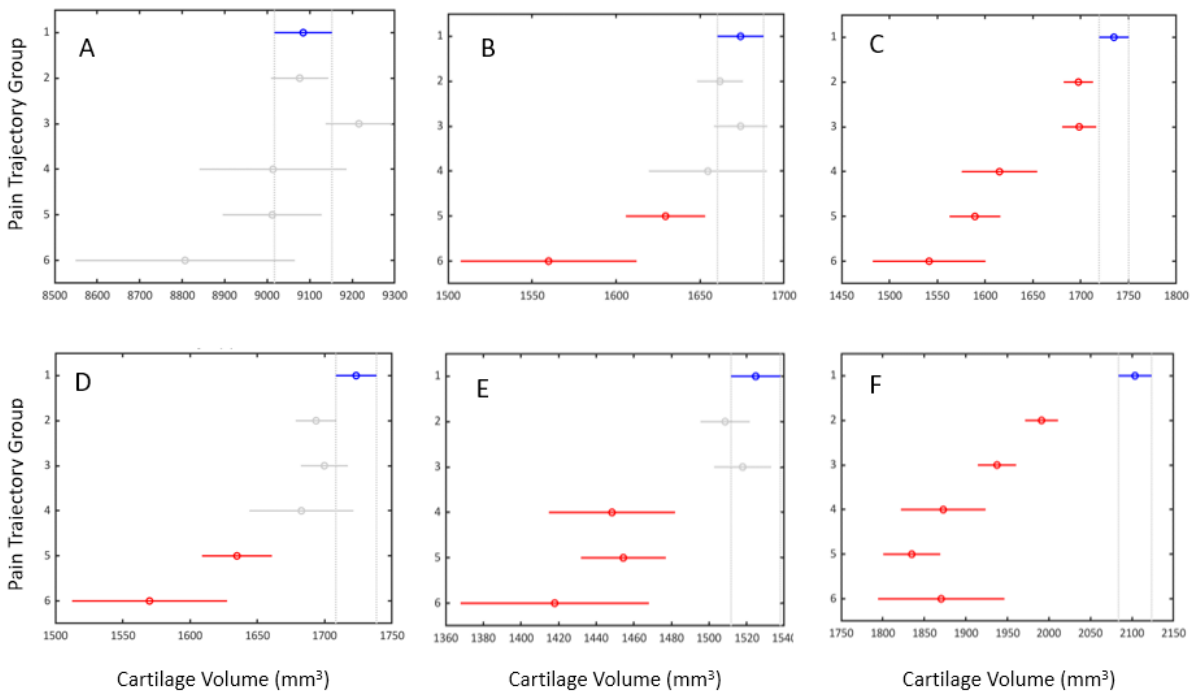


Figure 39: Cartilage volume association with knee pain trajectory
Differences between knee pain trajectory-groups in cartilage volume at different anatomical sites based on analysis of variance and Turkey Kramer multi-comparison correction. Presented are means and 95% confidence intervals for cartilage volumes per pain trajectory group, and comparisons with group 1(blue) as the reference and significant (red) or not significant (grey) differences based on the overlap of the confidence intervals. Pain trajectory groups: 1 – no pain phenotype, 2 – low-fluctuating pain, 3 – mild-increasing, 4 and 5 – moderate-treatment-sensitive pain, and 6 – severe-treatment-insensitive pain phenotype. A) Femur; B) Lateral Meniscus; C) Lateral Tibia; D) Medial Meniscus; E) Medial Tibia; F) Patella. For lateral tibial cartilage and patellar cartilage, there are significant differences for all pain trajectory groups compared to the no-pain phenotype, with over 10% cartilage loss in the most extreme pain trajectory groups.

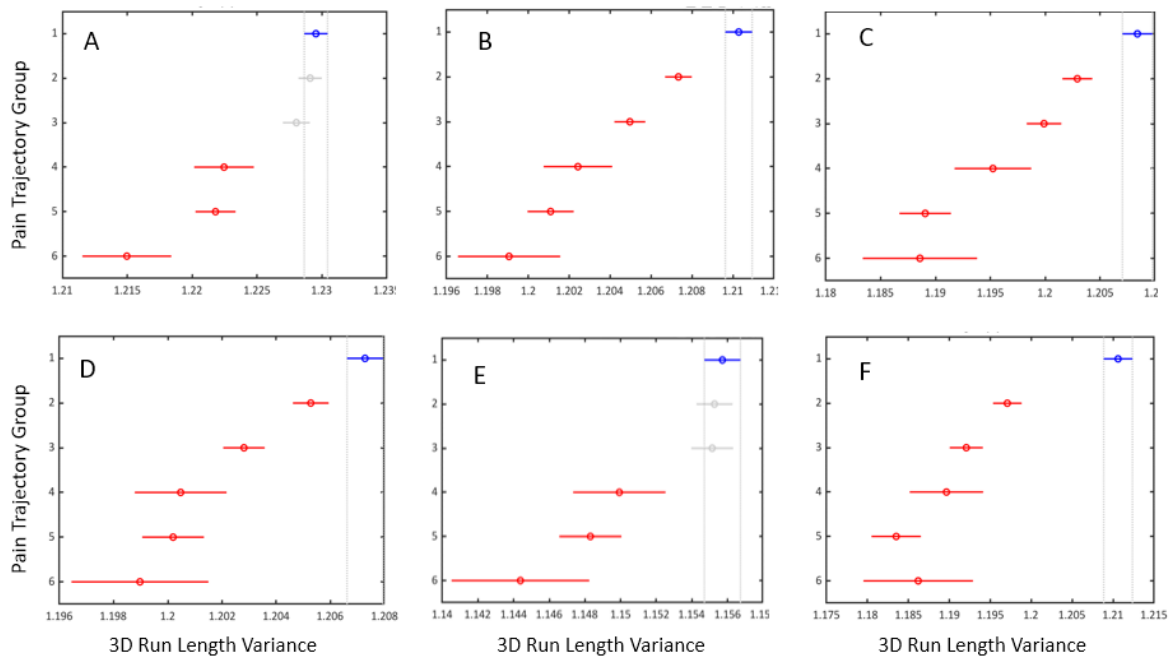


Figure 40: Cartilage 3D run length variance association with knee pain trajectory
Differences between knee pain trajectory-groups and the cartilage radiomic feature '3D run length variance' at different anatomical sites based on analysis of variance and Turkey Kramer multi-comparison correction. Presented are means and 95% confidence intervals for cartilage 3D run length variance per pain trajectory group, and comparisons with group 1(blue) as the reference and significant (red) or not significant (grey) differences based on the overlap of the confidence intervals. Pain trajectory groups: 1 – no pain phenotype, 2 – low-fluctuating pain, 3 – mild-increasing, 4 and 5 – moderate-treatment-sensitive pain, and 6 – severe-treatment-insensitive pain phenotype. A) Femur; B) Lateral Meniscus; C) Lateral Tibia; D) Medial Meniscus; E) Medial Tibia; F) Patella. All but two pain trajectory groups significantly differ in feature scores compared with the no-pain group, suggesting it is a more discriminative feature compared with cartilage volume.

6.3.2 Pain prediction models

Both models relied on the same set of PCA features to predict knee pain and pain trajectory. The explained variance for the first 10 features following PCA analysis is shown in Figure 41. Subsequent components have individually explained variance of < 2.3%.

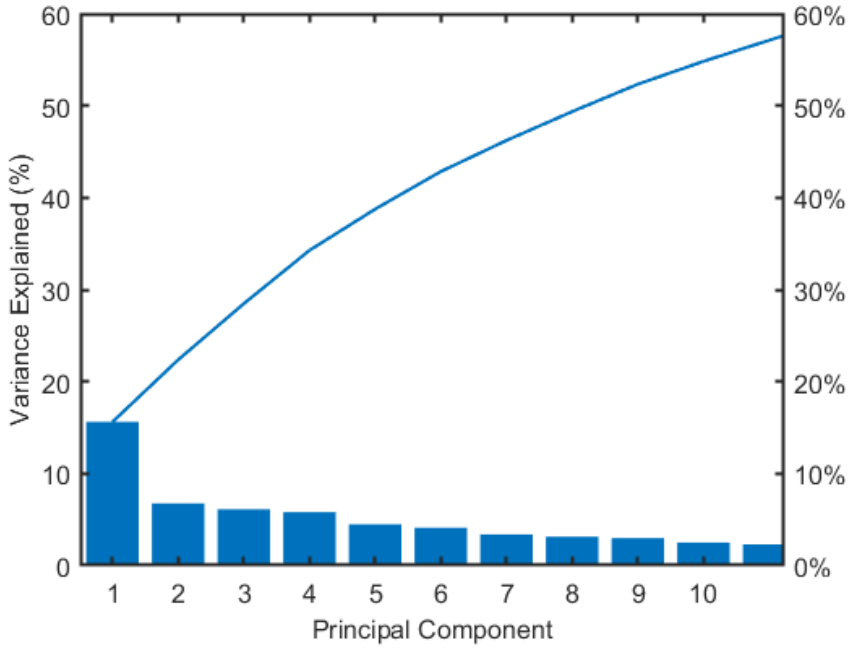


Figure 41: Principle component analysis of knee cartilage radiomics features
A pareto chart representing the variance explained by the first 10 principal components derived from a comprehensive set of 2922 knee cartilage radiomic features encompassing the Femoral, Lateral Meniscal, Lateral Tibial, Medial Meniscal, Medial Tibial, and Patella cartilage. PCA can significantly reduce dimensionality while still retaining 57% of explained variance with 10 features. The aim of reducing the dimensionality is to improve classification, given the large number of predictors vs the small number of outcomes.

6.3.2.1 Pain score

The performance of the linear classifiers for predicting OA pain groups is presented in Table 9. The model with the highest average area under the curve score (AUC = 57.7) was the linear discriminant analysis with a PCA level that explained 75% of the total variance. The Receiver Operator Characteristic (ROC) curve for this model is shown in Figure 42. Individual one-versus many AUC scores were group 1: 0.58, group 2: 0.62, group 3: 0.52, group 4: 0.56, group 5: 0.58, group 6: 0.60.

PCA: explained variance (%)	Number of Features	Area Under the Curve	
		LDA - Mean (Std)	SVM - Mean (Std)
99	305	0.53.5 (0.037)	0.555 (0.039)
95	131	0.55.8 (0.044)	0.540 (0.042)
90	60	0.57.8 (0.033)	0.535 (0.039)
75	23	**0.58.3 (0.039)**	0.548 (0.064)
50	9	0.57.5 (0.037)	0.500 (0.028)

Table 9: Predicting baseline WOMAC knee pain using machine learning
This table shows the performance of machine learning models, specifically Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM), to predict baseline WOMAC knee pain scores from a dataset of 2922 radiomic features. Average AUC values are evaluated for each pain group calculated as a one-verses-many comparison. The model with the highest average AUC is shown in bold. LDA = Linear discriminant analysis, SVM = Support vector machine, Std = Standard deviation. Overall performance very poor and is close to that of a random classifier (AUC = 0.5).

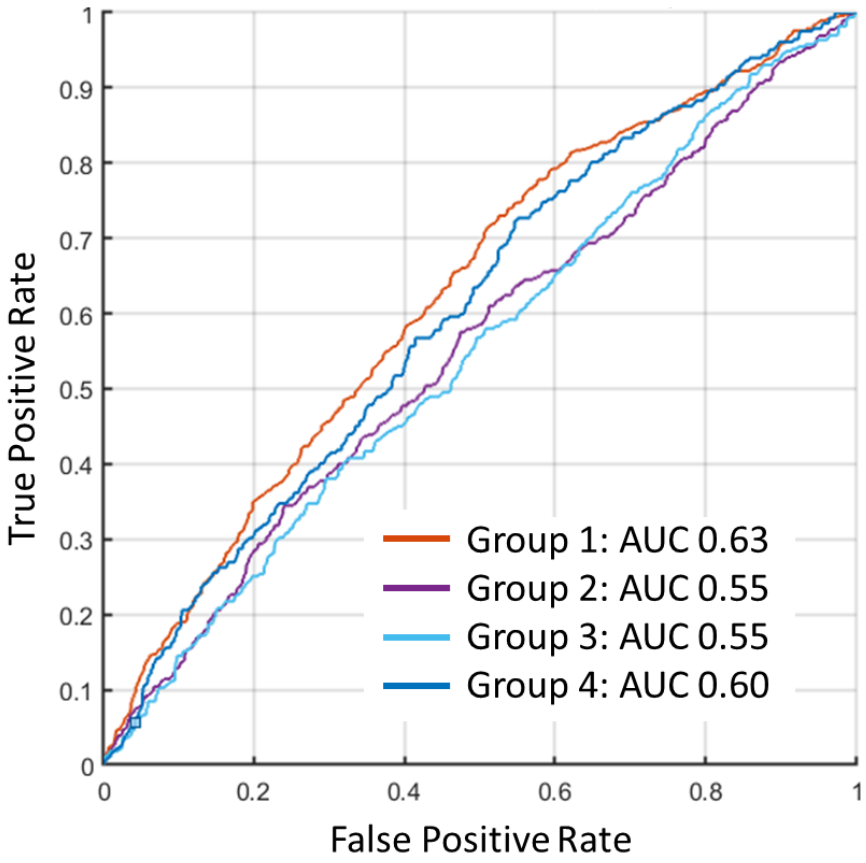


Figure 42: Predicting baseline WOMAC knee pain using machine learning
Receiver operator characteristic curve for predicting baseline WOMAC knee pain. AUC scores are one-versus-all for each group. Pain Scores are: group 1: Mild (1-3), group 2: Moderate (4-6), group 3: Severe (7-9) and group 4: Extreme (≥10). Performance of these models are limited, but demonstrate a greater ability to classify the extremities of the dataset – both the mild and extreme pain groups.

Figure 43 shows the ROC curves for predicting OA baseline pain using a shallow neural network with 10-fold cross-validated models. Individual one-versus-many AUC scores were group 1: 0.65, group 2: 0.59, group 3: 0.59, and group 4: 0.67. The mean AUC score across the groups was 0.6259 (0.040).

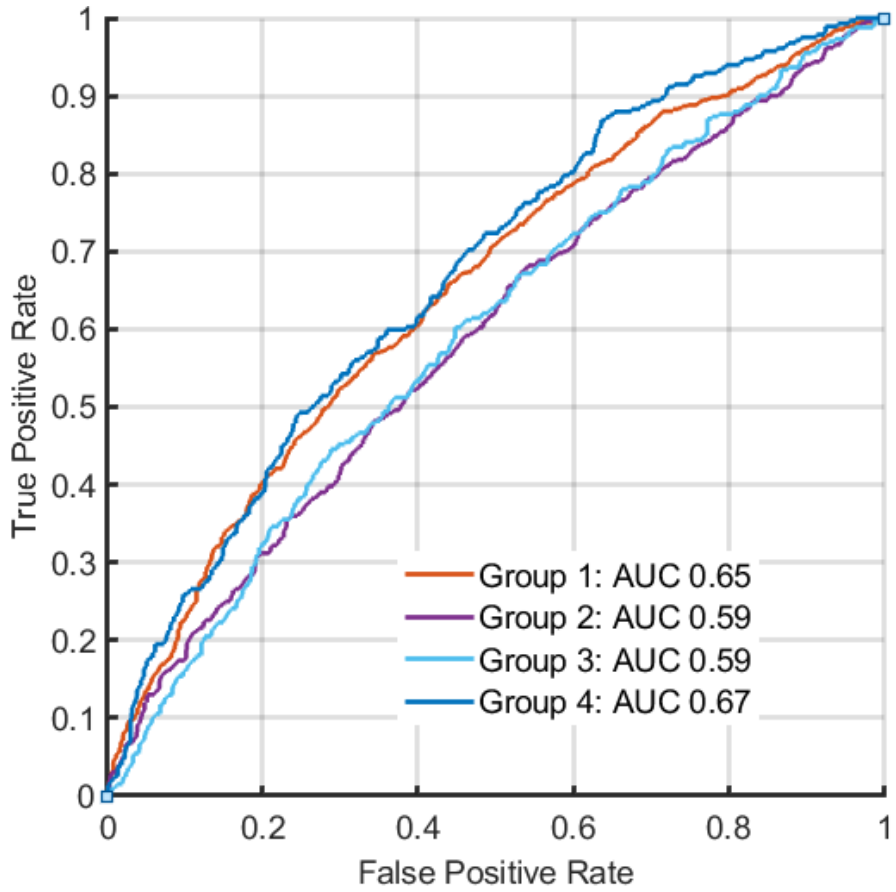


Figure 43: Predicting baseline pain scores using neural networks
Receiver operator characteristic curve for predicting baseline WOMAC knee pain. AUC scores are one-versus-all for each group. Pain Scores are: group 1: Mild (1-3), group 2: Moderate (4-6), group 3: Severe (7-9) and group 4: Extreme (≥10). Neural networks demonstrate a marginal improvement in performance versus the machine learning approaches but classification scores remain very low.

6.3.2.2 Pain trajectory

The performance of the linear classifiers for predicting OA pain trajectory are presented in Table 10. The model with the highest average area under the curve score (AUC = 57.7) was the support vector machine with a PCA level that explained 50% of the total variance. The Receiver Operator Characteristic (ROC) curve for this model is shown in Figure 44. Individual one-versus many AUC scores were group 1: 0.58, group 2: 0.62, group 3: 0.52, group 4: 0.56, group 5: 0.58, group 6: 0.60.

PCA: explained variance (%)	Number of Features	Area Under the Curve	
		LDA - Mean (Std)	SVM - Mean (Std)
99	305	0.552 (0.020)	0.557 (0.029)
95	131	0.553 (0.040)	0.570 (0.039)
90	60	0.557 (0.032)	0.575 (0.036)
75	23	0.575 (0.036)	0.568 (0.038)
50	9	0.567 (0.037)	**0.577 (0.034)**

Table 10: Predicting knee pain trajectory using machine learning
This table shows the performance of machine learning models, specifically Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM), to predict baseline WOMAC knee pain scores from a dataset of 2922 radiomic features. Average AUC values are evaluated for each pain trajectory calculated as a one-verses-many comparison. The model with the highest average AUC is shown in bold. LDA = Linear discriminant analysis, SVM = Support vector machine, Std = Standard deviation. Overall performance very poor and is close to that of a random classifier (AUC = 0.5).

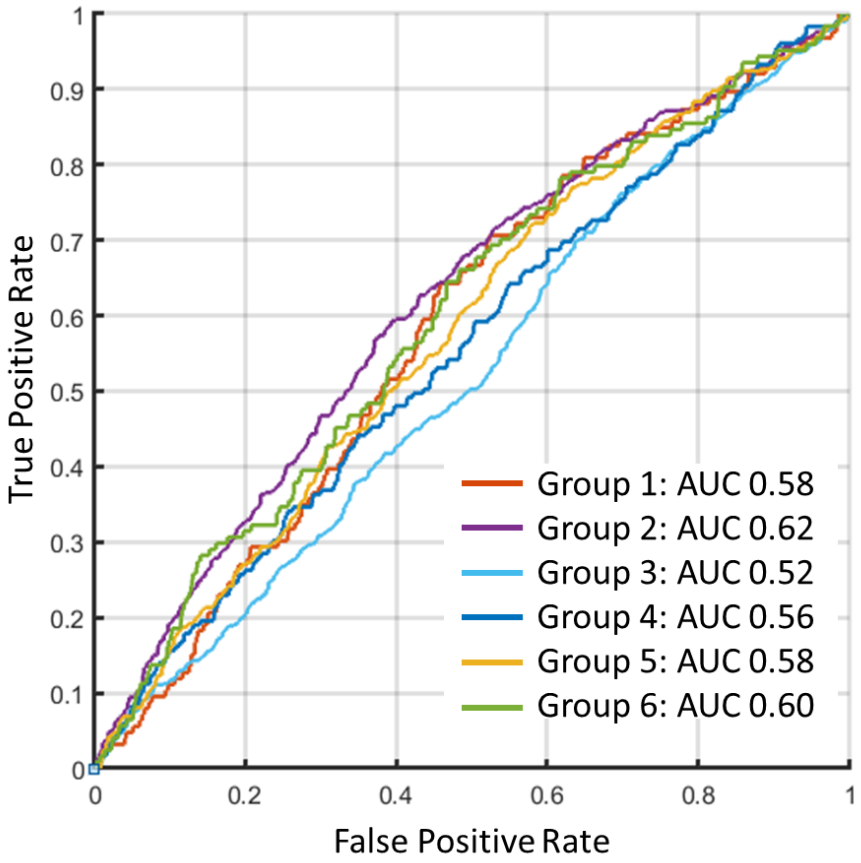


Figure 44: Predicting knee pain trajectory using machine learning
Receiver operator characteristic curve for predicting pain trajectories. AUC scores are one-versus-all for each group. Pain trajectories are group 1: no pain phenotype, group 2: low-fluctuating pain, group 3: mild-increasing pain, groups 4 and 5 present moderate-treatment-sensitive pain, and group 6: severe-treatment-insensitive pain phenotypes (386). Performance of these models are limited, their ability to classify central groups 3 and 4 is lower than the groups at the extremities (1,2, 5, and 6).

Figure 45 shows the ROC curves for predicting OA pain trajectory using a shallow neural network with 10-fold cross-validated models. Individual one-versus-many AUC scores were group 1: 0.66, group 2: 0.63, group 3: 0.57, group 4: 0.66, group 5: 0.62, and group 6: 0.67. The mean AUC score across the groups was 0.635 (0.037).

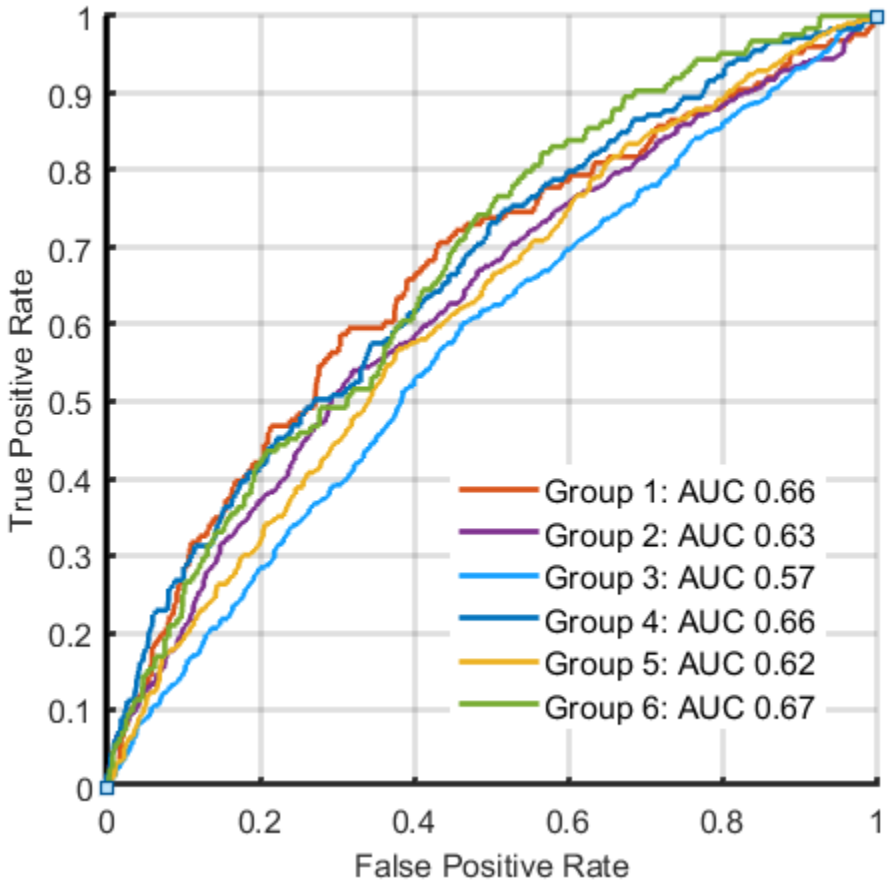


Figure 45: Predicting pain trajectories using neural networks

Receiver operator characteristic curve for predicting pain trajectories. AUC scores are one-versus-all for each group. Pain trajectories are: group 1: no pain phenotype, group 2: low-fluctuating pain, group 3: mild-increasing pain, groups 4 and 5 present moderate-treatment-sensitive pain, and group 6: severe-treatment-insensitive pain phenotypes (386). Neural networks demonstrate a marginal improvement in performance versus the machine learning approaches but classification scores remain very low.

6.4 Discussion

Knee cartilage radiomics can differentiate long-term knee pain phenotypes to some extent. Cartilage is a non-innervated tissue of the osteochondral structure. Presumably, the level of cartilage damage corresponds to the intensity of pain by sensitising other innervated joint structures, either the denuded bone directly or other joint structures indirectly by degradation products. Radiomics variables like cartilage volume and volume texture parameter (3D run length variance) are assumed to capture the cartilage damage and perform better as imaging biomarkers of OA than conventional markers such as volume.

A weak correlation was seen with cross-sectional and longitudinal pain outcome. However, even the best-performing radiomic feature only has a slightly stronger correlation than the volume. Though these were relatively small, the Turkey Kramer multi-comparison test showed some significant differences between groups. Also, cartilage volume had higher variance (wider confidence intervals) than the volume texture at all anatomical sites and increasingly with the severity of the pain phenotype.

Pain trajectories were mostly stable longitudinally, except for Group 4, which significantly improved pain over time. Data from the Supplementary information (386) showed individuals in this group were 2.5 and 4.5 times more likely to receive TKR than the other groups. This may be a strong explanation for why changes are seen in this group over time. In subsequent analysis, longitudinal pain trajectories are unlikely to provide any additional clinical insight, as other than for group 4, the trajectories do not significantly deviate from baseline pain values.

When using high throughput radiomic features to predict cross-sectional and longitudinal knee pain, the AUC values suggest low to poor performance in a one-versus-all approach. For pain trajectory prediction, the AUC score from linear machine learning models was improved using neural network classification (0.58 vs 0.64). Similarly, the AUC score from linear machine learning models was improved for predicting baseline WOMAC pain using neural network classification (0.58 vs 0.63). However, these improvements alone are insufficient to achieve the goal of accurate OA phenotyping for personalised medicine using knee cartilage. To improve the ability of radiomic features to differentiate OA pain phenotypes, additional structures known to contribute to knee pain should be included in subsequent modelling.

6.5 Conclusion

In this section, the relationship between baseline MRI knee cartilage features and baseline pain was examined as well as long-term knee OA pain trajectories and, consequently, the potential of cartilage radiomics to contribute to personalised medicine in OA. However, cartilage alone does not satisfactorily predict knee pain, suggesting that other knee targets may be needed to improve OA pain phenotyping.

7 Pain Phenotyping with other structural targets

7.1 Introduction

The investigation into cartilage as a pain target in chapter six revealed associations between cartilage and knee pain. However, these had limited predictive ability with modest AUC values. Including new targets covering anatomical regions contributing to knee pain pathology could produce better results.

Deep learning on whole images has recently successfully predicted pain progression in the osteoarthritis initiative cohort using knee radiographers (389). When predicting progression in the entire OAI cohort, defined as an increase in WOMAC Pain (9/100 points) at two time points in a 9-year follow-up with annual testing, an AUC of 0.77 was achieved. This increased to 0.81 when including clinic-demographic features.

Siamese convolutional neural networks have been used to compare MRI scans in knees with frequent unilateral knee pain to the contralateral knee, achieving an AUC of 0.808 (390). This binary classification task was a within-person comparison, with participants contributing a healthy and painful knee, an 'optimal' scenario for comparison. A radiologist reviewed a subset of the saliency maps, and effusion-synovitis generated 89% of activation hotspots. Other contributions were BMLs (5%), infrapatellar fat pad (4%), cartilage loss (2%) and meniscal damage (1%). This may indicate a tendency of the network to pick the most prominent feature in paired samples of painful and healthy knees, an approach that may not scale well to pain prediction in a symptomatic cohort.

Other pain prediction approaches include image features from T2 cartilage mapping for predicting patients with progression at three years vs asymptomatic matched controls (391,392). In these studies, 71 – 76% accuracy was achieved with image texture features, increasing to 84% using signal variation measures. Similarly to the Siamese CNN study (385), the cohort selection represents a scenario that maximises performance by comparing healthy individuals to those suffering from painful OA.

The prediction of knee pain in symptomatic patients is likely more difficult due to the high prevalence of OA pathology across the cohort. In chapter six, cartilage alone had limited success in predicting knee pain. Including new targets could improve the ability to localise the source of knee pain, allowing better stratification of patients. This imaged-based pain phenotyping may lead to more personalised treatment options. Examples of these given previously include targeted nerve ablation (381), compartmental knee replacement (382), preservation or resection of the infrapatellar fat pad during total knee replacement surgery (383), subchondroplasty for painful BMLs (384), and corticosteroid (385) injections for patients likely to experience pain progression.

Drawing on the benefits of the exploratory analysis performed in chapter six, the focus of chapter seven is to extend the cartilage pain model to include additional clinically relevant targets. The U-net method for segmentation used in chapter four would enable accurate segmentation of these new regions using MRI. However, the lack of annotated data prevents such a supervised learning approach without significant time investments.

Atlas-based image registration provides an alternative to segment new tissues in the knee. While atlas-based segmentation is less accurate compared with convolutional neural networks (285), this registration method has previously been used for knee pathology-pain mapping of bone marrow lesions (282–284). Additionally, registering images to an atlas performs a size normalisation, which may improve the effectiveness of radiomic features by removing the variation introduced by differences in knee size.

Two new targets (bone marrow lesions and effusion-synovitis in the infrapatellar fat pad) were added based on their associations with knee OA Pain. Both targets are potential earlier markers of knee OA compared to cartilage, enabling treatment when the disease may still be reversible (393). A nested case-controlled longitudinal MRI study found effusion-synovitis was a significant risk factor for the development of OA (389). It has been recognised that BMLs develop earlier than cartilage damage (390) and, once present, rarely regress (394).

While studies suggest that cartilage defects correlate strongly with bone marrow lesions, in older adults, BMLs are independently associated with pain (395). Also, a significant association has been demonstrated between changes in bone marrow lesions and pain in the community and knee OA populations (396,397). In both studies, a decrease in BML score was associated with decreased knee pain. Additionally, larger MRI bone marrow lesions in knee OA are present almost exclusively in persons with knee pain (35.9% vs 2%) (100). This suggests that BMLs have an important role in knee OA pain.

The second structural target is the infrapatellar fat pad, which has a role in inflammation that may influence disease development and progression (55). Effusion-synovitis links local inflammation with knee pain (398) and can be seen on MRI (399). The Infrapatellar fat pad texture scores have been used to predict knee OA incidence (397) and have been associated with pain sensitisation in knee OA (87). These factors point towards the infrapatellar fat pad as a potential tissue capable of driving pain sensitisation, and therefore, targeting synovitis might reduce pain (87).

In this chapter, new targets are added to existing cartilage radiomics to predict baseline knee pain in the progression cohort of the OAI. Based on the findings of chapter six, only the most successful model, the shallow neural network, is used to predict knee pain. Furthermore, only the cross-sectional pain is modelled, given the lack of additional insights provided by the pain trajectory classes when considering patients who undergo total knee replacement surgery. The use of a knee atlas for generating additional segmentations is explored, including the expected normalisation effect introduced when registering MR images to an atlas.

7.2 Method

7.2.1 Study participants

The progression cohort of the OAI was selected to examine the prediction of baseline knee OA pain. Knees within the progression cohort previously undergoing total knee replacement surgery were excluded (n = 20). The progression cohort consisted of 1370 participants with 1863 knees available with suitable MRI data. Clinical and demographic information is presented in Table 7.

7.2.2 Knee Atlas

A study-specific knee atlas was created using 10 participants from the incident cohort of the OAI with no knee pain, difficulty squatting or pivoting, a knee valgus alignment of 180°, a BMI <30 and age between 50 and 70 years. The right knees underwent an initial reflection along the median plane to have the same structure as the left knees. MRI images underwent N4 Bias field correction to remove low-frequency intensity non-uniformity in the MRI data (351). Also, the right knee underwent an initial reflection along the median plane to have the same structure as the left knee.

Rather than co-registering images, cartilage segmentations generated using the u-net described in chapter four were registered. All cartilage labels were initially set to 1, with background regions having a value of 0. Knees were iteratively co-registered to an average image using affine registration, and the average image was updated after each iteration. A total of 10 iterations were used to create the knee atlas.

The affine registration step was performed in Matlab using an intensity-based registration algorithm. The optimiser was a regular step gradient descent with a mean squares metric. The properties set for the optimiser were gradient magnitude tolerance = 1e-4, minimum step length = 1e-5, maximum step length = 6.25e-2, maximum iterations = 100, and relaxation factor 1e-1. The number of non-zero pixels was recorded to assess the registration progress at each iteration. The final MRI knee atlas was generated by applying the affine matrix transformations to the MRI images of each participant.

7.2.3 Radiomic features

In the knee Atlas, the femur and tibia sub-cartilage regions were segmented manually in ITK-snap. The infrapatellar fat pad was segmented using the same process. Registration of MRIs to the knee atlas may not always be completely accurate. Therefore, to reduce the effect of other tissues on the radiomics features from the atlas-generated segmentations, an erosion operation was applied using a sphere of 10 pixels in diameter.

For the progression cohort of the OAI (n = 1863), MRI images underwent N4 Bias field correction to remove low-frequency intensity non-uniformity in the MRI data (351). The right knees underwent an initial reflection along the median plane to have the same structure as the left knees. Cartilage regions of interest (ROI) were not used for the registration of the progression cohort, as damaged cartilage was unlikely to provide a reliable reference for the registration. Instead, Knee MRIs were registered to the knee atlas using an initial affine transform (FSL-FLIRT) followed by a non-linear transform (FSL-FNIRT) (400–402).

For each participant, the affine and non-linear transforms generated when registering Knee MRIs to the atlas were applied to the cartilage segmentations, which were then combined with the atlas-based segmentations. Nine labels were generated: the patellar, femoral, lateral tibial, medial tibial, lateral meniscal and medial meniscal cartilage, the infrapatellar fat pad and femoral and tibial sub cartilage regions. The cartilage segmentation masks were priorities in cases where there was an overlap between segmentations. In such cases, creating multiple volumes for the non-cartilage segmentation was possible. Post-processing

was applied to remove all but the largest region of interest for each segmentation to create more anatomically plausible segmentations.

The cartilage volume pre- and post-registration were compared to examine the effect on the mean and standard deviation of the cartilage volumes when registering knee MRIs to an atlas. Additionally, the two radiomics metrics from chapter six (Volume and 3D Run Length Volume) were compared to test for differences in their association with the WOMAC pain score. This was done using the Person correlation coefficient, and Fisher z-to-r transform for p-value.

The relationship between radiomic features and WOMAC pain scores was examined by calculating the Pearson correlation coefficient for cartilage features with and without the registration to the knee atlas. The relationship between WOMAC pain scores and the additional radiomic features obtained from the knee atlas was also examined.

7.2.4 Pain prediction model

In chapter six, the shallow neural network was the most accurate model for classifying baseline knee pain using the radiomics features as predictors. The same neural network was used to test the improved prediction using the additional radiomic features from the knee atlas.

The network is designed with a single hidden layer with ten activation nodes. A hyperbolic tangent sigmoid activation function was used to improve training convergence. The network was trained using a cross-entropy loss function and gradient descent with momentum optimization. A softmax activation function was used in the last layer to perform classification. The network was trained via k-fold cross-validation, with k equal to ten within each fold. The split between training, validation, and testing was 70%, 20%, and 10%, respectively.

7.3 Results

7.3.1 Knee Atlas

The training progress of the knee atlas generation is shown in Figure 46; as a global measure of progress, the number of non-zero pixels is shown at each iteration. The first iteration is the starting point, where segmentations of all the participants are used to create an average image. As the segmentations converge, the total number of non-zero pixels decreases. In this plot, the reduction in non-zero pixels is also plotted for each iteration defined as $N_{I+1} - N_I$, where N is the number of non-zero pixels and I is the iteration.

Figure 47 shows the alignment of the segmentations. As the iterations progress, the cartilage segmentations can be seen to align. A heatmap is used to demonstrate areas that are well aligned versus those that do not align as well. A higher value indicates a more consistent alignment across the 10 participants used to create the atlas. Figure 48 shows the resulting knee atlas after the affine transformations are applied to the MRIs.

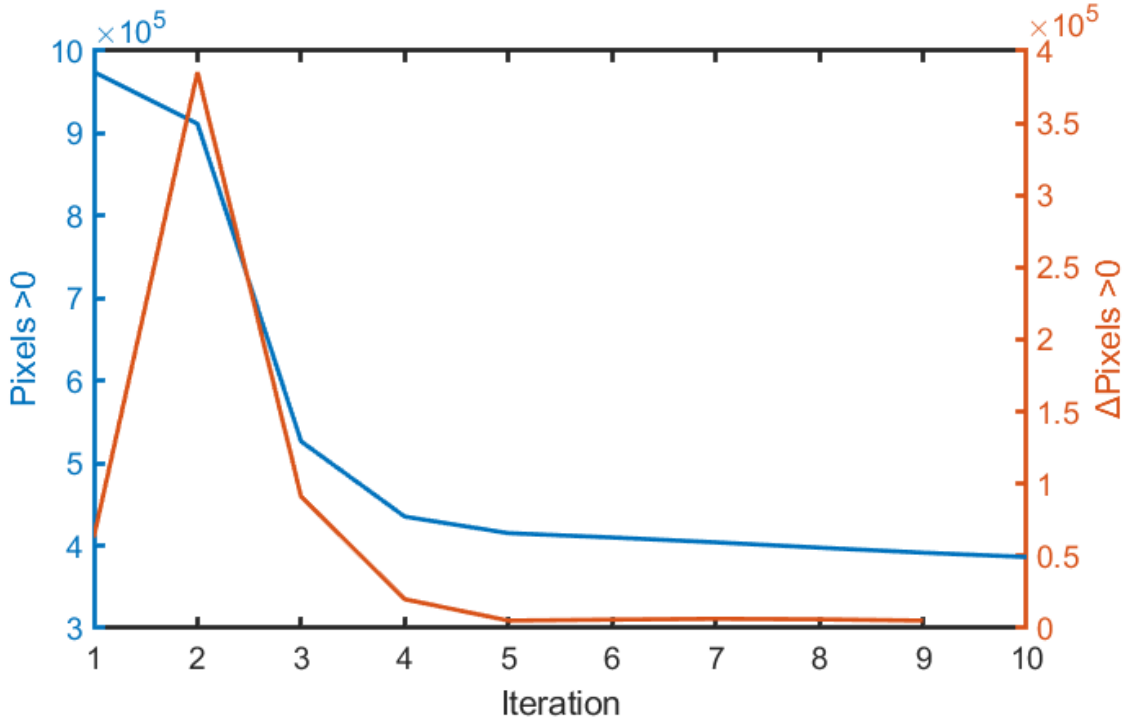


Figure 46: Training progress of knee atlas using cartilage segmentations
The number of non-zero pixels is plotted in blue (Left Axis), and the reduction in non-zero pixels is shown in orange (Right Axis). Initially, the different cartilage segmentations cover a larger area, but as the segmentations align to the atlas, the number of non-zero pixels decreases. After several iterations, the change in the number of non-zero pixels is close to zero, and training is stopped as there is no further convergence of the segmentation masks.

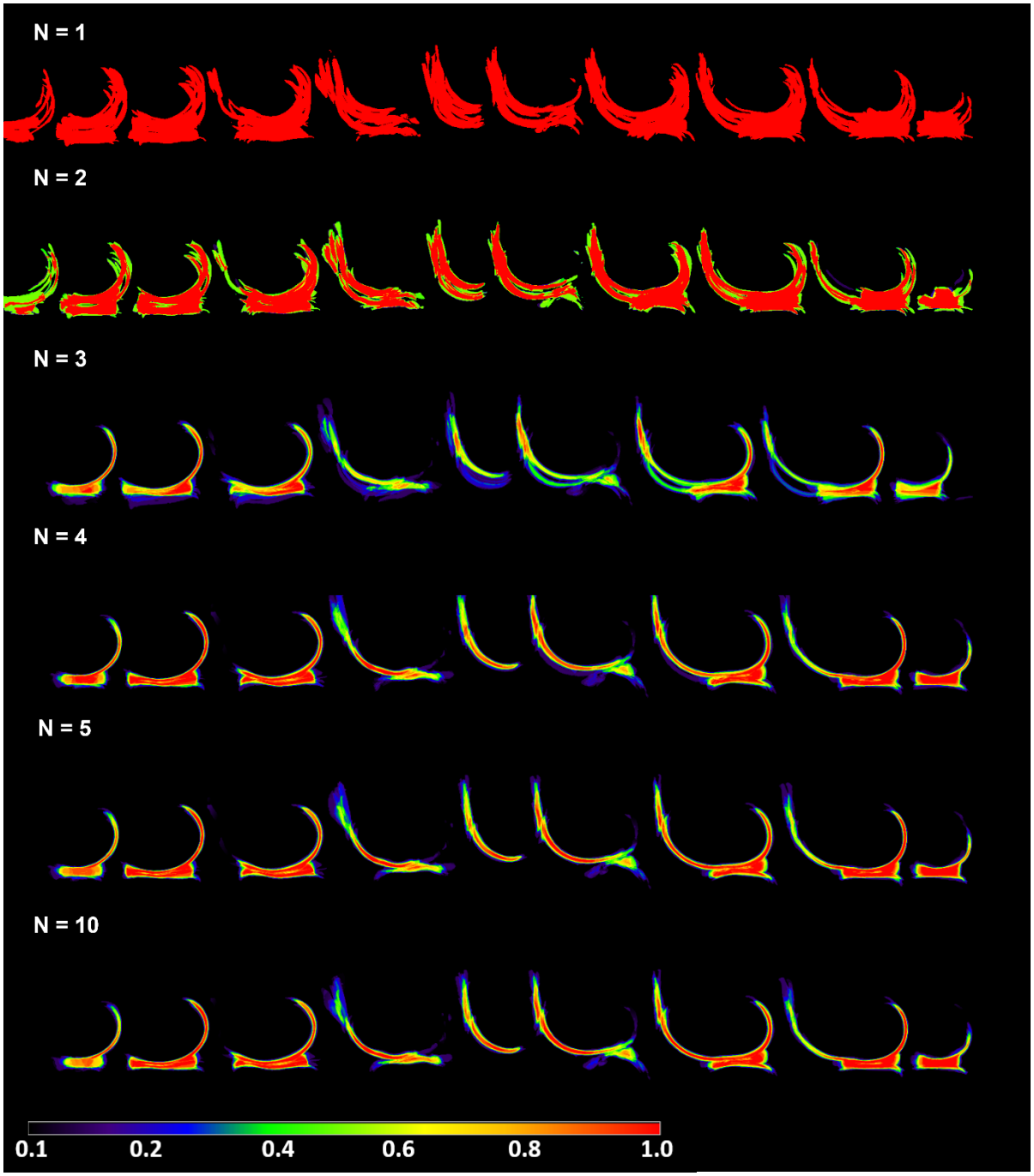


Figure 47: Sagittal view of knee atlas cartilage segmentation
Initially, the different cartilage segmentations covered a larger area, but as the segmentations aligned, the average image resembled healthy knee cartilage. The Patella has the lowest average pixel intensity, suggesting it is the least well-aligned target using affine registration.

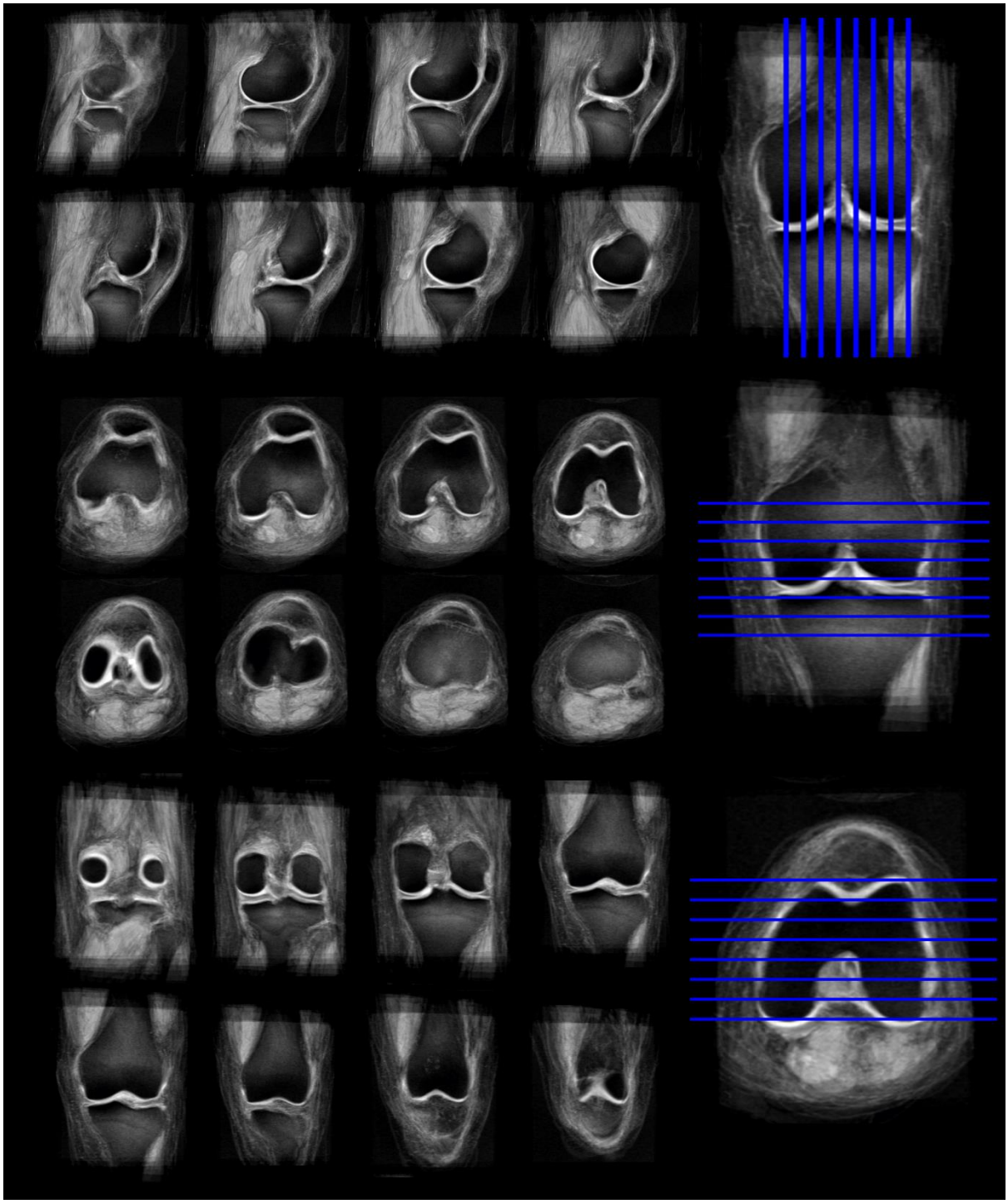


Figure 48: MRI Knee atlas
A knee atlas was generated using 10 participants from the OAI incident cohort. Cartilage segmentation drove the affine registrations, and the images were registered using the final set of affine transforms. Tibiofemoral cartilage appears well-defined, while cartilage at the patellofemoral joint is subject to more blurring.

7.3.2 Radiomic features

The additional segmentations created using the knee atlas are shown in Figure 49. Morphological erosion was applied to shrink the borders and reduce the likelihood of radiomic features being generated from non-target regions after the registration of the image to the atlas. In 95 cases (5.6% of the total number), the registration to the atlas image failed to converge, resulting in these images being rejected from the remainder of the analysis. Four hundred eighty-seven atlas-based radiomic features from the Image Biomarker Standardization Imitative (260) were extracted for each label, resulting in 4383 radiomic features per knee.

Results from comparing cartilage volume pre and post-registration to the knee atlas are shown for femoral cartilage in Figure 50. The mean and standard deviation for all cartilage volumes, pre and post-registration, are given in Table 11. The average difference in cartilage volume for a given cartilage label was -4.0% (range -7.7% to 1.2%) in the knee atlas registered images. The average standard deviation in cartilage volume for a given cartilage label was -30.2% (Range -49.1% to -16.9%) in the knee atlas registered images.

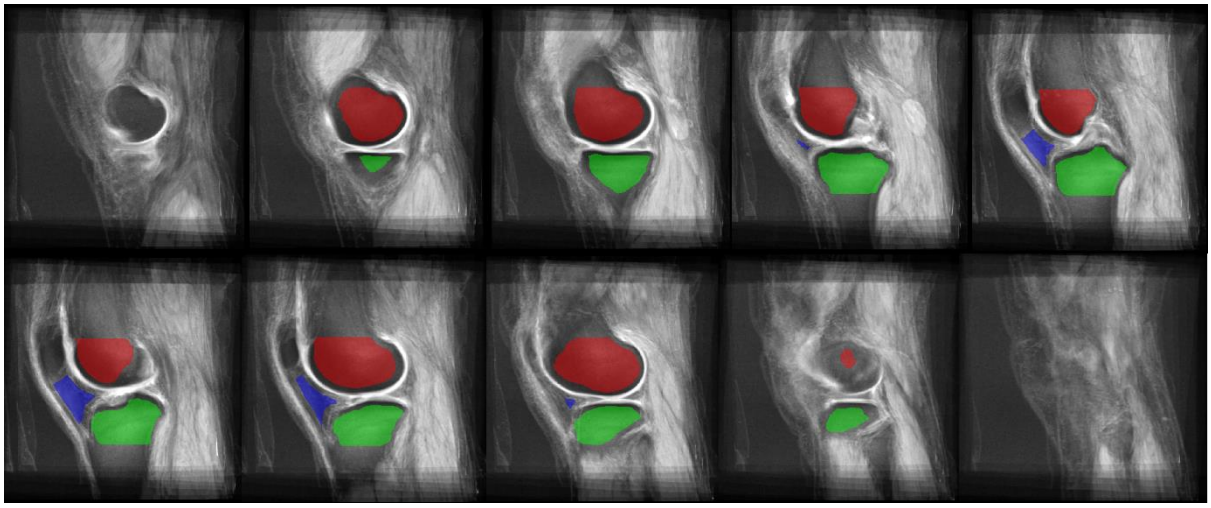


Figure 49: Atlas-based segmentation
Using the knee atlas, three new regions were segmented: the femoral and tibial sub-cartilage regions and the infrapatellar fat pad. The femoral and tibial regions were chosen as likely regions for BML formation, and the patella was selected as a potential source of local inflammation in the knee joint.

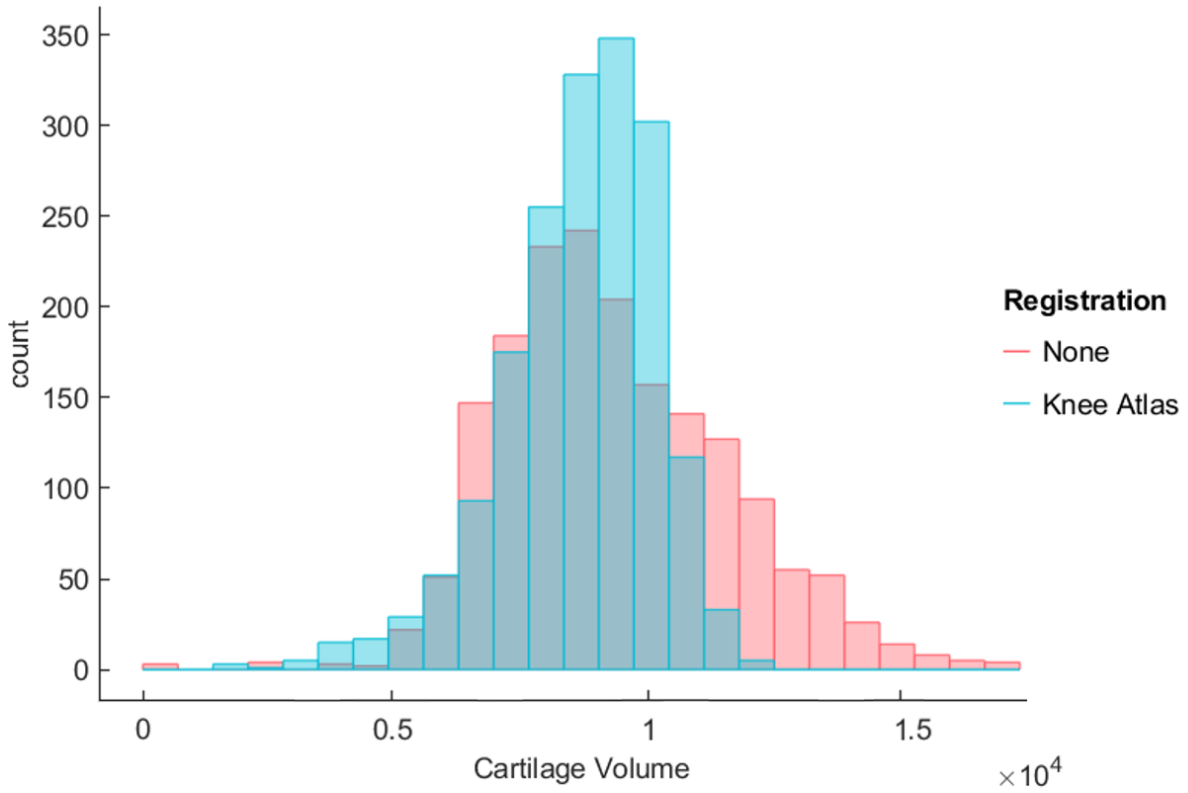


Figure 50: Effect of atlas-based registration on femoral cartilage volume
Femoral cartilage volume was lower in the knee atlas registered knees than in the original images. Volumes were calculated using an ensemble of 2D U-nets in the original knee MRIs. While the mean volume was 8% lower in the atlas registered knees, the standard deviation was reduced by 49%, indicating a more closely grouped set of segmentations after registration.

Registration	Femoral Cartilage		Lat. Meniscus		Lat. Tibial Cartilage		Medial Meniscus		Medial Tibial Cartilage		Patellar Cartilage	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
None	9275	2246	1600	534	1607	551	1616	530	1487	449	1810	734
Knee Atlas	8608	1506	1518	436	1590	432	1644	453	1506	305	1607	623
Difference (%)	-7.7	-49.1	-5.4	-22.4	-1.0	-27.5	1.7	-16.9	1.2	-47.5	-12.6	-17.8

Table 11: Cartilage volume in native and atlas-registered knees
Differences in cartilage volume were measured across all subjects using native imaging and knee atlas-registered images. The standard deviation (std) was lower in all cases, and the average volume was lower in all cases except for the medial meniscus in the knee atlas-registered images. This shows that the atlas registration progress slightly decreases knee volumes while tightening the range of volumes in the segmented cartilage structures.

		Femoral Cartilage	Lat. Meniscus	Lat. Tibial Cartilage	Medial Meniscus	Medial Tibial Cartilage	Patellar Cartilage
Volume	Original	**-0.06**	-0.05	-0.13	-0.06	**-0.12**	-0.05
	Knee Atlas	**-0.20**	-0.09	-0.18	-0.11	**-0.21**	-0.10
	p-value	**<0.01**	0.22	0.12	0.14	**<0.01**	0.18
3D Run Length Variance	Original	-0.13	-0.09	-0.15	**-0.08**	**-0.14**	**-0.04**
	Knee Atlas	-0.07	-0.07	-0.11	**0.00**	**-0.03**	**-0.12**
	p-value	0.08	0.68	0.27	**0.03**	**<0.01**	**0.02**

Table 12: Association between cartilage volume and WOMAC knee pain
Pearson's correlation coefficient for cartilage volumes and 3D run length variance for native and knee atlas registered images. Fisher's z-to-r transform was used to compare the correlation coefficients. Cartilage labels with $p<0.05$ are shown in bold. After registration to the knee atlas, several cartilage volumes were significantly more correlated with knee pain. The interpretation of 3D run length variance, another cartilage feature picked for its strong association with knee pain, shows registration to a knee atlas having a varied effect.

The effect of atlas registration on the association between cartilage volume and WOMAC knee pain is shown in Table 12. Figure 51 shows the relationship between radiomic features and WOMAC pain scores for images pre and post-registration. By comparing the cartilage radiomic features, there are more features with higher mean R-values across cartilage types when knee MRIs were registered to the Knee Atlas. Figure 52 compares the correlation between new radiomic targets and cartilage targets. The infrapatellar fat pad appeared to generate features that were, on average, better correlated with WOMAC pain than cartilage features. The opposite seemed to be true for the subchondral bone.

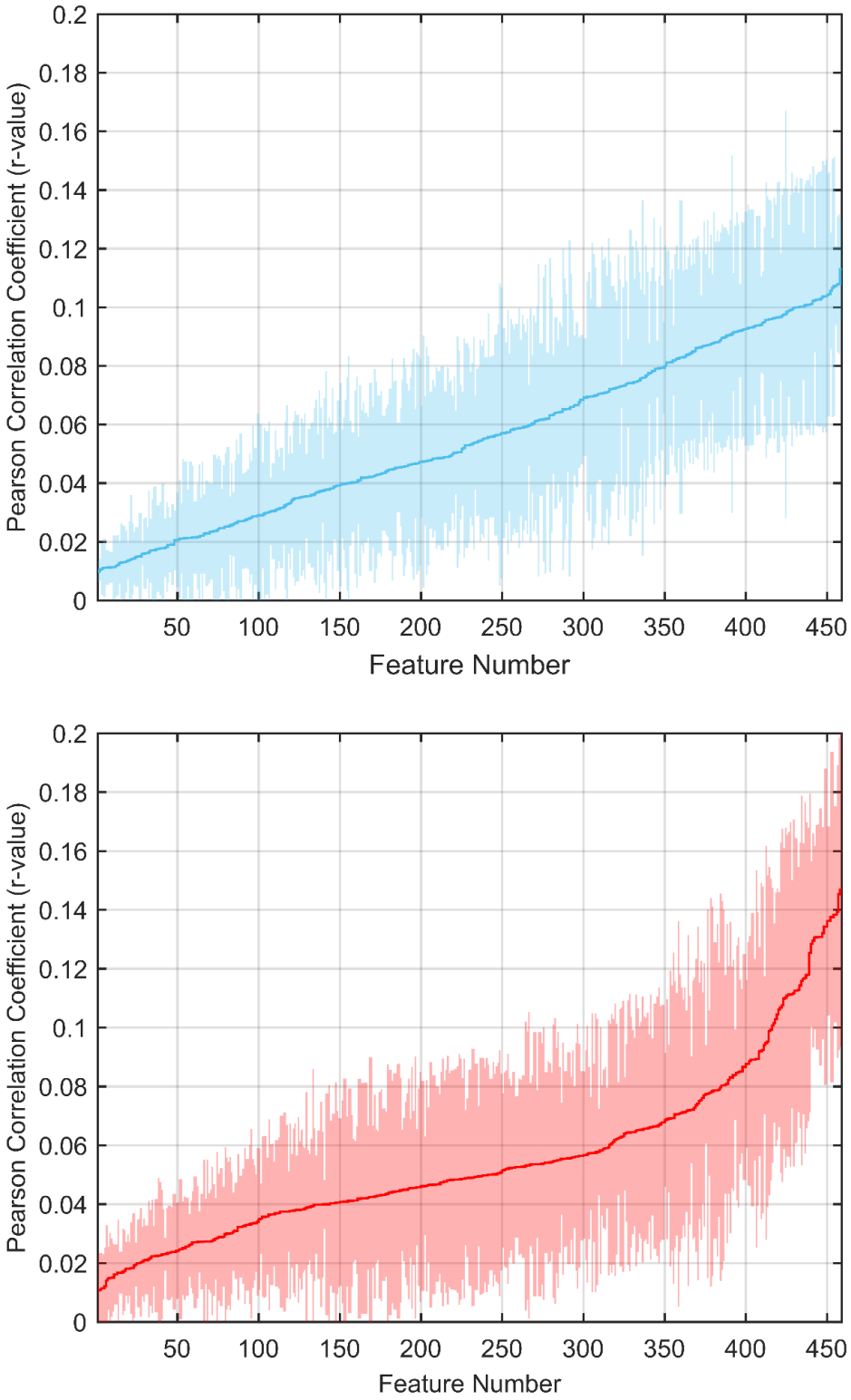


Figure 51: Effect of knee atlas registration on radiomic features

The absolute value of the Pearson correlation coefficient was calculated between cartilage radiomic features and WOMAC pain for patients in the progression cohort of the OAI. In each plot, features were sorted in ascending order. Top: mean and standard deviation R-values for original cartilage segmentations. Bottom: mean and standard deviation R-values for knee-atlas registered cartilage segmentations. The atlas registered features correlate more strongly to knee pain than the original features.

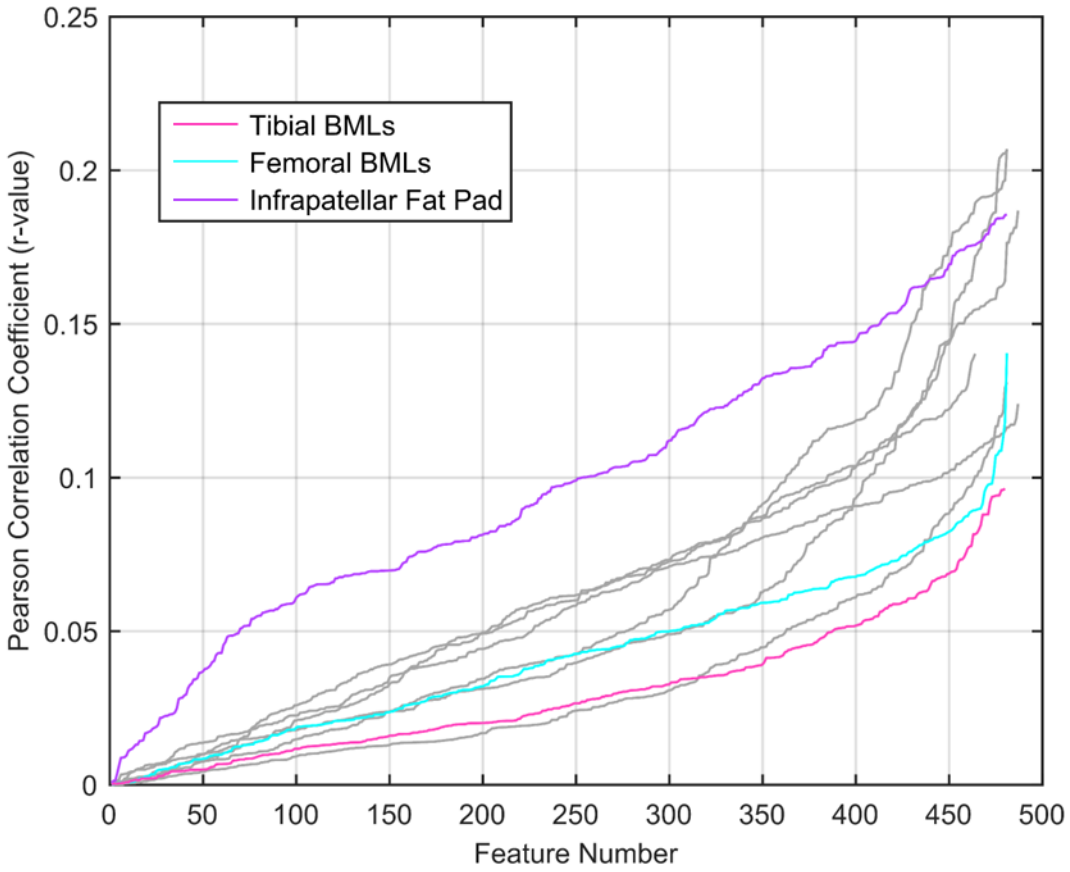


Figure 52: Non-cartilage radiomic targets correlation with pain
The absolute value of the Pearson correlation coefficient was calculated between radiomic features and WOMAC pain score for patients in the progression cohort of the OAI. Correlation coefficients were calculated after the registration of images to the knee Atlas. For each instance, features were sorted in ascending order, with individual cartilage radiomic features shown in grey. While the BML targets in the tibia and femur performed poorly, the infrapatellar fat pad showed a much stronger association with knee pain than the cartilage or bone regions.

7.3.2 Pain Prediction Model

Using the same shallow neural network architecture as in chapter six, baseline WOMAC pain was predicted using radiomics features from cartilage, femoral and tibial sub-cartilage regions and the infrapatellar fat pad. The results are shown in Figure 53. A comparison to the original cartilage radiomics model is made in Table 13; the mean AUC across the pain groups was 0.626 in the cartilage-only model and increased to 0.645 in the extended radiomics model.

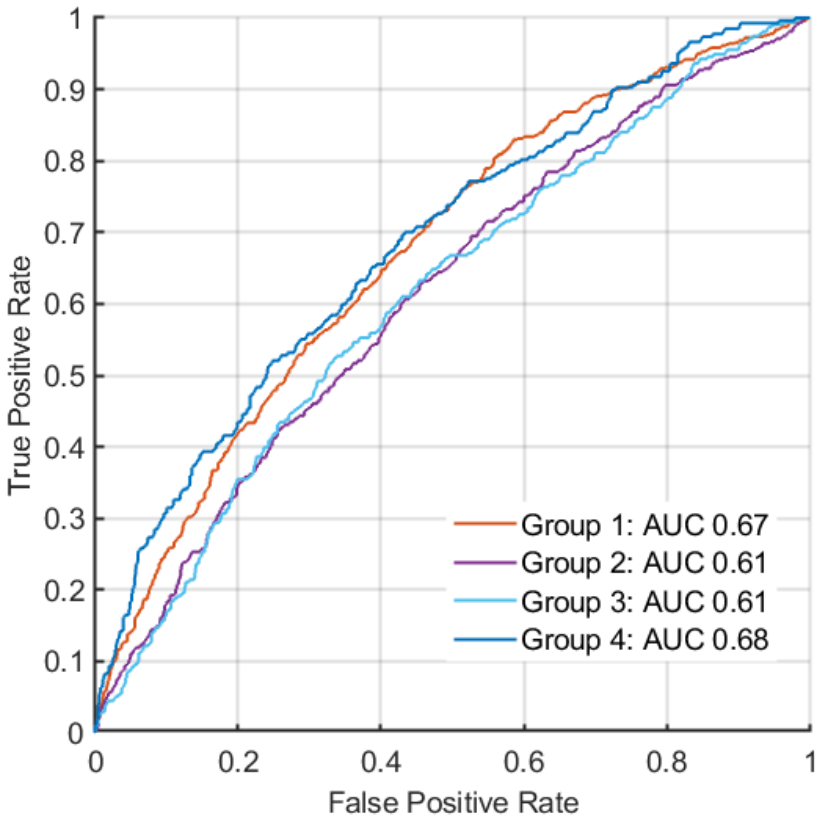


Figure 53: Predicting baseline pain scores using neural networks
Receiver operator characteristic curve for predicting baseline WOMAC knee pain using radiomic features from cartilage plus additional pain targets. AUC scores are one-versus-all for each group. Pain Scores are: Group 1: Mild (1-3), Group 2: Moderate (4-6), Group 3: Severe (7-9) and Group 4: Extreme (≥10). These results show only a marginal improvement in the predictions made without the additional structural targets (AUC 0.59 – 0.67).

Model	**Pain**			
	Mild (1-3)	Moderate (4-6)	Severe (7-9)	Extreme 10)
Cartilage Only	0.65	0.59	0.59	0.67
Knee Atlas-based	0.67	0.61	0.61	0.68

Table 13: Area Under the Curve values for radiomic knee pain models
The native cartilage model is compared to the knee atlas-based model incorporating additional radiomic targets, including the sub-cartilage regions in the femur and tibia and the infrapatellar fat pad. These results show only a marginal improvement in the predictions made without the additional structural targets.

7.4 Discussion

Extending the cartilage-pain model to include other targets led to several interesting observations. Firstly, developing a knee atlas was possible without performing registration on the MRIs. Instead, an atlas was generated by registering segmented cartilage. These aligned quickly, providing a robust knee atlas within only ten iterations. The effect of registering images to the atlas strongly affects cartilage volume measurements. While there was a slight offset in volume measurement (-4.0%), the standard deviation of the cartilage volume was much lower, reduced on average by -30.2%. This was beneficial to the association between cartilage volume and baseline WOMAC pain, leading to an average 90% increase in Pearson's correlation coefficient.

The effect of image registration on other radiomics features was explored. For the 3D run length variance example, two labels were significantly worse (medial meniscus and medial tibial cartilage), while one was considerably better (patellar cartilage). It is possible that features that depend on the variation in grayscale in the image were affected by the distortion introduced by the non-linear registration.

Segmentation to the knee atlas had a relatively high failure rate (5.6%). This could be a feature of the algorithms used (FSL-FLIRT and FSL-FNIRT), which have settings already optimized for brain registration. Differences in intensity modulation, bias field, and inter-subject variability may all have had an impact. When checking registered images manually to evaluate the suitability of this registration, the images appear well registered. However, with a registration failure rate of 1 in 18, further optimisation may have improved the quality of the final registration.

The level of regularisation in the knee atlas registered images was not explored quantitatively. Stronger regularisations (reduced local warping) may significantly affect the stability of radiomics features in the registered images. Removing the non-linear registration in favour of an affine-only registration may be preferable.

The addition of new labels, including the sub-cartilage regions in the tibia and femur to capture BML activity and the infrapatellar fat pad to include inflammatory changes, did not have as much of an effect as desired on the final predictive value of the prediction model. An average 0.02 increase in AUC was observed when using the extended Radiomics model compared to the cartilage-only model (mean AUC of 0.63).

Registration improves the average correlation of cartilage radiomic features with WOMAC pain. This highlights a potential improvement in the quality of radiomic features post-registration. Including new targets in the knee atlas space had a mixed effect. The radiomic features from the infrapatellar fat pad appeared to be more strongly correlated with knee pain than the cartilage features, suggesting it may be a better target for pain. Conversely, the subchondral bone proved to be a poor-quality target. This could be due to technical parameters, such as registration or reduction of the segmented volumes (erosion) to reduce possible overlaps with other structures. Deep learning-based segmentation methods may mitigate or reduce some of these effects but were not explored due to the time cost of generating manual segmentation.

Additional targets that may be difficult to segment using an atlas-based registration model may improve results. For example, the inclusion of osteophytes, effusion-synovitis, and ligaments. These additional structures could be added by training a U-net for automated segmentation, as done for knee cartilage in chapter four. Another benefit of deep learning-generated segmentations is their superiority over atlas-based registration, providing fewer errors in the segmented volumes. While U-net segmentations could extend the radiomics framework to additional targets, manual segmentation is time-consuming and could still miss significant targets that drive knee pain.

A deep learning convolutional neural network trained on the whole image may improve pain prediction, given the large number of potential regions contributing to peripheral pain. The advantage of such a model would be removing a priori knowledge, such as using handcrafted features (radiomics) or selecting regions of interest (segmentation).

The connections made within a neural network can be sensitised to patterns within the image that are not captured by the radiomics approach; for example, strong activation may be caused by a cluster of related changes that are spatially disparate, signalling an underlying pathology that would not easily have been identified via a radiomics approach. Convolutional neural networks have the added benefit of pinpointing the locations of the signals driving the decision-making via saliency maps created from the network's activation functions (369).

A challenge in predicting pain scores is its subjective nature, with many factors affecting a person's reporting of their experience. Psychological, environmental, and constitutional factors can affect osteoarthritis pain perception, with some authors suggesting OA pain intensity is not correlated with joint degeneration. Still, it should instead be viewed as a set of phenotypes: constant, intermittent or neuropathic, with or without central centralisation (403). Additionally, the high prevalence of OA features (19 – 43%, > 40 years old) in asymptomatic uninjured knees (404) is suggestive that not all OA pathology is a peripheral driver of pain.

Adding clinico-demographic data will likely improve the model's ability to predict knee pain. In a study using a deep learning model to predict pain progression based on knee radiographs (389), the inclusion of age, gender, race, BMI, WOMAC pain score, and KL grade increased the AUC score from 0.77 to 0.81 ($p<0.05$). An extension of the knee-atlas-based pain prediction model to include clinico-demographic confounders may improve the ability of the underlying imaging features to predict knee OA pain by including factors that correlate with the perception or intensity of an individual's knee pain.

7.5 Conclusion

Adding new segmentation to capture BMLs and inflammation in the infrapatellar fat pad only had a limited effect on the prediction of knee pain (AUC increase of 0.02) compared to a cartilage-only model. Future improvements in the model could include changing the regularisation parameters of the registration to reduce localised distortion or using fully U-net-derived segmentations for greater accuracy. Alternate approaches, such as end-to-end convolutional neural networks for pain prediction, may have a significant effect.

Registration of images to a knee atlas significantly improved the association between knee volume and pain score. Additionally, the registration improved the overall correlation of cartilage radiomic features with knee pain.

Previous studies reported good performance (AUC of 0.77 – 0.81) for predicting knee pain progression (389–392) in classifying 'healthy' or unaffected knees from those that experienced pain progression. The premise of the radiomics comparison presented in the chapter was identifying pain in a cohort selected from the 'affected' group. This was because a model developed in a patient cohort could have a more immediate clinical application. However, the model performance was insufficient to evaluate it in a clinical context.

The limited ability to classify WOMAC pain in this model may be representative of the exclusion of significant clinical and demographic covariates or the limited contribution of peripheral pain drives compared to central and neuropathic components of the disease. In the discussion, alternate models, including deep learning for segmentation or direct end-to-end learning, were discussed and may provide more compelling models for peripheral drivers of knee pain.

8 Conclusion

8.1 Contribution to knowledge

This thesis has investigated the role of magnetic resonance imaging biomarkers in predicting meaningful clinical scores and outcomes in knee osteoarthritis. The imaging markers developed could be used in the clinic to inform patients and their physicians or in research as mechanistic outcome markers. Significant contributions to the knowledge of this thesis include a new method for the development of anatomically plausible segmentations, an analysis pipeline for the prediction of clinical outcomes using high-throughput radiomic features, and assessment of these features in both their native space and registered to an anatomical template.

8.2 Cartilage segmentation

Chapter four, the first experimental chapter of the thesis, starts with developing the U-net model used throughout the remainder of the thesis. The development of the U-net segmentation model for robust knee cartilage segmentation in this thesis preceded the 2021 publication of the IWOAI (International Workshop on Osteoarthritis Imaging) Knee MRI Segmentation Challenge. Results were comparable to the best outcomes from other competing institutions (Table 14). Additionally, the segmentation pipeline for this thesis was designed to create cartilage regions with a single component per label, a critical step for radiomic analysis used in subsequent chapters.

1. Author	Femoral Cartilage	Tibial Cartilage *(Medial Lateral)*	Patellar Cartilage	Meniscal Cartilage *(Medial Lateral)*
IWOAI Challenge (279)	0.91(0.05)	0.90 (0.06) †	0.88(0.09)	0.87 (0.05) †
Peake et al. (this study)	0.88(0.03)	0.84(0.08) 0.88(0.04)	0.85(0.10)	0.85(0.05) 0.90(0.04)

Note. †Lateral and medial components were combined.

Table 14: Dice coefficients for cartilage segmentation
Results for automatic cartilage segmentation methods using the iMorphics dataset. Data are means with standard deviations in parentheses. For the IWOAI segmentation challenge, the top submission for each cartilage component was used as a reference. The segmentations achieved in this thesis have a similar DICE to the top outputs from an international challenge later released using the same dataset.

The U-net segmentations developed were used to reproduce the 2003 findings of Hunter et Al. comparing cartilage volume with symptoms in knee osteoarthritis. This showed a weak association of cartilage volume with knee pain. Rather than using manual segmentations, which are time-consuming, the fully automated cartilage segmentation allowed for a scaling-up of the number of cases from 133 to 9383. The time needed to segment each volume was reduced from 20 minutes to 14 seconds, primarily by removing human oversite from the algorithm execution.

The results from the original study were reproduced, with a small but significant correlation between cartilage volumes and WOMAC Scores, excluding femoral cartilage. However, the

finding that patellar cartilage was more strongly correlated with symptoms than other cartilage within the knee was not upheld.

This chapter introduced the concept of using radiomic features by comparing the intensity range with volumetric measurements of cartilage, showing that intensity may be a more sensitive marker for knee pain. This draws on the findings that show changes in quantitative T2 precede volumetric changes. These changes will likely affect DESS (Double Echo Steady State) MRI signal intensity, as contrast is determined by the T2/T1 ratio of the tissue.

In addition to having higher correlation scores with knee pain in half the cartilage segments, a case study demonstrates the advantage of using intensity descriptors for knee cartilage. In the case study, cartilage lesions demonstrate local intensity changes on DESS MRI that precede volumetric changes. It is proposed that having a range of handcrafted descriptors may improve the utility of cartilage segmentation for diagnosis or prediction.

The chapter firmly establishes the advantages of fully automated deep learning segmentations over manual methods regarding throughput and segmentation accuracy. Due to the parallelisation of computer learning tasks, using a high-performance computing cluster, the cartilage segmentation of over nine thousand knees was achievable in under 5 minutes, a task that would have taken over 3000 person-hours to complete using semi-automated tools.

An additional benefit of a deep learning-driven method was the accurate contouring of complex boundaries. Given the large volume of data (160 slices with a 384x384 matrix), this task is not readily achievable using conventional methods. One drawback highlighted in this chapter was the challenge of using overlap measures to validate the performance of segmentation models when accurate segmentation of the ground truth is limited.

8.3 Prediction of Total Knee Replacement

Chapter five examines the U-net segmentation method and the utility of high-throughput radiomics. This chapter aimed to predict total knee replacement within five years of a baseline MRI. While the intensity range provides improved sensitivity to WOMAC scores, it remained weakly correlated. Multiple distinct descriptors of the underlying cartilage MRI signal within these regions were obtained by scaling up the number of different image features derived from the cartilage segmentations. Radiomic analysis was used for high throughput image analysis, relying on many samples to determine significant and independent imaging features for data modelling.

Radiomics is used to develop a mechanical surrogate outcome marker to predict the likelihood of individuals undergoing total knee replacement within five years of their MRI. This application was intended to demonstrate the clinical utility of automated cartilage segmentation. The decision to have a total knee replacement usually comes down to a discussion between a patient and their clinician.

The determination is usually straightforward; joint replacement is recommended when patients have exhausted previous treatment options and pain or loss of joint function has reduced their living quality. However, many features that have successfully predicted total knee replacement, such as education and income, may introduce implicit bias where patients' treatment is affected by their ability to pay or self-advocate. The value of an MRI-based model increases as it excludes many unhelpful biases that limit access to healthcare.

In the future, the models developed in this thesis could be adapted to answer other important questions, such as predicting the success of total knee replacement. Such a tool could help prioritise long waiting lists and enhance clinical decision-making.

The model used in chapter five was developed explicitly for a setting where the number of predictors was much greater than the number of observations. The backbone of the model was a regularised Elastic Net regression, employing a grid search for optimal hyperparameters. The model was initially developed and validated in participants with symptomatic knee OA using internal cross-validation that incorporated the entire pipeline. The method was designed to be widely generalisable for applying radiomics data to other clinical problems.

In addition to providing a standard reporting of results via Area Under the Curve metrics and Receiver-Operator Characteristic curves, a unique contribution to the knowledge of this chapter stemmed from the use of Cox proportional hazard modelling. This enabled accurate reporting of the likelihood of individuals to undergo total knee replacement surgery; in this study, the risk groups were generated using a median split of the data to create 'high' and 'low' risk categories. Unlike other forms of analysis, Cox proportional hazard permits the use of right-censored data, preventing drop-out bias from affecting reported results.

Results showed that 'high-risk individuals were much more likely to undergo total knee replacement surgery compared to those at lower risk. To validate these models, the radiomics model translated best from the progression cohort, where it was developed, into the incident cohort, where it was validated. The thresholds for developing high & low-risk groups were not varied, yet the proportional hazard ratios remained relativity unchanged (Table 15), demonstrating good internal validity of the model.

	Model	**C-Statistic (95% CI)**	**Proportional hazard ratio. (95% CI, p-value)**
Progression Cohort	Clinical and Demographic	0.79 (0.75 – 0.83)	5.9 (5.6 – 6.2, p< .001)
	Radiomics	0.81 (0.77 – 0.84)	7.5 (7.1 – 7.9, p< .001)
	Combined	0.85 (0.82 – 0.88)	9.8 (9.4 – 10.2, p< .001)
Incident Cohort	Clinical and Demographic	0.75 (0.69 – 0.78)	4.8 (4.3 – 5.3, p< .001)
	Radiomics	0.82 (0.75 – 0.86)	9.6 (9.2 – 10.1, p< .001)
	Combined	0.75 (0.71 – 0.80)	4.8 (4.3 – 5.2, p< .001)

Table 15: Prediction of total knee replacement
The results for predicting total knee replacement within five years, as well as the progression cohort of the Osteoarthritis, were used to develop the models. Statistics were generated using internal cross-validation. The incident cohort of the Osteoarthritis Initiative was used as a test set to check the internal validity of the models. Given that both cohorts of the OAI have MRIs from the same centres but different clinical scores due to inclusion criteria, the radiomics model showed the most robust performance in the incident cohort.

Future developments could focus on why predictive models fail. It may be that many of the false positives arising from the radiomics model do require a knee replacement. Still, changes in circumstances such as new employment limiting time off work, ability to pay out-

of-pocket expenses or long waiting lists prevent the individual from getting surgery. Similarly, with false negatives, the opposite may be true; an opening on a long waiting list or improved financial circumstances may prompt an individual to get a knee replacement before they would usually be considered for one.

As with all data modelling, external validation of the finding would be needed before use in a clinical setting. The radiomics model may also be an imaging-based mechanistic outcome marker for clinical trials. Within randomised control trials, these could be secondary outcome markers used to test the performance of a new drug or intervention in knee osteoarthritis. As they are independent of clinico-demographic factors, they should provide a robust response to changes in cartilage appearance over time.

The long acquisition time for the cartilage-sensitive MRI sequence is a current limitation. Using existing clinical deep-learning image reconstruction techniques, it is likely possible to shorten the acquisition time by up to 50%. At the same time, further accelerations may be obtained by using dedicated sequence-specific image reconstruction techniques. For routine clinical use, with current advanced acceleration techniques, a 10 or 15-minute exam time would be sufficient to acquire a fast clinical knee protocol with the additional high resolution DESS sequence. The segmentations could be passed to the clinician alongside the risk score for TKR and individual scores of the most significant radiomic features to help aid the interpretation of the underlying algorithm.

Acquiring the DESS at a lower resolution could also be feasible to enable a shorter exam time. While the impact on the radiomics models has not been assessed in this thesis, some segmentation and prediction algorithm retraining would likely be required. However, at an additional 5 minutes for an accelerated DESS MRI, the greater challenge would be demonstrating and evaluating the clinical utility of this additional information rather than further shortening exam time.

8.4 Pain phenotyping using cartilage MRI

In chapter six, the association between knee pain and MRI-based cartilage radiomics was examined. A classification test was used to investigate the ability of these features to predict knee pain and pain phenotypes in the progression cohort of the OAI. Finding robust imaging markers for either pain or pain trajectory would be highly desirable for early treatment, and localization of peripheral drivers may facilitate more personalized treatments.

As cartilage radiomics provided a good prognosis for TKR, it was hypothesised that they may also predict baseline pain or longitudinal pain trajectory scores. Given the significant research investment in knee cartilage biomarkers and the limited association with knee pain in existing literature, this would be a welcome finding.

Previous studies reported good performance with Area Under Curve scores of 0.77 – 0.81 for predicting knee pain progression (360–363) in classifying ‘healthy’ or unaffected knees from those that experienced pain progression. However, more is needed to provide clinical efficacy or reliable outcome markers for clinical trials in populations exclusively suffering from osteoarthritis, especially given the high cost of MRI.

An initial examination found that cartilage radiomic features were weakly correlated with pain and pain trajectory scores in the Osteoarthritis Initiative cohort, with correlation scores of < 0.16. A follow-up analysis by Turkey Kramer's multi-comparison test showed some significant differences between groups, though they were small. Several classification methods were tested, and the most successful was neural network classification, which achieved Area Under the Curve scores of 0.63 for baseline pain prediction and 0.64 for pain trajectory prediction.

It was concluded that cartilage radiomics alone would be insufficient to predict baseline pain scores. Damage to other innervated tissues or sensitisation of innervated tissues via other joint structures may explain the limited correlation between pain and cartilage loss.

8.5 Pain phenotyping part extended methods

Chapter seven explores new methods to improve the prediction of knee pain in osteoarthritis, building on the knowledge from previous chapters. Two main changes were made to try and improve prediction accuracy. Methodologically, an atlas-based registration method was developed to reduce the variability of radiomic features due to knee size. This was intended as a form of 'size normalisation' for the radiomic features. Secondly, additional targets from areas shown to contribute to knee pain in OA were added to the radiomics model.

To simplify the analysis and enable a closer focus on the new methods, neural network classification was used exclusively in this chapter, as it was the most successful method in Chapter five. Additionally, only baseline pain was taken forward as a valid outcome measure. Pain trajectories were mostly stable over time, leading to no additional clinical information compared to baseline pain scores. Group 4 improved pain scores over time; however, this was driven mainly by including patients who received TKR over the 9-year follow-up. Data from the Supplementary information (http://links.lww.com/PAIN/B88) showed individuals in group 4 were 2.5 and 4.5 times more likely to receive TKR than the other groups.

Creating a high-resolution MRI knee Atlas from the incident cohort of the OAI showed plausible results with well-defined cartilage boundaries and smooth bone texture. Registration of images to the Atlas reduced the standard deviation of cartilage volumes by up to 49%, showing a volume normalisation effect. This size normalisation increased the correlation coefficient of cartilage volume with baseline knee pain by 90%.

Overall, features extracted from images registered to the knee atlas had higher correlation coefficients. Features from the infrapatellar fat pad were more strongly correlated compared to other targets, while the features from the trabecular bone were less strongly correlated compared to other targets.

Despite some improvements in the correlation of features with knee pain, including other knee targets and registration to the knee atlas, the area under the curve metric was only improved to 0.65, which is still below what would be needed for beneficial clinical or research activities.

While the classification model did not prove successful, this chapter had several key findings. Significantly, the knee atlas did improve the correlation of features with the outcome, and the infrapatellar fat pad was shown to be a more significant target for knee pain than cartilage or trabecular bone. Further developments were discussed for the atlas method. This could involve improving the registration and segmentation processes. Other deep learning methods, such as directly from the MRI or smaller segmented sub-volumes, may be more successful.

Pain is a complex experience, and a person's expertise influences perception. Pain can be intermittent, neuropathic, and with or without central centralisation. This is likely to limit the ability of peripheral drivers alone to predict an individual's pain scores. Including other clinico-demographic factors in a pain model may help improve pain prediction. If imaging only contributes fractionally to the overall outcome, alternative strategies may be needed to find suitable biomarkers for knee pain in osteoarthritis. If designing a new study, it may be anticipated that knee joint radiomics would correlate better with nociception than pain experience. Thus, the use of pain thresholds at the knee may be a better proxy for nociception.

8.6 Future developments

The radiomics model in chapter four may be the most likely to generalise across study populations, as it is not dependent on the clinicodemographic characteristics of a cohort. The internal validation of this data shows imaging-based predictors may be more stable across populations, enabling more accessible translation of this model compared to clinical and demographic predictors, which did not perform as well when evaluated. Further evaluation of this model could include external assessment of the radiomics model at a different centre, using scanners from different MRI manufacturers. This would test the robustness of the segmentation and feature extraction pipeline and the ability to reproduce images with similar contrast using MRI from a different vendor.

Data from an external evaluation may be sufficient to include the radiomics model as a secondary outcome measure for a clinical trial, with primary outcome markers likely to be changes in clinical scores (such as WOMAC). This would be a step towards improving the robustness of studies evaluating disease-modifying osteoarthritis drugs, especially those that target cartilage.

As well as serving as a potential mechanistic outcome marker, it would also be valuable to give clinicians access to automated segmentations developed in chapter four of this thesis. These could be used to explain the underlying cartilage damage to patients with knee OA, empowering service users to take control of their health and well-being.

Studies suggest this may be effective as visualising pathology can boost compliance with treatments, even if those treatments are challenging. Initially, this could be run as a controlled trial with outcomes such as patient engagement with fitness programmes and group interviews to assess the value of having a clinician-led programme that includes visualisation of a patient's cartilage.

Initially, as part of any evaluation of either the segmentation or the TKR prediction tool, results could be added to the hospital system, such as the Picture Archiving and Communication System (PACS) provided sites are supported with necessary domain expertise. If sufficient data were collected showing the utility of cartilage segmentation, this could provide the impetus needed to develop these tools as either a third-party solution or as a product from the MRI manufacturer. Both cases require significant investment to achieve regulatory approval (CE or UKCA marking). However, an ever-growing number of artificial intelligence products are reaching this threshold and finding their way into internationally regulated markets. Work presented in this thesis may be used to power future studies to examine the clinical efficacy of a TKR prediction tool as a step towards clinical implementation.

9 Acknowledgements

This research was funded by the NIHR Nottingham Biomedical Research Centre. The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Data and/or research tools used in the preparation of this manuscript were obtained and analysed from the controlled access datasets distributed from the Osteoarthritis Initiative (OAI), a data repository housed within the NIMH Data Archive (NDA). OAI is a collaborative informatics system created by the National Institute of Mental Health and the National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS) to provide a worldwide resource to quicken the pace of biomarker identification, scientific investigation and OA drug development.

10 References

1. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis. Nat Rev Rheumatol [Internet]. 2014 [cited 2021 Dec 17];10(7):437–41. Available from: https://pubmed.ncbi.nlm.nih.gov/24662640/

2. Vos T, Allen C, Arora M, Barber RM, Brown A, Carter A, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet [Internet]. 2016 Oct 8 [cited 2021 Dec 17];388(10053):1545–602. Available from: http://www.thelancet.com/article/S0140673616316786/fulltext

3. Age Structure - Our World in Data [Internet]. [cited 2023 Jan 9]. Available from: https://ourworldindata.org/age-structure?country=

4. Greenaway C, Hargreaves S, Barkati S, Coyle CM, Gobbi F, Veizis A, et al. COVID-19: Exposing and addressing health disparities among ethnic minorities and migrants. J Travel Med [Internet]. 2020 Nov 9 [cited 2023 Jan 9];27(7):1–3. Available from: https://academic.oup.com/jtm/article/27/7/taaa113/5875716

5. IMPROVING OUTCOMES THROUGH PERSONALISED MEDICINE Working at the cutting edge of science to improve patients' lives.

6. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 Feb 2;542(7639):115–8.

7. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019 Jun 1;25(6):954–61.

8. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. J Am Med Assoc. 2016 Dec 13;316(22):2402–10.

9. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. Nature 2020 577:7788 [Internet]. 2020 Jan 1 [cited 2023 Jan 6];577(7788):89–94. Available from: https://www.nature.com/articles/s41586-019-1799-6

10. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. Lancet Digit Health [Internet]. 2021 Mar 1 [cited 2023 Jan 9];3(3):e195–203. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

11. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. Ann Rheum Dis [Internet]. 2001 [cited 2021 Dec 20];60(2):91. Available from: /pmc/articles/PMC1753462/

12. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Vol. 393, The Lancet. Lancet Publishing Group; 2019. p. 1745–59.

13. Bruyère O, Cooper C, Pelletier J, … EMS in arthritis and, 2016 undefined. A consensus statement on the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO) algorithm for the. Elsevier [Internet]. [cited 2023 Jan 4]; Available from: https://www.sciencedirect.com/science/article/pii/S0049017215002887

14. Bruyère O, Cooper C, Pelletier JP, Branco J, Luisa Brandi M, Guillemin F, et al. An algorithm recommendation for the management of knee osteoarthritis in Europe and internationally: A report from a task force of the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). Semin Arthritis Rheum. 2014 Dec 1;44(3):253–63.

15. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: The chronic osteoarthritis management initiative of the U.S. bone and joint initiative. Semin Arthritis Rheum [Internet]. 2014 [cited 2023 Jan 5];43(6):701–12. Available from: https://pubmed.ncbi.nlm.nih.gov/24387819/

16. Bijlsma JWJJ, Berenbaum F, Lafeber FPJGJG. Osteoarthritis: An update with relevance for clinical practice. The Lancet [Internet]. 2011 Jun 18 [cited 2018 Oct 11];377(9783):2115–26. Available from: https://www.sciencedirect.com/science/article/pii/S0140673611602432

17. Centers for Disease Control and Prevention (CDC). Prevalence and most common causes of disability among adults--United States, 2005. MMWR Morb Mortal Wkly Rep [Internet]. 2009 May 1 [cited 2023 Jan 5];58(16):421–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19407734

18. Schiphof D, van Middelkoop M, de Klerk BM, Oei EHG, Hofman A, Koes BW, et al. Crepitus is a first indication of patellofemoral osteoarthritis (and not of tibiofemoral osteoarthritis). Osteoarthritis Cartilage [Internet]. 2014 May 1 [cited 2023 Jan 5];22(5):631–8. Available from: http://www.oarsijournal.com/article/S106345841400048X/fulltext

19. Chapleau J, Lambert BS, Sullivan TC, Clyburn TA, Incavo SJ. Impact of Valgus vs Varus Mechanical Axis Correction During Primary Total Knee Arthroplasty on Postoperative Periarticular Bone Mineral Density. J Arthroplasty. 2021 May 1;36(5):1792–8.

20. Prieto-Alhambra D, Judge A, Javaid MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: Influences of age, gender and osteoarthritis affecting other joints. Ann Rheum Dis. 2014 Sep;73(9):1659–64.

21. Turkiewicz A, Petersson IF, Björk J, Hawker G, Dahlberg LE, Lohmander LS, et al. Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. Osteoarthritis Cartilage. 2014 Nov 1;22(11):1826–32.

22. Pereira D, Peleteiro B, Araújo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: A systematic review [Internet]. Vol. 19, Osteoarthritis and Cartilage. 2011 [cited 2020 Mar 23]. p. 1270–85. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21907813

23. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet [Internet]. 2012 [cited 2023 Jan 5];380(9859):2163–96. Available from: https://pubmed.ncbi.nlm.nih.gov/23245607/

24. Cui A, Li H, Wang D, Zhong J, Chen Y, Lu H. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine [Internet]. 2020 Dec 1 [cited 2023 Jan 4];29–30. Available from: https://pubmed.ncbi.nlm.nih.gov/34505846/

25. Dahaghin S, Bierma-Zeinstra SMA, Ginai AZ, Pols HAP, Hazes JMW, Koes BW. Prevalence and pattern of radiographic hand osteoarthritis and association with pain and disability (the Rotterdam study). Ann Rheum Dis [Internet]. 2005 May [cited 2021 Dec 17];64(5):682. Available from: /pmc/articles/PMC1755481/

26. Yu D, Peat G, Bedson J, Jordan KP. Annual consultation incidence of osteoarthritis estimated from population-based health care data in England. Rheumatology (Oxford) [Internet]. 2015 Aug 6 [cited 2021 Dec 17];54(11):2051. Available from: /pmc/articles/PMC4603278/

27. Leal J, Gray AM, Prieto-Alhambra D, Arden NK, Cooper C, Javaid MK, et al. Impact of hip fracture on hospital care costs: a population-based study. Osteoporosis International [Internet]. 2016 Feb 1 [cited 2021 Dec 17];27(2):549. Available from: /pmc/articles/PMC4740562/

28. Chen A, Gupte C, Akhtar K, Smith P, Cobb J. How the UK Compares. Article ID. 2012;2012.

29. McKee M, Dunnell K, Anderson M, Brayne C, Charlesworth A, Johnston-Webber C, et al. The changing health needs of the UK population. The Lancet. 2021 May 22;397(10288):1979–91.

30. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis Rheum [Internet]. 1986 [cited 2021 Dec 30];29(8):1039–49. Available from: https://pubmed.ncbi.nlm.nih.gov/3741515/

31. Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, et al. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. Ann Rheum Dis [Internet]. 2010 Mar 1 [cited 2018 Oct 18];69(3):483–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19762361

32. Sakellariou G, Conaghan PG, Zhang W, Bijlsma JWJ, Boyesen P, D'Agostino MA, et al. EULAR recommendations for the use of imaging in the clinical management of peripheral joint osteoarthritis. Ann Rheum Dis [Internet]. 2017 Sep 1 [cited 2021 Dec 30];76(9):1484–94. Available from: https://pubmed.ncbi.nlm.nih.gov/28389554/

33. Bijlsma JWJ, Berenbaum F, Lafeber FPJG. Osteoarthritis: an update with relevance for clinical practice. Lancet. 2011;377(9783):2115–26.

34. Michael Nevitt AC, Felson DT, Lester G. OAI Protocol THE OSTEOARTHRITIS INITIATIVE PROTOCOL FOR THE COHORT STUDY.

35. Zhang Y, Jordan JM. Epidemiology of Osteoarthritis. Clin Geriatr Med [Internet]. 2010 Aug [cited 2021 Dec 20];26(3):355. Available from: /pmc/articles/PMC2920533/

36. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. Osteoarthritis Cartilage [Internet]. 2005 Sep [cited 2021 Dec 20];13(9):769–81. Available from: https://pubmed.ncbi.nlm.nih.gov/15978850/

37. Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. Osteoarthritis Cartilage [Internet]. 2015 Apr 1 [cited 2021 Dec 20];23(4):507–15. Available from: http://www.oarsijournal.com/article/S1063458414013429/fulltext

38. Brouwer GM, van Tol AW, Bergink AP, Belo JN, Bernsen RMD, Reijman M, et al. Association between valgus and varus alignment and the development and progression of radiographic osteoarthritis of the knee. Arthritis Rheum [Internet]. 2007 Apr [cited 2021 Dec 20];56(4):1204–11. Available from: https://pubmed.ncbi.nlm.nih.gov/17393449/

39. Runhaar J, van Middelkoop M, Reijman M, Vroegindeweij D, Oei EHG, Bierma-Zeinstra SMA. Malalignment: a possible target for prevention of incident knee osteoarthritis in overweight and obese women. Rheumatology (Oxford) [Internet]. 2014 [cited 2021 Dec 20];53(9):1618–24. Available from: https://pubmed.ncbi.nlm.nih.gov/24706992/

40. Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. Osteoarthritis Cartilage. 2004 Jan 1;12(SUPLL.):39–44.

41. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Vol. 393, The Lancet. Lancet Publishing Group; 2019. p. 1745–59.

42. Hunter DJ, Eckstein F. Exercise and osteoarthritis. J Anat [Internet]. 2009 [cited 2021 Dec 20];214(2):197. Available from: /pmc/articles/PMC2667877/

43. Gelber AC, Hochberg MC, Mead LA, Wang NY, Wigley FM, Klag MJ. Joint injury in young adults and risk for subsequent knee and hip osteoarthritis. Ann Intern Med [Internet]. 2000 Sep 5 [cited 2021 Dec 20];133(5). Available from: https://pubmed.ncbi.nlm.nih.gov/10979876/

44. McMillan G, Nichols L. Osteoarthritis and meniscus disorders of the knee as occupational diseases of miners. Occup Environ Med [Internet]. 2005 Aug 1 [cited 2021 Dec 20];62(8):567–75. Available from: https://oem.bmj.com/content/62/8/567

45. Cooper C, McAlindon T, Coggon D, Egger P, Dieppe P. Occupational activity and osteoarthritis of the knee. Ann Rheum Dis [Internet]. 1994 Feb 1 [cited 2021 Dec 20];53(2):90–3. Available from: https://ard.bmj.com/content/53/2/90

46. Iannone F, Lapadula G. Obesity and inflammation--targets for OA therapy. Curr Drug Targets [Internet]. 2010 Mar 25 [cited 2021 Dec 23];11(5):586–98. Available from: https://pubmed.ncbi.nlm.nih.gov/20199391/

47. Wang H, Bai J, He B, Hu X, Liu D. Osteoarthritis and the risk of cardiovascular disease: a meta-analysis of observational studies. Sci Rep [Internet]. 2016 Dec 22 [cited 2021 Dec 23];6. Available from: /pmc/articles/PMC5177921/

48. Veronese N, Cereda E, Maggi S, Luchini C, Solmi M, Smith T, et al. Osteoarthritis and mortality: A prospective cohort study and systematic review with meta-analysis. Semin Arthritis Rheum [Internet]. 2016 Oct 1 [cited 2021 Dec 23];46(2):160–7. Available from: https://pubmed.ncbi.nlm.nih.gov/27179749/

49. Louati K, Vidal C, Berenbaum F, Sellam J. Association between diabetes mellitus and osteoarthritis: Systematic literature review and meta-analysis. RMD Open [Internet]. 2015 Jan 1 [cited 2021 Dec 23];1(1):e000077. Available from: https://rmdopen.bmj.com/content/1/1/e000077

50. Marshall DA, Liu X, Barnabe C, Yee K, Faris PD, Barber C, et al. Existing comorbidities in people with osteoarthritis: a retrospective analysis of a population-based cohort in Alberta, Canada. BMJ Open [Internet]. 2019 Nov 1 [cited 2021 Dec 23];9(11):e033334. Available from: https://bmjopen.bmj.com/content/9/11/e033334

51. Calders P, van Ginckel A. Presence of comorbidities and prognosis of clinical symptoms in knee and/or hip osteoarthritis: A systematic review and meta-analysis. Semin Arthritis Rheum [Internet]. 2018 Jun 1 [cited 2021 Dec 23];47(6):805–13. Available from: https://pubmed.ncbi.nlm.nih.gov/29157670/

52. Swain S, Sarmanova A, Coupland C, Doherty M, Zhang W. Comorbidities in Osteoarthritis: A Systematic Review and Meta-Analysis of Observational Studies. Arthritis Care Res (Hoboken) [Internet]. 2020 Jul 1 [cited 2023 Jan 4];72(7):991–1000. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/acr.24008

53. Brandt KD, Radin EL, Dieppe PA, van de Putte L. Yet more evidence that osteoarthritis is not a cartilage disease. Ann Rheum Dis [Internet]. 2006 Oct [cited 2022 Jan 14];65(10):1261. Available from: /pmc/articles/PMC1798332/

54. Pelletier JP, Martel-Pelletier J, Abramson SB. Osteoarthritis, an inflammatory disease: Potential implication for the selection of new therapeutic targets. Vol. 44, Arthritis and Rheumatism. 2001. p. 1237–47.

55. Ioan-Facsinay A, Kloppenburg M. An emerging player in knee osteoarthritis: the infrapatellar fat pad [Internet]. Vol. 15, Arthritis Research and Therapy. 2013 Dec [cited 2019 Jan 23]. Available from: http://arthritis-research.com/content/15/6/225

56. Berenbaum F. Osteoarthritis as an inflammatory disease (osteoarthritis is not osteoarthrosis!). Osteoarthritis Cartilage. 2013 Jan 1;21(1):16–21.

57. Koonce RC, Bravman JT. Obesity and osteoarthritis: More than just wear and tear. Journal of the American Academy of Orthopaedic Surgeons [Internet]. 2013 Mar [cited 2022 Jan 14];21(3):161–9. Available from: https://journals.lww.com/jaaos/Fulltext/2013/03020/Obesity_and_Osteoarthritis__More_Than_Just_Wear.5.aspx

58. Fu K, Robbins SR, McDougall JJ. Osteoarthritis: the genesis of pain. Rheumatology [Internet]. 2018 May 1 [cited 2022 Jan 14];57(suppl_4):iv43–50. Available from: https://academic.oup.com/rheumatology/article/57/suppl_4/iv43/4753702

59. Loeser RF, Collins JA, Diekman BO. Ageing and the pathogenesis of osteoarthritis. Nat Rev Rheumatol [Internet]. 2016 Jul 1 [cited 2022 Jan 14];12(7):412. Available from: /pmc/articles/PMC4938009/

60. Mow VC, Holmes MH, Michael Lai W. Fluid transport and mechanical properties of articular cartilage: a review. J Biomech [Internet]. 1984 [cited 2022 Jan 14];17(5):377–94. Available from: https://pubmed.ncbi.nlm.nih.gov/6376512/

61. Sasisekharan R, Raman R, Prabhakar V. Glycomics approach to structure-function relationships of glycosaminoglycans. Annu Rev Biomed Eng [Internet]. 2006 [cited 2022 Jan 14];8:181–231. Available from: https://pubmed.ncbi.nlm.nih.gov/16834555/

62. Plaas AHK, West LA, Midura RJ. Keratan sulfate disaccharide composition determined by FACE analysis of keratanase II and endo-beta-galactosidase digestion products. Glycobiology [Internet]. 2001 [cited 2022 Jan 14];11(10):779–90. Available from: https://pubmed.ncbi.nlm.nih.gov/11588154/

63. Goldring M, Goldring MB. The role of the chondrocyte in osteoarthritis. Arthritis Rheum. 2000;43(9):1916–26.

64. Akkiraju H, Nohe A. Role of chondrocytes in cartilage formation, progression of osteoarthritis and cartilage regeneration [Internet]. Vol. 3, Journal of Developmental Biology. NIH Public Access; 2015 [cited 2022 Jan 14]. p. 177–92. Available from: /pmc/articles/PMC4916494/

65. Goldring MB. Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. Ther Adv Musculoskelet Dis [Internet]. 2012 [cited 2022 Jan 14];4(4):269–85. Available from: https://pubmed.ncbi.nlm.nih.gov/22859926/

66. Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. Vol. 393, The Lancet. Lancet Publishing Group; 2019. p. 1745–59.

67. Davies-Tuck ML, Wluka AE, Forbes A, Wang Y, English DR, Giles GG, et al. Development of bone marrow lesions is associated with adverse effects on knee cartilage while resolution is associated with improvement - a potential target for prevention of knee osteoarthritis: A longitudinal study. Arthritis Res Ther. 2010 Jan 19;12(1).

68. van der Kraan PM, van den Berg WB. Osteophytes: relevance and biology. Osteoarthritis Cartilage. 2007 Mar 1;15(3):237–44.

69. Deveza LA, Loeser RF. Is osteoarthritis one disease or a collection of many? Rheumatology [Internet]. 2018 May 1 [cited 2022 Jan 15];57(suppl_4):iv34–42. Available from: https://academic.oup.com/rheumatology/article/57/suppl_4/iv34/4753701

70. Glyn-Jones S, Palmer AJR, Agricola R, Price AJ, Vincent TL, Weinans H, et al. Osteoarthritis. The Lancet. 2015 Jul 25;386(9991):376–87.

71. Liu A, Kendzerska T, Stanaitis I, Hawker G. The relationship between knee pain characteristics and symptom state acceptability in people with knee osteoarthritis. Osteoarthritis Cartilage [Internet]. 2014 Feb [cited 2021 Dec 27];22(2):178–83. Available from: https://pubmed.ncbi.nlm.nih.gov/24300776/

72. Parry EL, Thomas MJ, Peat G. Defining acute flares in knee osteoarthritis: a systematic review. BMJ Open [Internet]. 2018 Jul 1 [cited 2021 Dec 27];8(7):e019804. Available from: https://bmjopen.bmj.com/content/8/7/e019804

73. Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S, et al. The Revised IASP definition of pain: concepts, challenges, and compromises. Pain [Internet]. 2020 Sep 9 [cited 2023 Jan 4];161(9):1976. Available from: /pmc/articles/PMC7680716/

74. Millan MJ. The induction of pain: an integrative review. Prog Neurobiol. 1999 Jan 1;57(1):1–164.

75. Englund M, Roemer FW, Hayashi D, Crema MD, Guermazi A. Meniscus pathology, osteoarthritis and the treatment controversy. Nature Reviews Rheumatology 2012 8:7 [Internet]. 2012 May 22 [cited 2021 Dec 26];8(7):412–9. Available from: https://www.nature.com/articles/nrrheum.2012.69

76. Ashraf S, Wibberley H, Mapp PI, Hill R, Wilson D, Walsh DA. Increased vascular penetration and nerve growth in the meniscus: a potential source of pain in osteoarthritis. Ann Rheum Dis [Internet]. 2011 Mar 1 [cited 2021 Dec 26];70(3):523–9. Available from: https://ard.bmj.com/content/70/3/523

77. Schaible HG, Ebersberger A, Segond Von Banchet G. Mechanisms of Pain in Arthritis. Ann N Y Acad Sci [Internet]. 2002 Jun 1 [cited 2021 Dec 26];966(1):343–54. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1749-6632.2002.tb04234.x

78. Eitner A, Hofmann GO, Schaible HG. Mechanisms of osteoarthritic pain. Studies in humans and experimental models. Front Mol Neurosci. 2017 Nov 3;10:349.

79. Mease PJ, Hanna S, Frakes EP, Altman RD. Pain Mechanisms in Osteoarthritis: Understanding the Role of Central Pain and Current Approaches to Its Treatment. J Rheumatol [Internet]. 2011 Aug 1 [cited 2021 Dec 26];38(8):1546–51. Available from: https://www.jrheum.org/content/38/8/1546

80. Latremoliere A, Woolf CJ. Central Sensitization: A Generator of Pain Hypersensitivity by Central Neural Plasticity. The journal of pain : official journal of the American Pain Society [Internet]. 2009 Sep [cited 2021 Dec 26];10(9):895. Available from: /pmc/articles/PMC2750819/

81. Treede RD, Jensen TS, Campbell JN, Cruccu G, Dostrovsky JO, Griffin JW, et al. Neuropathic pain. Neurology [Internet]. 2008 Apr 29 [cited 2021 Dec 26];70(18):1630–5. Available from: https://n.neurology.org/content/70/18/1630

82. Dimitroulas T, Duarte R v., Behura A, Kitas GD, Raphael JH. Neuropathic pain in osteoarthritis: A review of pathophysiological mechanisms and implications for treatment. Semin Arthritis Rheum. 2014 Oct 1;44(2):145–54.

83. Fingleton C, Smart K, Moloney N, Fullen BM, Doody C. Pain sensitization in people with knee osteoarthritis: a systematic review and meta-analysis. Osteoarthritis Cartilage [Internet]. 2015 Jul 1 [cited 2021 Dec 27];23(7):1043–56. Available from: https://pubmed.ncbi.nlm.nih.gov/25749012/

84. French HP, Smart KM, Doyle F. Prevalence of neuropathic pain in knee or hip osteoarthritis: A systematic review and meta-analysis. Semin Arthritis Rheum [Internet]. 2017 Aug 1 [cited 2021 Dec 27];47(1):1–8. Available from: https://pubmed.ncbi.nlm.nih.gov/28320529/

85. Hochman JR, Davis AM, Elkayam J, Gagliese L, Hawker GA. Neuropathic pain symptoms on the modified painDETECT correlate with signs of central sensitization in knee osteoarthritis. Osteoarthritis Cartilage [Internet]. 2013 Sep [cited 2021 Dec 27];21(9):1236–42. Available from: https://pubmed.ncbi.nlm.nih.gov/23973136/

86. Neogi T, Frey-Law L, Scholz J, Niu J, Arendt-Nielsen L, Woolf C, et al. Sensitivity and sensitisation in relation to pain severity in knee osteoarthritis: trait or state? Ann Rheum Dis [Internet]. 2015 Apr 1 [cited 2021 Dec 27];74(4):682–8. Available from: https://pubmed.ncbi.nlm.nih.gov/24351516/

87. Neogi T, Guermazi A, Roemer F, Nevitt MC, Scholz J, Arendt-Nielsen L, et al. Association of Joint Inflammation With Pain Sensitization in Knee Osteoarthritis: The Multicenter Osteoarthritis Study. Arthritis Rheumatol [Internet]. 2016 Mar 1 [cited 2021 Dec 27];68(3):654–61. Available from: https://pubmed.ncbi.nlm.nih.gov/26554395/

88. Tonge DP, Pearson MJ, Jones SW. The hallmarks of osteoarthritis and the potential to develop personalised disease-modifying pharmacological therapeutics. Osteoarthritis Cartilage [Internet]. 2014 [cited 2022 Jan 11];22(5):609–21. Available from: https://pubmed.ncbi.nlm.nih.gov/24632293/

89. Grässel S, Muschter D. Recent advances in the treatment of osteoarthritis. F1000Res [Internet]. 2020 [cited 2022 Jan 11];9. Available from: /pmc/articles/PMC7199286/

90. Karsdal MA, Michaelis M, Ladel C, Siebuhr AS, Bihlet AR, Andersen JR, et al. Disease-modifying treatments for osteoarthritis (DMOADs) of the knee and hip: lessons learned from failures and opportunities for the future. Osteoarthritis Cartilage [Internet]. 2016 Dec 1 [cited 2022 Jan 11];24(12):2013–21. Available from: https://pubmed.ncbi.nlm.nih.gov/27492463/

91. Guehring H, Moreau F, Daelken B, Ladel C, Guenther O, Bihlet AR, et al. The effects of sprifermin on symptoms and structure in a subgroup at risk of progression in the FORWARD knee osteoarthritis trial. Semin Arthritis Rheum [Internet]. 2021 Apr 1 [cited 2022 Jan 11];51(2):450–6. Available from: https://pubmed.ncbi.nlm.nih.gov/33752164/

92. Hochberg MC, Guermazi A, Guehring H, Aydemir A, Wax S, Fleuranceau-Morel P, et al. Effect of Intra-Articular Sprifermin vs Placebo on Femorotibial Joint Cartilage Thickness in Patients With Osteoarthritis: The FORWARD Randomized Clinical Trial. JAMA [Internet]. 2019 Oct 8 [cited 2022 Jan 11];322(14):1360–70. Available from: https://pubmed.ncbi.nlm.nih.gov/31593273/

93. Double-Blind, Randomized, Single Dose Escalation Safety Study of Intraarticular Bone Morphogenic Protein (38A BMP-7) in Subjects With Osteoarthritis (OA) of the Knee - Full Text View - ClinicalTrials.gov [Internet]. [cited 2022 Jan 11]. Available from: https://clinicaltrials.gov/ct2/show/NCT01133613

94. Dose Finding Study of Bone Morphogenetic Protein 7 (BMP-7) in Subjects With Osteoarthritis (OA) of the Knee - Full Text View - ClinicalTrials.gov [Internet]. [cited 2022 Jan 11]. Available from: https://clinicaltrials.gov/ct2/show/NCT01111045

95. Friedlaender GE, Perry CR, Cole JD, Cook SD, Cierny G, Muschler GF, et al. Osteogenic Protein-1 (Bone Morphogenetic Protein-7) in the Treatment of Tibial Nonunions: A Prospective, Randomized Clinical Trial Comparing rhOP-1 with Fresh Bone Autograft*. J Bone Joint Surg Am [Internet]. 2001 [cited 2022 Jan 11];83-A Suppl 1(Pt 2):S151. Available from: /pmc/articles/PMC1425155/

96. Malekipour F, Whitton C, Oetomo D, Lee PVS. Shock absorbing ability of articular cartilage and subchondral bone under impact compression. J Mech Behav Biomed Mater [Internet]. 2013 Oct [cited 2022 Jan 11];26:127–35. Available from: https://pubmed.ncbi.nlm.nih.gov/23746699/

97. Li G, Yin J, Gao J, Cheng TS, Pavlos NJ, Zhang C, et al. Subchondral bone in osteoarthritis: Insight into risk factors and microstructural changes. Arthritis Res Ther. 2013 Dec 9;15(6).

98. Buckland-Wright C. Subchondral bone changes in hand and knee osteoarthritis detected by radiography. Osteoarthritis Cartilage [Internet]. 2004 [cited 2022 Jan 11];12 Suppl A(SUPLL.):10–9. Available from: https://pubmed.ncbi.nlm.nih.gov/14698636/

99. Burr DB, Gallant MA. Bone remodelling in osteoarthritis. Nat Rev Rheumatol [Internet]. 2012 Nov [cited 2022 Jan 11];8(11):665–73. Available from: https://pubmed.ncbi.nlm.nih.gov/22868925/

100. Felson DT, Chaisson CE, Hill CL, Totterman SMS, Gale ME, Skinner KM, et al. The association of bone marrow lesions with pain in knee osteoarthritis. Ann Intern Med [Internet]. 2001 Apr 3 [cited 2021 Dec 23];134(7):541–9. Available from: www.annals.org

101. Spector TD. Bisphosphonates: Potential therapeutic agents for disease modification in osteoarthritis. Aging Clin Exp Res [Internet]. 2003 Oct 10 [cited 2022 Jan 13];15(5):413–8. Available from: https://link.springer.com/article/10.1007/BF03327362

102. RANKL-blockade for the Treatment of Erosive Osteoarthritis (OA) of Interphalangeal Finger Joints - Full Text View - ClinicalTrials.gov [Internet]. [cited 2022 Jan 13]. Available from: https://clinicaltrials.gov/ct2/show/NCT02771860

103. Teriparatide as a Chondroregenerative Therapy in OA - Full Text View - ClinicalTrials.gov [Internet]. [cited 2022 Jan 13]. Available from: https://clinicaltrials.gov/ct2/show/NCT03072147

104. Knadler MP, Lobo E, Chappell J, Bergstrom R. Duloxetine: Clinical pharmacokinetics and drug interactions. Clin Pharmacokinet [Internet]. 2011 Sep 13 [cited 2021 Dec 18];50(5):281–94. Available from: https://link.springer.com/article/10.2165/11539240-000000000-00000

105. Cepeda MS, Camargo F, Zea C, Valencia L. Tramadol for osteoarthritis. Cochrane Database of Systematic Reviews. 2006 Jul 19;2006(3).

106. Ashraf S, Bouhana KS, Pheneger J, Andrews SW, Walsh DA. Selective inhibition of tropomyosin-receptor-kinase A (TrkA) reduces pain and joint damage in two rat models of inflammatory arthritis. Arthritis Res Ther [Internet]. 2016 May 4 [cited 2022 Jan 13];18(1). Available from: https://pubmed.ncbi.nlm.nih.gov/27145816/

107. Iwakura N, Ohtori S, Orita S, Yamashita M, Takahashi K, Kuniyoshi K. Role of low-affinity nerve growth factor receptor inhibitory antibody in reducing pain behavior and calcitonin gene-related Peptide expression in a rat model of wrist joint inflammatory pain. J Hand Surg Am [Internet]. 2010 Feb [cited 2022 Jan 13];35(2):267–73. Available from: https://pubmed.ncbi.nlm.nih.gov/20060234/

108. First-in-human Study With Ascending Single Intra-articular Doses of GZ389988 in Patients With Painful Osteoarthritis of the Knee - Full Text View - ClinicalTrials.gov [Internet]. [cited 2022 Jan 13]. Available from: https://clinicaltrials.gov/ct2/show/NCT02424942

109. Krupka E, Jiang GL, Jan C. Efficacy and safety of intra-articular injection of tropomyosin receptor kinase A inhibitor in painful knee osteoarthritis: a randomized, double-blind and placebo-controlled study. Osteoarthritis Cartilage [Internet]. 2019 Nov 1 [cited 2022 Jan 13];27(11):1599–607. Available from: https://clinicaltrials.gov/ct2/show/NCT02845271

110. Safety, Tolerability & Pharmacokinetics of LEVI-04 in Healthy Volunteers and Patients With Osteoarthritis Knee Pain - Full Text View - ClinicalTrials.gov [Internet]. [cited 2022 Jan 13]. Available from: https://clinicaltrials.gov/ct2/show/NCT03227796

111. H. Le Graverand MP, Mazzuca S, Lassere M, Guermazi A, Pickering E, Brandt K, et al. Assessment of the radioanatomic positioning of the osteoarthritic knee in serial radiographs:

comparison of three acquisition techniques. Osteoarthritis Cartilage. 2006 Jan 1;14(SUPPL. 1):37–43.

112. Braun HJ, Gold GE. Diagnosis of osteoarthritis: Imaging. Bone. 2012 Aug 1;51(2):278–88.

113. KELLGREN JH, LAWRENCE JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis. 1957;16(4):494–502.

114. Kohn MD, Sassoon AA, Fernando ND. Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. Clin Orthop Relat Res [Internet]. 2016 Aug 1 [cited 2021 Dec 23];474(8):1886. Available from: /pmc/articles/PMC4925407/

115. Sekiya I, Katano H, Guermazi A, Miura Y, Okanouchi N, Tomita M, et al. Association of AI-determined Kellgren–Lawrence grade with medial meniscus extrusion and cartilage thickness by AI-based 3D MRI analysis in early knee osteoarthritis. Scientific Reports 2023 13:1 [Internet]. 2023 Nov 16 [cited 2024 Jan 10];13(1):1–14. Available from: https://www.nature.com/articles/s41598-023-46953-9

116. Peterfy CG, Guermazi A, Zaim S, Tirman PFJ, Miaux Y, White D, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. Osteoarthritis Cartilage. 2004 Mar;12(3):177–90.

117. Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston-Leeds Osteoarthritis Knee Score). Ann Rheum Dis. 2008;67(2):206–11.

118. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage. 2011 Aug 1;19(8):990–1002.

119. Sophia Fox AJ, Bedi A, Rodeo SA. The Basic Science of Articular Cartilage: Structure, Composition, and Function. Sports Health [Internet]. 2009 Nov [cited 2021 Dec 23];1(6):461. Available from: /pmc/articles/PMC3445147/

120. Link TM, Stahl R, Woertler K. Cartilage imaging: Motivation, techniques, current and future significance. Eur Radiol. 2007 May 9;17(5):1135–46.

121. Shakoor D, Guermazi A, Kijowski R, Fritz J, Jalali-Farahani S, Mohajer B, et al. Diagnostic performance of three-dimensional MRI for depicting cartilage defects in the knee: A meta-analysis. Radiology. 2018 Oct 1;289(1):71–82.

122. Hardy PA, Recht MP, Piraino D, Thomasson D. Optimization of a Dual Echo in the Steady State (DESS) free-precession sequence for imaging cartilage. Journal of Magnetic Resonance Imaging. 1996 Mar 1;6(2):329–35.

123. Roemer FW, Kwoh CK, Hannon MJ, Crema MD, Moore CE, Jakicic JM, et al. Semiquantitative assessment of focal cartilage damage at 3 T MRI: A comparative study of dual echo at steady state (DESS) and intermediate-weighted (IW) fat suppressed fast spin echo sequences. Eur J Radiol [Internet]. 2011 Nov 1 [cited 2021 Dec 23];80(2):e126–31. Available from: http://www.ejradiology.com/article/S0720048X10003645/fulltext

124. Chen CA, Kijowski R, Shapiro LM, Tuite MJ, Davis KW, Klaers JL, et al. Cartilage Morphology at 3.0T: Assessment of Three-Dimensional MR Imaging Techniques. J Magn Reson Imaging [Internet]. 2010 [cited 2021 Dec 23];32(1):173. Available from: /pmc/articles/PMC3065186/

125. Kijowski R, Davis KW, Woods MA, Lindstrom MJ, de Smet AA, Gold GE, et al. Knee joint: Comprehensive assessment with 3D isotropic resolution fast spin-echo MR imaging - Diagnostic performance compared with that of conventional MR imaging at 3.0 T. Radiology [Internet]. 2009 Aug 1 [cited 2021 Dec 23];252(2):486–95. Available from: https://pubs.rsna.org/doi/abs/10.1148/radiol.2523090028

126. Hunter DJ, Zhang YQ, Niu JB, Tu X, Amin S, Clancy M, et al. The association of meniscal pathologic changes with cartilage loss in symptomatic knee osteoarthritis. Arthritis Rheum. 2006 Mar;54(3):795–801.

127. Felson DT, McLaughlin S, Goggins J, LaValley MP, Gale ME, Totterman S, et al. Bone marrow edema and its relation to progression of knee osteoarthritis. Ann Intern Med [Internet]. 2003 Sep 2 [cited 2021 Dec 23];139(5 Pt 1). Available from: https://pubmed.ncbi.nlm.nih.gov/12965941/

128. Felson DT, Gale DR, Elon Gale M, Niu J, Hunter DJ, Goggins J, et al. Osteophytes and progression of knee osteoarthritis. Rheumatology [Internet]. 2005 Jan 1 [cited 2021 Dec 23];44(1):100–4. Available from: https://academic.oup.com/rheumatology/article/44/1/100/2899172

129. Cicuttini F, Baker J, Hart D, cartilage TSO and, 1996 undefined. Association of pain with radiological changes in different compartments and views of the knee joint. Elsevier [Internet]. [cited 2021 Dec 23]; Available from: https://www.sciencedirect.com/science/article/pii/S1063458405803231

130. Lanyon P, O'Reilly S, … AJA of the rheumatic, 1998 undefined. Radiographic assessment of symptomatic knee osteoarthritis in the community: definitions and normal joint space. ard.bmj.com [Internet]. [cited 2021 Dec 23]; Available from: https://ard.bmj.com/content/57/10/595.short

131. van der Heijden RA, de Kanter JLM, Bierma-Zeinstra SMA, Verhaar JAN, van Veldhoven PLJ, Krestin GP, et al. Structural abnormalities on magnetic resonance imaging in patients with patellofemoral pain: A cross-sectional case-control study. American Journal of Sports Medicine [Internet]. 2016 Sep 1 [cited 2021 Mar 16];44(9):2339–46. Available from: https://pubmed.ncbi.nlm.nih.gov/27206691/

132. Ioan-Facsinay A, Kloppenburg M. An emerging player in knee osteoarthritis: The infrapatellar fat pad. Vol. 15, Arthritis Research and Therapy. 2013.

133. Draghi F, Ferrozzi G, Urciuoli L, Bortolotto C, Bianchi S. Hoffa's fat pad abnormalities, knee pain and magnetic resonance imaging in daily practice. Vol. 7, Insights into Imaging. Springer Verlag; 2016. p. 373–83.

134. Atukorala I, Kwoh CK, Guermazi A, Roemer FW, Boudreau RM, Hannon MJ, et al. Synovitis in knee osteoarthritis: a precursor of disease?

135. Han W, Aitken D, Zhu Z, Halliday A, Wang X, Antony B, et al. Signal intensity alteration in the infrapatellar fat pad at baseline for the prediction of knee symptoms and structure in older adults: a cohort study.

136. Ruhdorfer A, Haniel F, … TPO and, 2017 undefined. Between-group differences in infra-patellar fat pad size and signal in symptomatic and radiographic progression of knee osteoarthritis vs non-progressive. Elsevier [Internet]. [cited 2021 Mar 16]; Available from: https://www.sciencedirect.com/science/article/pii/S106345841730852X

137. Wang K, Ding C, Hannon MJ, Chen Z, Kwoh CK, Lynch J, et al. Signal intensity alteration within infrapatellar fat pad predicts knee replacement within 5 years: data from the Osteoarthritis Initiative. Osteoarthritis and Cartilage [Internet]. 2018 Oct 1 [cited 2018 Oct 9];1345–50. Available from: https://www.sciencedirect.com/science/article/pii/S1063458418312895

138. Wang K, Ding C, Hannon MJ, Chen Z, Kwoh CK, Hunter DJ. Quantitative Signal Intensity Alteration in Infrapatellar Fat Pad Predicts Incident Radiographic Osteoarthritis: The Osteoarthritis Initiative. Arthritis Care Res (Hoboken) [Internet]. 2019 Jan 1 [cited 2021 Mar 16];71(1):30–8. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/acr.23577

139. Hill CL, Hunter DJ, Niu J, Clancy M, Guermazi A, Genant H, et al. Synovitis detected on magnetic resonance imaging and its relation to pain and cartilage loss in knee osteoarthritis. ard.bmj.com [Internet]. 2007 [cited 2021 Dec 23];66:1599–603. Available from: https://ard.bmj.com/content/66/12/1599.short?casa_token=8ZpZRmvg_cQAAAAA:ZBwQEACU0kmy1zn6uFbxwX9SfmmJw2R5zwFp6RQUUqnnzPGS_LMwfyWneCaRoPenQzZliy1MO8J7

140. Loeuille D, Rat AC, Goebel JC, Champigneulle J, Blum A, Netter P, et al. Magnetic resonance imaging in osteoarthritis: which method best reflects synovial membrane inflammation?: Correlations with clinical, macroscopic and microscopic features. Osteoarthritis Cartilage. 2009 Sep 1;17(9):1186–92.

141. Roemer FW, Kassim Javaid M, Guermazi A, Thomas M, Kiran A, Keen R, et al. Anatomical distribution of synovitis in knee osteoarthritis and its association with joint effusion assessed on non-enhanced and contrast-enhanced MRI. Osteoarthritis Cartilage [Internet]. 2010 Oct 1 [cited 2021 Dec 26];18(10):1269–74. Available from: http://www.oarsijournal.com/article/S1063458410002438/fulltext

142. Guermazi A, Roemer FW, Hayashi D, Crema MD, Niu J, Zhang Y, et al. Assessment of synovitis with contrast-enhanced MRI using a whole-joint semiquantitative scoring system in people with, or at high risk of, knee osteoarthritis: the MOST study. Ann Rheum Dis [Internet]. 2011 May [cited 2021 Dec 26];70(5):805. Available from: /pmc/articles/PMC4180232/

143. Hill CL, Seo GS, Gale D, Totterman S, Gale ME, Felson DT. Cruciate ligament integrity in osteoarthritis of the knee. Arthritis Rheum [Internet]. 2005 Mar 1 [cited 2021 Dec 26];52(3):794–9. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.20943

144. Lohmander LS, Englund PM, Dahl LL, Roos EM. The long-term consequence of anterior cruciate ligament and meniscus injuries: Osteoarthritis. American Journal of Sports Medicine [Internet]. 2007 Oct 30 [cited 2021 Dec 26];35(10):1756–69. Available from: https://journals.sagepub.com/doi/full/10.1177/0363546507307396?casa_token=el4EjH6pmwkAAAAA%3AEsXZNvmFkNC79tO775dwJ1bqypBbvS4k2BRSFBzZdjkfvpD6MRgZFKBn8HtTLlwuo86MZCRxAZsBDw

145. Schaefer FKW, Schaefer PJ, Brossmann J, Frahm C, Muhle C, Hilgert RE, et al. Value of fat-suppressed PD-weighted TSE-sequences for detection of anterior and posterior cruciate ligament lesions—Comparison to arthroscopy. Eur J Radiol [Internet]. 2006 Jun 1 [cited 2021

Dec 26];58(3):411–5. Available from: http://www.ejradiology.com/article/S0720048X06000064/fulltext

146. Seedhom BB, Dowson D, Wright V. Proceedings: Functions of the menisci. A preliminary study. Ann Rheum Dis. 1974;33(1):111.

147. Englund M, Guermazi A, Lohmander SL. The role of the meniscus in knee osteoarthritis: a cause or consequence? Radiol Clin North Am [Internet]. 2009 Jul [cited 2021 Dec 26];47(4):703–12. Available from: https://pubmed.ncbi.nlm.nih.gov/19631077/

148. Jung JY, Yoon YC, Kwon JW, Ahn JH, Choe BK. Diagnosis of internal derangement of the knee at 3.0-T MR imaging: 3D isotropic intermediate-weighted versus 2D sequences. Radiology [Internet]. 2009 Dec [cited 2021 Dec 26];253(3):780–7. Available from: https://pubmed.ncbi.nlm.nih.gov/19789228/

149. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: The chronic osteoarthritis management initiative of the U.S. bone and joint initiative. Semin Arthritis Rheum. 2014;43(6):701–12.

150. Zeng C, Wei J, Persson MSM, Sarmanova A, Doherty M, Xie D, et al. Relative efficacy and safety of topical non-steroidal anti-inflammatory drugs for osteoarthritis: a systematic review and network meta-analysis of randomised controlled trials and observational studies. Br J Sports Med [Internet]. 2018 May 1 [cited 2021 Dec 18];52(10):642–50. Available from: https://pubmed.ncbi.nlm.nih.gov/29436380/

151. Wang ZY, Shi SY, Li SJ, Chen F, Chen H, Lin HZ, et al. Efficacy and Safety of Duloxetine on Osteoarthritis Knee Pain: A Meta-Analysis of Randomized Controlled Trials. Pain Med [Internet]. 2015 [cited 2021 Dec 18];16(7):1373–85. Available from: https://pubmed.ncbi.nlm.nih.gov/26176791/

152. Jevsevar D, Donnelly P, Brown GA, Cummins DS. Viscosupplementation for Osteoarthritis of the Knee: A Systematic Review of the Evidence. J Bone Joint Surg Am [Internet]. 2015 Sep 2 [cited 2021 Dec 18];97(24):2047–60. Available from: https://pubmed.ncbi.nlm.nih.gov/26677239/

153. Johansen M, Bahrt H, Altman RD, Bartels EM, Juhl CB, Bliddal H, et al. Exploring reasons for the observed inconsistent trial reports on intra-articular injections with hyaluronic acid in the treatment of osteoarthritis: Meta-regression analyses of randomized trials. Semin Arthritis Rheum [Internet]. 2016 Aug 1 [cited 2021 Dec 18];46(1):34–48. Available from: https://pubmed.ncbi.nlm.nih.gov/27139169/

154. AAOS AA of OS. Management of Osteoarthritis of the Knee (Non-Arthroplasty) Evidence-Based Clinical Practice Guideline. 2021 [cited 2021 Dec 20]; Available from: https://www.aaos.org/oak3cpg

155. Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM, et al. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. BMJ [Internet]. 2013 Sep 20 [cited 2021 Dec 20];347(7928). Available from: https://www.bmj.com/content/347/bmj.f5555

156. Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. Ann Rheum Dis

[Internet]. 2007 Apr 1 [cited 2021 Dec 20];66(4):433–9. Available from: https://ard.bmj.com/content/66/4/433

157. Omidi A, Jamshidi Zanganeh M, Khodaveisi M, Mohammadi Y. The effect of self-management training on pain intensity in patients with knee osteoarthritis referring to orthopedic clinic of Imam Hossein Hospital in Malayer. Natl J Physiol Pharm Pharmacol [Internet]. 2018 [cited 2021 Dec 20]; Available from: www.njppp.com

158. Hurley M v., Walsh NE, Mitchell HL, Pimm TJ, Patel A, Williamson E, et al. Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: a cluster randomized trial. Arthritis Rheum [Internet]. 2007 Oct 15 [cited 2021 Dec 20];57(7):1211–9. Available from: https://pubmed.ncbi.nlm.nih.gov/17907147/

159. Saffari M, Emami Meybodi MK, Sanaeinasab H, Karami A, Pakpour AH, Koenig HG. A theory of planned behavior-based intervention to improve quality of life in patients with knee/hip osteoarthritis: a randomized controlled trial. Clin Rheumatol [Internet]. 2018 Sep 1 [cited 2021 Dec 20];37(9):2505–15. Available from: https://pubmed.ncbi.nlm.nih.gov/29744608/

160. Somers TJ, Blumenthal JA, Guilak F, Kraus VB, Schmitt DO, Babyak MA, et al. Pain coping skills training and lifestyle behavioral weight management in patients with knee osteoarthritis: a randomized controlled study. Pain [Internet]. 2012 Jun [cited 2021 Dec 20];153(6):1199–209. Available from: https://pubmed.ncbi.nlm.nih.gov/22503223/

161. Kroon FP, van der Burg LR, Buchbinder R, Osborne RH, Johnston R v., Pitt V. Self-management education programmes for osteoarthritis. Cochrane Database of Systematic Reviews [Internet]. 2014 Jan 15 [cited 2022 Dec 30];2014(1). Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD008963.pub2/full

162. Baker K, Goggins J, Xie H, Szumowski K, LaValley M, Hunter DJ, et al. A randomized crossover trial of a wedged insole for treatment of knee osteoarthritis. Arthritis Rheum [Internet]. 2007 Apr [cited 2021 Dec 20];56(4):1198–203. Available from: https://pubmed.ncbi.nlm.nih.gov/17393448/

163. Felson DT, Parkes M, Carter S, Liu A, Callaghan MJ, Hodgson R, et al. The Efficacy of a Lateral Wedge Insole for Painful Medial Knee Osteoarthritis After Prescreening: A Randomized Clinical Trial. Arthritis Rheumatol [Internet]. 2019 Jun 1 [cited 2021 Dec 20];71(6):908–15. Available from: https://pubmed.ncbi.nlm.nih.gov/30615299/

164. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis [Internet]. 2005 Jan 1 [cited 2022 Jun 22];64(1):29–33. Available from: https://ard.bmj.com/content/64/1/29

165. Thoumie P, Marty M, Avouac B, Pallez A, Vaumousse A, Pipet LPT, et al. Effect of unloading brace treatment on pain and function in patients with symptomatic knee osteoarthritis: the ROTOR randomized clinical trial. Scientific Reports 2018 8:1 [Internet]. 2018 Jul 12 [cited 2021 Dec 20];8(1):1–9. Available from: https://www.nature.com/articles/s41598-018-28782-3

166. Block JA. Osteoarthritis: OA guidelines: Improving care or merely codifying practice? Vol. 10, Nature Reviews Rheumatology. Nature Publishing Group; 2014. p. 324–6.

167. Manheimer E, Cheng K, Linde K, Lao L, Yoo J, Wieland S, et al. Acupuncture for peripheral joint osteoarthritis. Cochrane Database of Systematic Reviews. 2010 Jan 20;(1).

168. Stausholm MB, Naterstad IF, Joensen J, Lopes-Martins RÁB, Sæbø H, Lund H, et al. Efficacy of low-level laser therapy on pain and disability in knee osteoarthritis: systematic review and meta-analysis of randomised placebo-controlled trials. BMJ Open [Internet]. 2019 Oct 1 [cited 2021 Dec 20];9(10):e031142. Available from: https://bmjopen.bmj.com/content/9/10/e031142

169. Brosseau L, Robinson V, Wells GA, deBie R, Gam A, Harman K, et al. Low level laser therapy (Classes I, II and III) for treating osteoarthritis. Cochrane Database of Systematic Reviews [Internet]. 2004 Jul 19 [cited 2021 Dec 20];(3). Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD002046.pub2/full

170. Higashi H, Barendregt JJ. Cost-effectiveness of total hip and knee replacements for the Australian population with osteoarthritis: Discrete-event simulation model. PLoS One. 2011 Sep 23;6(9).

171. Culliford DJ, Maskell J, Kiran A, Judge A, Javaid MK, Cooper C, et al. The lifetime risk of total hip and knee arthroplasty: results from the UK general practice research database. Osteoarthritis Cartilage [Internet]. 2012 Jun [cited 2021 Dec 19];20(6):519–24. Available from: https://pubmed.ncbi.nlm.nih.gov/22395038/

172. Murphy LB, Helmick CG, Schwartz TA, Renner JB, Tudor G, Koch GG, et al. One in four people may develop symptomatic hip osteoarthritis in his or her lifetime. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society [Internet]. 2010 Nov [cited 2022 Jun 22];18(11):1372. Available from: /pmc/articles/PMC2998063/

173. Murphy L, Schwartz TA, Helmick CG, Renner JB, Tudor G, Koch G, et al. Lifetime risk of symptomatic knee osteoarthritis. Arthritis Care Res (Hoboken) [Internet]. 2008 Sep 15 [cited 2022 Jun 22];59(9):1207–13. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.24021

174. NHS. National Institute for Health and Clinical Excellence. Osteoarthritis: The care and management of osteoarthritis in adults. Clinical guideline 59. 2008 p. 25.

175. Singh JA, Lewallen D. Predictors of pain and use of pain medications following primary Total Hip Arthroplasty (THA): 5,707 THAs at 2-years and 3,289 THAs at 5-years. BMC Musculoskelet Disord [Internet]. 2010 [cited 2021 Dec 19];11(5):90. Available from: /pmc/articles/PMC2881019/

176. Wylde V, Hewlett S, Learmonth ID, Dieppe P. Persistent pain after joint replacement: prevalence, sensory qualities, and postoperative determinants. Pain [Internet]. 2011 Mar [cited 2021 Dec 19];152(3):566–72. Available from: https://pubmed.ncbi.nlm.nih.gov/21239114/

177. Beswick AD, Wylde V, Gooberman-Hill R, Blom A, Dieppe P. What proportion of patients report long-term pain after total hip or knee replacement for osteoarthritis? A systematic review of prospective studies in unselected patients. BMJ Open [Internet]. 2012 Jan 1 [cited 2022 Jun 22];2(1):e000435. Available from: https://bmjopen.bmj.com/content/2/1/e000435

178. Price AJ, Alvand A, Troelsen A, Katz JN, Hooper G, Gray A, et al. Hip and knee replacement 2 Knee replacement. www.thelancet.com [Internet]. 2018 [cited 2021 Dec 19];392. Available from: www.thelancet.com

179. Burn E, Liddle AD, Hamilton TW, Judge A, Pandit HG, Murray DW, et al. Cost-effectiveness of unicompartmental compared with total knee replacement: A population-based study using data from the National Joint Registry for England and Wales. BMJ Open [Internet]. 2018 Apr 1 [cited 2021 Dec 19];8(4):e020977. Available from: https://bmjopen.bmj.com/content/8/4/e020977

180. Price AJ, Alvand A, Troelsen A, Katz JN, Hooper G, Gray A, et al. Knee replacement. The Lancet [Internet]. 2018 Nov 3 [cited 2021 Dec 19];392(10158):1672–82. Available from: http://www.thelancet.com/article/S0140673618323444/fulltext

181. Brouwer RW, Huizinga MR, Duivenvoorden T, van Raaij TM, Verhagen AP, Bierma-Zeinstra SMA, et al. Osteotomy for treating knee osteoarthritis [Internet]. Vol. 2014, Cochrane Database of Systematic Reviews. John Wiley and Sons Ltd; 2014 [cited 2021 Dec 20]. Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD004019.pub4/full

182. Cao ZW, Mai XJ, Wang J, Feng EH, Huang YM. Uni-compartmental knee arthroplasty versus high tibial osteotomy for knee osteoarthritis: a systematic review and meta-analysis. Vol. 33, Journal of Arthroplasty. Churchill Livingstone; 2018. p. 952–9.

183. Siemieniuk RAC, Harris IA, Agoritsas T, Poolman RW, Brignardello-Petersen R, van de Velde S, et al. Arthroscopic surgery for degenerative knee arthritis and meniscal tears: a clinical practice guideline. BMJ [Internet]. 2017 May 10 [cited 2021 Dec 20];357. Available from: https://www.bmj.com/content/357/bmj.j1982

184. Richmond J, Hunter D, Irrgang J, Jones MH, Levy B, Marx R, et al. Treatment of osteoarthritis of the knee (Nonarthroplasty. Journal of the American Academy of Orthopaedic Surgeons [Internet]. 2009 [cited 2021 Dec 20];17(9):591–600. Available from: https://pubmed.ncbi.nlm.nih.gov/19726743/

185. Kise NJ, Risberg MA, Stensrud S, Ranstam J, Engebretsen L, Roos EM. Exercise therapy versus arthroscopic partial meniscectomy for degenerative meniscal tear in middle aged patients: randomised controlled trial with two year follow-up. BMJ [Internet]. 2016 [cited 2021 Dec 20];354. Available from: https://pubmed.ncbi.nlm.nih.gov/27440192/

186. Khan M, Evaniew N, Bedi A, Ayeni OR, Bhandari M. Arthroscopic surgery for degenerative tears of the meniscus: A systematic review and meta-analysis. CMAJ [Internet]. 2014 Oct 7 [cited 2021 Dec 20];186(14):1057–64. Available from: https://www.cmaj.ca/content/186/14/1057

187. Hawker G, Guan J, Judge A, Dieppe P. Knee arthroscopy in England and Ontario: Patterns of use, changes over time, and relationship to total knee replacement. Journal of Bone and Joint Surgery - Series A [Internet]. 2008 Nov 1 [cited 2021 Dec 20];90(11):2337–45. Available from: https://pubmed.ncbi.nlm.nih.gov/18978402/

188. Brignardello-Petersen R, Guyatt GH, Buchbinder R, Poolman RW, Schandelmaier S, Chang Y, et al. Knee arthroscopy versus conservative management in patients with degenerative knee

disease: a systematic review. BMJ Open [Internet]. 2017 May 1 [cited 2021 Dec 20];7(5):e016114. Available from: https://bmjopen.bmj.com/content/7/5/e016114

189. Bohensky MA, Desteiger R, Kondogiannis C, Sundararajan V, Andrianopoulos N, Bucknill A, et al. Adverse outcomes associated with elective knee arthroscopy: a population-based cohort study. Arthroscopy [Internet]. 2013 Apr [cited 2021 Dec 20];29(4):716–25. Available from: https://pubmed.ncbi.nlm.nih.gov/23395251/

190. Chapple CM, Nicholson H, Baxter GD, Abbott JH. Patient characteristics that predict progression of knee osteoarthritis: A systematic review of prognostic studies. Arthritis Care Res (Hoboken). 2011 Aug;63(8):1115–25.

191. Wesseling J, Dekker J, van den Berg WB, Bierma-Zeinstra SMA, Boers M, Cats HA, et al. CHECK (Cohort Hip and Cohort Knee): similarities and differences with the Osteoarthritis Initiative. Ann Rheum Dis [Internet]. 2009 Sep [cited 2021 Nov 26];68(9):1413–9. Available from: https://pubmed.ncbi.nlm.nih.gov/18772189/

192. Karsdal MA, Bay-Jensen AC, Lories RJ, Abramson S, Spector T, Pastoureau P, et al. The coupling of bone and cartilage turnover in osteoarthritis: opportunities for bone antiresorptives and anabolics as potential treatments? Ann Rheum Dis [Internet]. 2014 Feb [cited 2021 Nov 26];73(2):336–48. Available from: https://pubmed.ncbi.nlm.nih.gov/24285494/

193. Deveza LA, Melo L, Yamato TP, Mills K, Ravi V, Hunter DJ. Knee osteoarthritis phenotypes and their relevance for outcomes: a systematic review. Osteoarthritis Cartilage. 2017 Dec 1;25(12):1926–41.

194. Bijlsma JWJ, Berenbaum F, Lafeber FPJG. Osteoarthritis: An update with relevance for clinical practice. The Lancet. 2011 Jun 18;377(9783):2115–26.

195. Vadivelu N, Kai AM, Kodumudi V, Sramcik J, Kaye AD. The Opioid Crisis: a Comprehensive Overview. Current Pain and Headache Reports 2018 22:3 [Internet]. 2018 Feb 23 [cited 2021 Nov 26];22(3):1–6. Available from: https://link.springer.com/article/10.1007/s11916-018-0670-z

196. Lee AS, Ellman MB, Yan D, Kroin JS, Cole BJ, van Wijnen AJ, et al. A current review of molecular mechanisms regarding osteoarthritis and pain. Gene [Internet]. 2013 Sep 25 [cited 2021 Nov 26];527(2):440–7. Available from: https://pubmed.ncbi.nlm.nih.gov/23830938/

197. Hayashi D, Roemer FW, Jarraya M, Guermazi A. Imaging of Osteoarthritis. Geriatric Imaging [Internet]. 2013 Oct 1 [cited 2021 Dec 27];9783642355790:93–121. Available from: https://link.springer.com/chapter/10.1007/978-3-642-35579-0_5

198. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage. 2008 Dec;16(12):1433–41.

199. Shapiro L, Staroswiecki E, Gold G. MRI of the Knee: Optimizing 3T Imaging. Semin Roentgenol [Internet]. 2010 Oct [cited 2021 Dec 27];45(4):238. Available from: /pmc/articles/PMC2941506/

200. del Grande F, Rashidi A, Luna R, Delcogliano M, Stern SE, Dalili D, et al. Five-minute five-sequence knee MRI using combined simultaneous multislice and parallel imaging

acceleration: Comparison with 10-minute parallel imaging knee MRI. Radiology [Internet]. 2021 Jun 1 [cited 2021 Dec 27];299(3):635–46. Available from: https://pubs.rsna.org/doi/abs/10.1148/radiol.2021203655

201. Compressed sensing | IEEE Journals & Magazine | IEEE Xplore [Internet]. [cited 2021 Dec 27]. Available from: https://ieeexplore.ieee.org/abstract/document/4472246

202. Peng X, Ying L, Liu Y, Yuan J, Liu X, Liang D. Accelerated exponential parameterization of T2 relaxation with model-driven low rank and sparsity priors (MORASA). Magn Reson Med [Internet]. 2016 Dec 1 [cited 2021 Dec 27];76(6):1865–78. Available from: https://pubmed.ncbi.nlm.nih.gov/26762702/

203. Zibetti MVW, Sharafi A, Otazo R, Regatte RR. Accelerating 3D-T 1ρ mapping of cartilage using compressed sensing with different sparse and low rank models. Magn Reson Med [Internet]. 2018 Oct 1 [cited 2021 Dec 27];80(4):1475–91. Available from: https://pubmed.ncbi.nlm.nih.gov/29479738/

204. Kijowski R, Rosas H, Samsonov A, King K, Peters R, Liu F. Knee imaging: Rapid three-dimensional fast spin-echo using compressed sensing. Journal of Magnetic Resonance Imaging [Internet]. 2017 Jun 1 [cited 2021 Dec 27];45(6):1712–22. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/jmri.25507

205. Terzis R, Dratsch T, Hahnfeldt R, Basten L, Rauen P, Sonnabend K, et al. Five-minute knee MRI: An AI-based super resolution reconstruction approach for compressed sensing. A validation study on healthy volunteers. Eur J Radiol. 2024 Jun 1;175:111418.

206. Jaspan ON, Fleysher R, Lipton ML. Compressed sensing MRI: A review of the clinical literature [Internet]. Vol. 88, British Journal of Radiology. British Institute of Radiology; 2015 [cited 2021 Dec 28]. Available from: /pmc/articles/PMC4984938/

207. Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, et al. Learning a Variational Network for Reconstruction of Accelerated MRI Data. Magn Reson Med [Internet]. 2018 Jun 1 [cited 2024 Apr 18];79(6):3055. Available from: /pmc/articles/PMC5902683/

208. Pillai K. K213693. [cited 2024 Apr 18]; Available from: https://www.fda.gov/medical-

209. Cicuttini F. Osteoarthritis year in review: Imaging. Osteoarthritis Cartilage [Internet]. 2015 [cited 2021 Dec 28];23:A24. Available from: https://www.sciencedirect.com/science/article/pii/S106345841831584X

210. Toffanin R, Mlynárik V, Russo S, Szomolányi P, Piras A, Vittura F. Proteoglycan Depletion and Magnetic Resonance Parameters of Articular Cartilage. Arch Biochem Biophys. 2001 Jun 15;390(2):235–42.

211. Regatte RR, Akella SVS, Borthakur A, Kneeland JB, Reddy R. Proteoglycan Depletion–Induced Changes in Transverse Relaxation Maps of Cartilage: Comparison of T2 and T1ρ. Acad Radiol. 2002 Dec 1;9(12):1388–94.

212. Gray ML, Burstein D, Xia Y. Biochemical (and functional) imaging of articular cartilage. Semin Musculoskelet Radiol [Internet]. 2001 [cited 2024 Jan 10];5(4):329–43. Available from: http://www.thieme-connect.com/products/ejournals/html/10.1055/s-2001-19043

213. Li X, Benjamin Ma C, Link TM, Castillo DD, Blumenkrantz G, Lozano J, et al. In vivo T1ρ and T2 mapping of articular cartilage in osteoarthritis of the knee using 3 T MRI. Osteoarthritis Cartilage. 2007 Jul 1;15(7):789–97.

214. Hilbert T, Sumpf TJ, Weiland E, Frahm J, Thiran JP, Meuli R, et al. Accelerated T2 mapping combining parallel MRI and model-based reconstruction: GRAPPATINI. Wiley Online Library [Internet]. 2018 Aug 1 [cited 2022 Jan 10];48(2):359–68. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.25972?casa_token=xJA_ZDztx-8AAAAA:gZdJsXPfGZFjeCYdc1XrjIKtHyiW3o5hbUN5GG8lYx3CpiEqLtHV_VPj0lMm_gA2V8B30xmtUeJ19w

215. Braun HJ, Gold GE. Diagnosis of osteoarthritis: Imaging. Bone. 2012 Aug 1;51(2):278–88.

216. Wheaton AJ, Casey FL, Gougoutas AJ, Dodge GR, Borthakur A, Lonner JH, et al. Correlation of T1ρ with fixed charge density in cartilage. Journal of Magnetic Resonance Imaging [Internet]. 2004 Sep 1 [cited 2022 Jan 10];20(3):519–25. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/jmri.20148

217. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J [Internet]. 2019 Jun [cited 2022 Jun 22];6(2):94. Available from: /pmc/articles/PMC6616181/

218. Turing AM. Computing machinery and intelligence. Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer. 2009;23–65.

219. Krizhevsky A, … ISA in neural, 2012 undefined. Imagenet classification with deep convolutional neural networks. proceedings.neurips.cc [Internet]. [cited 2022 Jun 22]; Available from: https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

220. Rawat W, Wang Z. Deep convolutional neural networks for image classification: A comprehensive review. Neural Comput. 2017 Sep 1;29(9):2352–449.

221. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med [Internet]. [cited 2022 Jun 22]; Available from: https://doi.org/10.1038/s41591-018-0300-7

222. Mayer H, Gomez F, Wierstra D, Nagy I, Knoll A, Schmidhuber J. A System for Robotic Heart Surgery that Learns to Tie Knots Using Recurrent Neural Networks. 2006;

223. Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti A. A primer on deep learning in genomics. Nature Genetics 2018 51:1 [Internet]. 2018 Nov 26 [cited 2022 Jun 22];51(1):12–8. Available from: https://www.nature.com/articles/s41588-018-0295-5

224. Biotechnology ESN, 2017 undefined. AI-powered drug discovery captures pharma interest. go.gale.com [Internet]. [cited 2022 Jun 22]; Available from: https://go.gale.com/ps/i.do?id=GALE%7CA498337436&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=10870156&p=HRCA&sw=w

225. Banaee H, Ahmed MU, Loutfi A. Data Mining for Wearable Sensors in Health Monitoring Systems: A Review of Recent Trends and Challenges. Sensors 2013, Vol 13, Pages 17472-17500 [Internet]. 2013 Dec 17 [cited 2022 Jun 22];13(12):17472–500. Available from: https://www.mdpi.com/1424-8220/13/12/17472/htm

226. Hirschmann A, Cyriac J, Stieltjes B, Kober T, Richiardi J, Omoumi P. Artificial Intelligence in Musculoskeletal Imaging: Review of Current Literature, Challenges, and Trends. Semin Musculoskelet Radiol [Internet]. 2019 Jun 4 [cited 2021 Dec 30];23(3):304–11. Available from: http://www.thieme-connect.com/products/ejournals/html/10.1055/s-0039-1684024

227. (MRI[MeSH Major Topic]) AND ((artificial intelligence[MeSH Major Topic])) - Search Results - PubMed [Internet]. [cited 2021 Dec 30]. Available from: https://pubmed.ncbi.nlm.nih.gov/?term=%28MRI%5BMeSH+Major+Topic%5D%29+AND+%28%28artificial+intelligence%5BMeSH+Major+Topic%5D%29%29

228. Groot OQ, Bongers MER, Ogink PT, Senders JT, Karhade A v., Bramer JAM, et al. Does Artificial Intelligence Outperform Natural Intelligence in Interpreting Musculoskeletal Radiological Studies? A Systematic Review. Clin Orthop Relat Res [Internet]. 2020 Dec 1 [cited 2021 Dec 30];478(12):2751–64. Available from: https://pubmed.ncbi.nlm.nih.gov/32740477/

229. Lindsey R, Daluiski A, Chopra S, Lachapelle A, Mozer M, Sicular S, et al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci U S A [Internet]. 2018 Nov 6 [cited 2021 Dec 30];115(45):11591–6. Available from: https://www.pnas.org/content/115/45/11591

230. Bien N, Rajpurkar P, Ball RL, Irvin J, Park A, Jones E, et al. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: Development and retrospective validation of MRNet. PLoS Med [Internet]. 2018 Nov 1 [cited 2021 Dec 30];15(11):e1002699. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002699

231. von Schacky CE, Sohn JH, Liu F, Ozhinsky E, Jungmann PM, Nardo L, et al. Development and validation of a multitask deep learning model for severity grading of hip osteoarthritis features on radiographs. Radiology [Internet]. 2020 Feb 4 [cited 2021 Dec 31];295(1):139–45. Available from: https://pubs.rsna.org/doi/abs/10.1148/radiol.2020190925

232. Krogue JD, Cheng K v., Hwang KM, Toogood P, Meinberg EG, Geiger EJ, et al. Automatic Hip Fracture Identification and Functional Subclassification with Deep Learning. Radiol Artif Intell [Internet]. 2020 Mar 1 [cited 2021 Dec 31];2(2):e190023. Available from: https://pubs.rsna.org/doi/abs/10.1148/ryai.2020190023

233. Galbusera F, Bassani T, Casaroli G, Gitto S, Zanchetta E, Costa F, et al. Generative models: an upcoming innovation in musculoskeletal radiology? A preliminary test in spine imaging. Eur Radiol Exp [Internet]. 2018 Dec 1 [cited 2021 Dec 31];2(1):1–13. Available from: https://eurradiolexp.springeropen.com/articles/10.1186/s41747-018-0060-7

234. Jans LBO, Chen M, Elewaut D, van den Bosch F, Carron P, Jacques P, et al. MRI-based Synthetic CT in the Detection of Structural Lesions in Patients with Suspected Sacroiliitis: Comparison with MRI. Radiology [Internet]. 2021 Feb 1 [cited 2021 Dec 31];298(2):343–9. Available from: https://pubs.rsna.org/doi/abs/10.1148/radiol.2020201537

235. Calivá F, Leynes AP, Shah R, Bharadwaj UU, Majumdar S, Larson PEZ, et al. Breaking Speed Limits with Simultaneous Ultra-Fast MRI Reconstruction and Tissue Segmentation [Internet]. Vol. 121, Proceedings of Machine Learning Research. PMLR; 2020 [cited 2021 Dec 27]. p. 94–110. Available from: https://proceedings.mlr.press/v121/caliva20a.html

236. Peters RD, Marketing GP. The clinical benefits of AIR ™ Recon DL for MR image reconstruction The limitations of.

237. Lebel RM, Healthcare GE. Performance characterization of a novel deep learning-based MR image reconstruction pipeline.

238. Antun V, Renna F, Poon C, Adcock B, Hansen AC. On instabilities of deep learning in image reconstruction and the potential costs of AI. Proc Natl Acad Sci U S A [Internet]. 2020 Dec 1 [cited 2024 Apr 18];117(48):30088–95. Available from: https://pubmed.ncbi.nlm.nih.gov/32393633/

239. Case Studies - IPEM [Internet]. [cited 2024 Apr 18]. Available from: https://www.ipem.ac.uk/resources/mr-advanced-acceleration-technology/case-studies/

240. Michael Nevitt AC, Felson DT, Lester G. OAI Protocol THE OSTEOARTHRITIS INITIATIVE PROTOCOL FOR THE COHORT STUDY [Internet]. [cited 2019 Jul 30]. Available from: https://oai.epi-ucsf.org/datarelease/docs/StudyDesignProtocol.pdf

241. MA K, I B, P A, A B, JR A, BJ R, et al. Treatment of symptomatic knee osteoarthritis with oral salmon calcitonin: results from two phase 3 trials. Osteoarthritis Cartilage [Internet]. 2015 Apr 1 [cited 2021 Oct 27];23(4):532–43. Available from: https://pubmed.ncbi.nlm.nih.gov/25582279/

242. TE M, MP L, WF H, LL P, JB D, M Z, et al. Effect of Intra-articular Triamcinolone vs Saline on Knee Cartilage Volume and Pain in Patients With Knee Osteoarthritis: A Randomized Clinical Trial. JAMA [Internet]. 2017 May 16 [cited 2021 Oct 27];317(19):1967–75. Available from: https://pubmed.ncbi.nlm.nih.gov/28510679/

243. Hunter DJ, March L, Sambrook PN. The association of cartilage volume with knee pain. Osteoarthritis Cartilage. 2003 Oct 1;11(10):725–9.

244. OAI:DATA FROM 3D CARTILAGE/MENISCUS SEGMENTATIONS OF KNEE MRI SCANS [Internet]. [cited 2020 Apr 2]. Available from: https://oai.epi-ucsf.org/datarelease/iMorphics.asp

245. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology [Internet]. 2020 May 1 [cited 2020 Sep 17];295(2):328–38. Available from: https://pubs.rsna.org/doi/abs/10.1148/radiol.2020191145

246. Zhang WJ, Yang G, Lin Y, Ji C, Gupta MM. On Definition of Deep Learning. World Automation Congress Proceedings. 2018 Aug 8;2018-June:232–6.

247. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May 27;521(7553):436–44.

248. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks [Internet]. [cited 2019 Aug 5]. Available from: http://code.google.com/p/cuda-convnet/

249. Learning Hierarchical Features for Scene Labeling [Internet]. [cited 2019 Aug 5]. Available from: http://yann.lecun.com/exdb/publis/pdf/farabet-pami-13.pdf

250. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going Deeper with Convolutions. 2014 Sep 16 [cited 2019 Aug 5]; Available from: http://arxiv.org/abs/1409.4842

251. Mikolov T, Deoras A, Povey D, Lukáš B, #4 J", Honza " ˇ, et al. Strategies for Training Large Scale Neural Network Language Models [Internet]. [cited 2019 Aug 5]. Available from: https://www.microsoft.com/en-us/research/wp-content/uploads/2011/12/ASRU-2011.pdf

252. Hinton G, Deng L, Yu D, Dahl G, Mohamed AR, Jaitly N, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition [Internet]. [cited 2019 Aug 5]. Available from: https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/38131.pdf

253. Sainath TN, Mohamed AR, Kingsbury B, Ramabhadran B. DEEP CONVOLUTIONAL NEURAL NETWORKS FOR LVCSR [Internet]. [cited 2019 Aug 5]. Available from: http://www.cs.toronto.edu/~asamir/papers/icassp13_cnn.pdf

254. Zhang L, Tan J, Han D, Zhu H. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug Discov Today [Internet]. 2017 Nov [cited 2019 Aug 5];22(11):1680–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28881183

255. Deep learning for genomics. Nat Genet [Internet]. 2019 Jan 21 [cited 2019 Aug 5];51(1):1–1. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30578416

256. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature [Internet]. 2018 Mar 22 [cited 2019 Aug 1];555(7697):487–92. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29565357

257. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. Med Phys [Internet]. 2017 Apr [cited 2019 Aug 5];44(4):1408–19. Available from: http://doi.wiley.com/10.1002/mp.12155

258. Chaudhari AS, Fang Z, Kogan F, Wood J, Stevens KJ, Gibbons EK, et al. Super-resolution musculoskeletal MRI using deep learning. Magn Reson Med [Internet]. 2018 Nov 1 [cited 2019 Jul 30];80(5):2139–54. Available from: http://doi.wiley.com/10.1002/mrm.27178

259. Shen D, Wu G, Suk HI. Deep Learning in Medical Image Analysis. Annu Rev Biomed Eng. 2017;19:221–48.

260. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. 2016 Dec 21;

261. Artificial Neural Network | Brilliant Math & Science Wiki [Internet]. [cited 2019 Aug 7]. Available from: https://brilliant.org/wiki/artificial-neural-network/#citation-7

262. Activation functions in Neural Networks - GeeksforGeeks [Internet]. [cited 2019 Aug 7]. Available from: https://www.geeksforgeeks.org/activation-functions-neural-networks/

263. Backpropagation | Brilliant Math & Science Wiki [Internet]. [cited 2019 Aug 6]. Available from: https://brilliant.org/wiki/backpropagation/

264. Gulli A. Deep learning with Keras : implement neural networks with Keras on Theano and TensorFlow [Internet]. [cited 2019 Aug 5]. Available from: https://books.google.co.uk/books?hl=en&lr=&id=20EwDwAAQBAJ&oi=fnd&pg=PP1&dq=keras&ots=lHeA4hbOT2&sig=HX0_aYYcSYXPi8R4r_LR4RQZnTI#v=onepage&q=keras&f=false

265. Liang X, Wang X, Lei Z, Liao S, Li SZ. Soft-Margin Softmax for Deep Classification. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) [Internet]. 2017 Nov 14 [cited 2021 Dec 31];10635 LNCS:413–21. Available from: https://link.springer.com/chapter/10.1007/978-3-319-70096-0_43

266. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer [Internet]. 2012 Mar [cited 2022 Jan 3];48(4):441–6. Available from: /pmc/articles/PMC4533986/

267. Micheel CM, Nass SJ, Omenn GS. Committee on the Review of Omics-Based Tests for Predicting Patient Outcomes in Clinical Trials; Board on Health Care Services; Board on Health Sciences Policy. 2012 Mar 23 [cited 2022 Jan 3]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK202160/

268. Gillies RJ, Anderson AR, Gatenby RA, Morse DL. The biology underlying molecular imaging in oncology: from genome to anatome and back again. Clin Radiol [Internet]. 2010 Jul [cited 2022 Jan 3];65(7):517–21. Available from: https://pubmed.ncbi.nlm.nih.gov/20541651/

269. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. Radiology. 2016 Feb 1;278(2):563–77.

270. Lee G, Lee HY, Park H, Schiebler ML, van Beek EJR, Ohno Y, et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. Eur J Radiol [Internet]. 2017 Jan 1 [cited 2022 Jan 3];86:297–307. Available from: https://pubmed.ncbi.nlm.nih.gov/27638103/

271. Gatenby RA, Grove O, Gillies RJ. Quantitative imaging in cancer evolution and ecology. Radiology [Internet]. 2013 Oct 1 [cited 2022 Jan 3];269(1):8–15. Available from: https://pubs.rsna.org/doi/abs/10.1148/radiol.13122697

272. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol [Internet]. 2017 Dec 1 [cited 2022 Jan 3];14(12):749–62. Available from: https://pubmed.ncbi.nlm.nih.gov/28975929/

273. Kim S, Lee D, Park S, Oh KS, Chung SW, Kim Y. Automatic segmentation of supraspinatus from MRI by internal shape fitting and autocorrection. Comput Methods Programs Biomed [Internet]. 2017 Mar 1 [cited 2022 Jan 3];140:165–74. Available from: https://pubmed.ncbi.nlm.nih.gov/28254072/

274. Dalvi R, Abugharbieh R, Wilson DC, Wilson DR. Multi-contrast MR for enhanced bone imaging and segmentation. Annu Int Conf IEEE Eng Med Biol Soc [Internet]. 2007 [cited 2022 Jan 3];2007:5620–3. Available from: https://pubmed.ncbi.nlm.nih.gov/18003287/

275. Zhang J, Yan CH, Chui CK, Ong SH. Fast segmentation of bone in CT images using 3D adaptive thresholding. Comput Biol Med [Internet]. 2010 Feb [cited 2022 Jan 3];40(2):231–6. Available from: https://pubmed.ncbi.nlm.nih.gov/20053396/

276. Li W, Abram F, Pelletier JP, Raynauld JP, Dorais M, d'Anjou MA, et al. Fully automated system for the quantification of human osteoarthritic knee joint effusion volume using magnetic resonance imaging. Arthritis Res Ther [Internet]. 2010 Sep 16 [cited 2022 Jan 3];12(5):R173. Available from: /pmc/articles/PMC2991000/

277. Li X, Ma BC, Bolbos RI, Stahl R, Lozano J, Zuo J, et al. Quantitative assessment of bone marrow edema-like lesion and overlying cartilage in knees with osteoarthritis and anterior cruciate ligament tear using MR imaging and spectroscopic imaging at 3 Tesla. Journal of Magnetic

Resonance Imaging [Internet]. 2008 Aug 1 [cited 2022 Jan 3];28(2):453–61. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/jmri.21437

278. Heimann T, Morrison B. Segmentation of knee images: A grand challenge. … Image Analysis for the … [Internet]. 2010 [cited 2020 Mar 16];207–14. Available from: www.ski10.org

279. Desai AD, Caliva F, Iriondo C, Khosravan N, Mortazi A, Jambawalikar S, et al. The international workshop on osteoarthritis imaging knee MRI segmentation challenge: A multi-institute evaluation and analysis framework on a standardized dataset [Internet]. arXiv. arXiv; 2020 [cited 2021 Mar 12]. Available from: https://pubs.rsna.org/doi/abs/10.1148/ryai.2021200078

280. Cabezas M, Oliver A, Lladó X, Freixenet J, Bach Cuadra M. A review of atlas-based segmentation for magnetic resonance brain images. Comput Methods Programs Biomed. 2011 Dec 1;104(3):e158–77.

281. Shan L, Zach C, Charles C, Niethammer M. Automatic atlas-based three-label cartilage segmentation from MR knee images. Med Image Anal. 2014 Oct 1;18(7):1233–46.

282. Carballido-Gamio J, Majumdar S. Atlas-based knee cartilage assessment. Magn Reson Med [Internet]. 2011 Aug 1 [cited 2022 Jan 5];66(2):575–81. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/mrm.22836

283. Shan L, Charles C, Niethammer M. Automatic multi-atlas-based cartilage segmentation from knee MR images. In: Proceedings - International Symposium on Biomedical Imaging. NIH Public Access; 2012. p. 1028–31.

284. Arthofer C, Zedan Y, Tench C, Kurien T, Scammell B, Kerslake R, et al. Construction of an anatomical template of the knee and voxel-based knee pathology-pain mapping. Osteoarthritis Cartilage. 2021 Apr 1;29:S348–9.

285. Liu F, Zhou Z, Jang H, Samsonov A, Zhao G, Kijowski R. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. Magn Reson Med [Internet]. 2018 Apr 1 [cited 2022 Jan 5];79(4):2379–91. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/mrm.26841

286. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Advances in Neural Information Processing Systems [Internet]. 2014 [cited 2022 Jan 6]. p. 2672–80. Available from: http://www.github.com/goodfeli/adversarial

287. Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Trans Pattern Anal Mach Intell. 2017 Dec 1;39(12):2481–95.

288. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag; 2015. p. 234–41.

289. Johnstone IM, Titterington DM. Statistical challenges of high-dimensional data [Internet]. Vol. 367, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. The Royal Society Publishing; 2009 [cited 2022 Jan 7]. p. 4237–53. Available from: https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2009.0159

290. Han Z, Wei B, Mercado A, Leung S, Li S. Spine-GAN: Semantic segmentation of multiple spinal structures. Med Image Anal [Internet]. 2018 Dec 1 [cited 2022 Jan 7];50:23–35. Available from: https://pubmed.ncbi.nlm.nih.gov/30176546/

291. Zhou Z, Zhao G, Kijowski R, Liu F. Deep convolutional neural network for segmentation of knee joint anatomy. Magn Reson Med. 2018 Dec 1;80(6):2759–70.

292. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging [Internet]. 2015 Aug 12 [cited 2021 Dec 31];15(1):29. Available from: /pmc/articles/PMC4533825/

293. Shamir RR, Duchin Y, Kim J, Sapiro G, Harel N. Continuous Dice Coefficient: a Method for Evaluating Probabilistic Segmentations. 2019 [cited 2021 Dec 31]; Available from: http://dx.doi.org/10.1101/306977

294. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. Eur J Nucl Med Mol Imaging [Internet]. 2019 Dec 1 [cited 2022 Jan 3];46(13):2638–55. Available from: https://link.springer.com/article/10.1007/s00259-019-04391-8

295. Diamant A, Chatterjee A, Vallières M, Shenouda G, Seuntjens J. Deep learning in head & neck cancer outcome prediction. Scientific Reports 2019 9:1 [Internet]. 2019 Feb 26 [cited 2022 Jan 3];9(1):1–10. Available from: https://www.nature.com/articles/s41598-019-39206-1

296. Forghani R, Savadjiev P, Chatterjee A, Muthukrishnan N, Reinhold C, Forghani B. Radiomics and Artificial Intelligence for Biomarker and Prediction Model Development in Oncology. Comput Struct Biotechnol J [Internet]. 2019 Jan 1 [cited 2022 Jan 3];17:995. Available from: /pmc/articles/PMC6667772/

297. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. International Journal of Radiation Oncology*Biology*Physics. 2018 Nov 15;102(4):1143–58.

298. Polan DF, Brady SL, Kaufman RA. Tissue segmentation of computed tomography images using a Random Forest algorithm: A feasibility study. Phys Med Biol [Internet]. 2016 Aug 17 [cited 2022 Jan 3];61(17):6553–69. Available from: https://iopscience.iop.org/article/10.1088/0031-9155/61/17/6553

299. Bellman RE. Adaptive Control Processes. Adaptive Control Processes. 1961 Dec 31;

300. Jiang Y, Guo P. Regularization Versus Dimension Reduction, Which Is Better? Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) [Internet]. 2007 [cited 2022 Jan 7];4492 LNCS(PART 2):474–82. Available from: https://link.springer.com/chapter/10.1007/978-3-540-72393-6_57

301. Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 521:7553. 2015 May 27;521(7553):436–44.

302. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and Intelligent Laboratory Systems. 1987 Aug 1;2(1–3):37–52.

303. Principal Component Analysis. 2002 [cited 2022 Jan 3]; Available from: http://link.springer.com/10.1007/b98835

304. Lever J, Krzywinski M, Altman N. Principal component analysis. Vol. 14, Nature Methods. Nature Publishing Group; 2017. p. 641–2.

305. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography [Internet]. 2013 Jan 1 [cited 2022 Jan 3];36(1):27–46. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1600-0587.2012.07348.x

306. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol [Internet]. 2005 Apr 1 [cited 2022 Jan 3];67(2):301–20. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1467-9868.2005.00503.x

307. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. J Stat Softw. 2011;39(5):1–13.

308. Ogutu JO, Schulz-Streeck T, Piepho HP. Genomic selection using regularized linear regression models: Ridge regression, lasso, elastic net and their extensions. BMC Proc. 2012;6(SUPPL. 2).

309. Park H, Lim Y, Ko ES, Cho HH, Lee JE, Han BK, et al. Radiomics Signature on Magnetic Resonance Imaging: Association with Disease-Free Survival in Patients with Invasive Breast Cancer. Clinical Cancer Research [Internet]. 2018 Oct 1 [cited 2022 Jan 4];24(19):4705–14. Available from: https://clincancerres.aacrjournals.org/content/24/19/4705

310. Huang S ying, Franc BL, Harnish RJ, Liu G, Mitra D, Copeland TP, et al. Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis. npj Breast Cancer 2018 4:1 [Internet]. 2018 Aug 16 [cited 2022 Jan 4];4(1):1–13. Available from: https://www.nature.com/articles/s41523-018-0078-2

311. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society: Series B (Methodological). 1972 Jan;34(2):187–202.

312. Fox J, to SWAR and SP companion, 2002 undefined. Cox proportional-hazards regression for survival data. socialsciences.mcmaster.ca [Internet]. [cited 2022 Jan 4]; Available from: https://socialsciences.mcmaster.ca/jfox/Books/Companion-2E/appendix/Appendix-Cox-Regression.pdf

313. Conaghan PG, D'Agostino MA, le Bars M, Baron G, Schmidely N, Wakefield R, et al. Clinical and ultrasonographic predictors of joint replacement for knee osteoarthritis: results from a large, 3-year, prospective EULAR study. Ann Rheum Dis [Internet]. 2010 Apr 1 [cited 2022 Jan 4];69(4):644–7. Available from: https://ard.bmj.com/content/69/4/644

314. Swanton JK, Rovira A, Tintore M, Altmann DR, Barkhof F, Filippi M, et al. MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study. Lancet Neurol. 2007 Aug 1;6(8):677–86.

315. Desikan RS, Cabral HJ, Settecase F, Hess CP, Dillon WP, Glastonbury CM, et al. Automated MRI measures predict progression to Alzheimer's disease. Neurobiol Aging. 2010 Aug 1;31(8):1364–74.

316. Heinze G, Wallisch C, Dunkler D. Variable selection – A review and recommendations for the practicing statistician. Biom J [Internet]. 2018 May 1 [cited 2022 Jan 4];60(3):431. Available from: /pmc/articles/PMC5969114/

317. Hayes DF, Bast RC, Desch CE, Fritsche H, Kemeny NE, Jessup JM, et al. Tumor marker utility grading system: A framework to evaluate clinical utility of tumor markers. J Natl Cancer Inst [Internet]. 1996 Oct 16 [cited 2022 Jan 3];88(20):1456–66. Available from: https://pubmed.ncbi.nlm.nih.gov/8841020/

318. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst [Internet]. 2009 Nov [cited 2022 Jan 3];101(21):1446–52. Available from: https://pubmed.ncbi.nlm.nih.gov/19815849/

319. Morley J, Floridi L. NHS AI Lab: Why We Need to Be Ethically Mindful About AI for Healthcare. SSRN Electronic Journal [Internet]. 2019 Aug 30 [cited 2022 Jan 3]; Available from: https://papers.ssrn.com/abstract=3445421

320. Cho H, of HP 2017 39th AIC, 2017 undefined. Classification of low-grade and high-grade glioma using multi-modal image radiomics features. ieeexplore.ieee.org [Internet]. [cited 2022 Dec 30]; Available from: https://ieeexplore.ieee.org/abstract/document/8037508/

321. Sun P, Wang D, Mok V, Access LSI, 2019 undefined. Comparison of feature selection methods and machine learning classifiers for radiomics analysis in glioma grading. ieeexplore.ieee.org [Internet]. [cited 2022 Dec 30]; Available from: https://ieeexplore.ieee.org/abstract/document/8763934/

322. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. Eur Radiol Exp. 2018 Dec 1;2(1).

323. Eckstein F, Cicuttini F, Raynauld JP, Waterton JC, Peterfy C. Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): morphological assessment. Osteoarthritis Cartilage. 2006;14(SUPPL. 1):46–75.

324. Wang Y, Wluka AE, Cicuttini FM, Jones G, Ding C. Use magnetic resonance imaging to assess articular cartilage. Vol. 4, Therapeutic Advances in Musculoskeletal Disease. SAGE Publications; 2012. p. 77–97.

325. Eckstein F, Cotofana S, Wirth W, Nevitt M, John MR, Dreher D, et al. Greater rates of cartilage loss in painful knees than in pain-free knees after adjustment for radiographic disease stage: Data from the osteoarthritis initiative. Arthritis Rheum [Internet]. 2011 Aug [cited 2020 Apr 21];63(8):2257–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21520009

326. Calvo E, Palacios I, Delgado E, Ruiz-Cabello J, Hernández P, Sánchez-Pernaute O, et al. High-resolution MRI detects cartilage swelling at the early stages of experimental osteoarthritis. Osteoarthritis Cartilage [Internet]. 2001 [cited 2021 Mar 9];9(5):463–72. Available from: https://pubmed.ncbi.nlm.nih.gov/11467895/

327. Watson PJ, Carpenter TA, Hall LD, Tyler JA. Cartilage swelling and loss in a spontaneous model of osteoarthritis visualized by magnetic resonance imaging. In: Osteoarthritis and Cartilage. W.B. Saunders Ltd; 1996. p. 197–207.

328. Guermazi A, Alizai H, Crema MD, Trattnig S, Regatte RR, Roemer FW. Compositional MRI techniques for evaluation of cartilage degeneration in osteoarthritis. Vol. 23, Osteoarthritis and Cartilage. W.B. Saunders Ltd; 2015. p. 1639–53.

329. Baum T, Joseph GB, Karampinos DC, Jungmann PM, Link TM, Bauer JS. Cartilage and meniscal T2 relaxation time as non-invasive biomarker for knee osteoarthritis and cartilage repair procedures. Vol. 21, Osteoarthritis and Cartilage. W.B. Saunders; 2013. p. 1474–84.

330. Gatti AA. NEURALSEG: state-of-the-art cartilage segmentation using deep learning – analyses of data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2018 Apr 1;26:S47–8.

331. Zhou Z, Zhao G, Kijowski R, Liu F. Deep convolutional neural network for segmentation of knee joint anatomy. Magn Reson Med. 2018 Dec;80(6):2759–70.

332. Tack A, Mukhopadhyay A, Zachow S. Knee menisci segmentation using convolutional neural networks: data from the Osteoarthritis Initiative. Osteoarthritis Cartilage. 2018 May 1;26(5):680–8.

333. Ambellan F, Tack A, Ehlke M, Zachow S. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the Osteoarthritis Initiative. Med Image Anal. 2019 Feb 1;52:109–18.

334. Tack A, Zachow S. Accurate Automated Volumetry of Cartilage of the Knee using Convolutional Neural Networks: Data from the Osteoarthritis Initiative. 2019;

335. Siorpaes K, Wenger A, Bloecker K, Wirth W, Hudelmaier M, Eckstein F. Interobserver reproducibility of quantitative meniscus analysis using coronal multiplanar DESS and IWTSE MR imaging. Magn Reson Med. 2012 May;67(5):1419–26.

336. Prasoon A, Petersen K, Igel C, Lauze F, Dam E, Nielsen M. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2013. p. 246–53.

337. Norman B, Pedoia V, Majumdar S. Use of 2D U-net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry. Radiology. 2018 Jul 27;288(1):177–85.

338. Raj A, Vishwanathan S, Ajani B, Krishnan K, Agarwal H. Automatic knee cartilage segmentation using fully volumetric convolutional neural networks for evaluation of osteoarthritis. In: Proceedings - International Symposium on Biomedical Imaging. IEEE Computer Society; 2018. p. 851–4.

339. Panfilov E, Tiulpin A, Klein S, Nieminen MT, Saarakkala S. Improving Robustness of Deep Learning Based Knee MRI Segmentation: Mixup and Adversarial Domain Adaptation. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW) [Internet]. IEEE; 2019 [cited 2020 Mar 15]. p. 450–9. Available from: https://ieeexplore.ieee.org/document/9022164/

340. Desai AD, Gold GE, Hargreaves BA, Chaudhari AS. Technical Considerations for Semantic Segmentation in MRI using Convolutional Neural Networks. 2019 [cited 2020 Mar 15]; Available from: http://oai.epi-ucsf.org

341. Novikov AA, Lenis D, Major D, Hladůvka J, Wimmer M, Bühler K. Fully Convolutional Architectures for Multi-Class Segmentation in Chest Radiographs.

342. Wu S, Li H, Quang D, Guan Y. Three-Plane–assembled Deep Learning Segmentation of Gliomas. Radiol Artif Intell. 2020 Mar 1;2(2):e190011.

343. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-net: Learning dense volumetric segmentation from sparse annotation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag; 2016. p. 424–32.

344. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation.

345. Kamnitsas K, Bai W, Ferrante E, McDonagh S, Sinclair M, Pawlowski N, et al. Ensembles of Multiple Models and Architectures for Robust Brain Tumour Segmentation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) [Internet]. 2017 Nov 4 [cited 2021 Apr 6];10670 LNCS:450–62. Available from: http://arxiv.org/abs/1711.01468

346. Li H, Jiang G, Zhang J, Wang R, Wang Z, Zheng WS, et al. Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images. Neuroimage. 2018 Dec 1;183:650–65.

347. Welsch GH, Scheffler K, Mamisch TC, Hughes T, Millington S, Deimling M, et al. Rapid estimation of cartilage T2 based on double echo at steady state (DESS) with 3 Tesla. Magn Reson Med [Internet]. 2009 Aug 1 [cited 2021 Apr 6];62(2):544–9. Available from: http://doi.wiley.com/10.1002/mrm.22036

348. Sveinsson B, Chaudhari AS, Gold GE, Hargreaves BA. A simple analytic method for estimating T2 in the knee from DESS. Magn Reson Imaging [Internet]. 2017 May [cited 2019 Jul 11];38:63–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28017730

349. Michael Nevitt AC, Felson DT, Lester G. OAI Protocol THE OSTEOARTHRITIS INITIATIVE PROTOCOL FOR THE COHORT STUDY.

350. Paproki A, Engstrom C, Chandra SS, Neubert A, Fripp J, Crozier S. Automated segmentation and analysis of normal and osteoarthritic knee menisci from magnetic resonance images--data from the Osteoarthritis Initiative. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society. 2014 Sep 1;22(9):1259–70.

351. Tustison NJ, Avants BB, Cook PA, Yuanjie Zheng, Egan A, Yushkevich PA, et al. N4ITK: Improved N3 Bias Correction. IEEE Trans Med Imaging. 2010 Jun;29(6):1310–20.

352. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. 2014 Dec 22 [cited 2019 Aug 2]; Available from: http://arxiv.org/abs/1412.6980

353. Precision Imaging [Internet]. [cited 2024 Apr 18]. Available from: https://www.nottingham.ac.uk/research/beacons-of-excellence/precision-imaging/index.aspx

354. Tack A, Mukhopadhyay A, Zachow S. Knee menisci segmentation using convolutional neural networks: data from the Osteoarthritis Initiative. Osteoarthritis Cartilage. 2018 May;26(5):680–8.

355. Carballido-Gamio J, Blumenkrantz G, Lynch JA, Link TM, Majumdar S. Longitudinal analysis of MRI T2 knee cartilage laminar organization in a subset of patients from the osteoarthritis initiative. Magn Reson Med [Internet]. 2010 [cited 2021 Mar 11];63(2):465–72. Available from: https://pubmed.ncbi.nlm.nih.gov/19918905/

356. Schneider E, NessAiver M, White D, Purdy D, Martin L, Fanella L, et al. The osteoarthritis initiative (OAI) magnetic resonance imaging quality assurance methods and results. Osteoarthritis Cartilage. 2008 Sep 1;16(9):994–1004.

357. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage. 2008 Dec;16(12):1433–41.

358. Ahlbäck S. Osteoarthrosis of the knee. A radiographic investigation. Acta Radiol Diagn (Stockh). 1968;

359. Siorpaes K, Wenger A, Bloecker K, Wirth W, Hudelmaier M, Eckstein F. Interobserver reproducibility of quantitative meniscus analysis using coronal multiplanar DESS and IWTSE MR imaging. Magn Reson Med. 2012 May;67(5):1419–26.

360. Hunter DJ, Bowden JL. Therapy: Are you managing osteoarthritis appropriately? [Internet]. Vol. 13, Nature Reviews Rheumatology. Nature Publishing Group; 2017 [cited 2020 Sep 16]. p. 703–4. Available from: https://pubmed.ncbi.nlm.nih.gov/28931953/

361. Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al. CareTrack: Assessing the appropriateness of health care delivery in Australia [Internet]. Vol. 197, Medical Journal of Australia. 2012 [cited 2020 Sep 16]. p. 100–5. Available from: https://www.mja.com.au/journal/2012/197/2/caretrack-assessing-appropriateness-health-care-delivery-australia

362. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: The Chronic Osteoarthritis Management Initiative of the U.S. Bone and Joint Initiative. Semin Arthritis Rheum. 2014;43(6):701–12.

363. Manno RL, Bingham CO, Paternotte S, Gossec L, Halhol H, Giacovelli G, et al. OARSI-OMERACT initiative: Defining thresholds for symptomatic severity and structural changes in disease modifying osteoarthritis drug (DMOAD) clinical trials. Osteoarthritis Cartilage [Internet]. 2012 Feb [cited 2020 Sep 17];20(2):93–101. Available from: /pmc/articles/PMC4375956/?report=abstract

364. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage. 2011 Aug;19(8):990–1002.

365. Cai G, Cicuttini F, Aitken D, Laslett LL, Zhu Z, Winzenberg T, et al. Comparison of radiographic and MRI osteoarthritis definitions and their combination for prediction of tibial cartilage loss, knee symptoms and total knee replacement: a longitudinal study. Osteoarthritis Cartilage [Internet]. 2020 Aug 1 [cited 2020 Sep 17];28(8):1062–70. Available from: https://pubmed.ncbi.nlm.nih.gov/32413465/

366. Peake EJ, Chevasson R, Pszczolkowski S, Auer DP, Arthofer C. Ensemble learning for robust knee cartilage segmentation: data from the osteoarthritis initiative. bioRxiv [Internet]. 2020 Sep 1 [cited 2020 Sep 17];2020.09.01.267872. Available from: https://doi.org/10.1101/2020.09.01.267872

367. Yu D, Jordan KP, Snell KIE, Riley RD, Bedson J, Edwards JJ, et al. Development and validation of prediction models to estimate risk of primary total hip and knee replacements using data

from the UK: Two prospective open cohorts using the UK Clinical Practice Research Datalink. Ann Rheum Dis. 2019 Jan 1;78(1):91–9.

368. Kwoh CK, Guehring H, Aydemir A, Hannon MJ, Eckstein F, Hochberg MC. Predicting knee replacement in participants eligible for disease-modifying osteoarthritis drug treatment with structural endpoints. Osteoarthritis Cartilage [Internet]. 2020 Jun 1 [cited 2021 Mar 31];28(6):782–91. Available from: https://pubmed.ncbi.nlm.nih.gov/32247871/

369. Tolpadi AA, Lee JJ, Pedoia V, Majumdar S. Deep Learning Predicts Total Knee Replacement from Magnetic Resonance Images. Sci Rep [Internet]. 2020 Dec 1 [cited 2020 Sep 17];10(1):1–12. Available from: https://doi.org/10.1038/s41598-020-63395-9

370. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis [Internet]. Vol. 2, European Radiology Experimental. Springer; 2018 [cited 2020 Sep 17]. Available from: /pmc/articles/PMC6234198/?report=abstract

371. Yu D, Jordan KP, Snell KIE, Riley RD, Bedson J, Edwards JJ, et al. Development and validation of prediction models to estimate risk of primary total hip and knee replacements using data from the UK: Two prospective open cohorts using the UK Clinical Practice Research Datalink. Ann Rheum Dis. 2019 Jan 1;78(1):91–9.

372. OAI [Internet]. [cited 2020 Nov 30]. Available from: https://nda.nih.gov/oai/

373. Ashrafinia S, 0000-0002-9592-0437. QUANTITATIVE NUCLEAR MEDICINE IMAGING USING ADVANCED IMAGE RECONSTRUCTION AND RADIOMICS. 2019 Mar 26 [cited 2020 Sep 18]; Available from: http://dspace-prod.mse.jhu.edu:8080/handle/1774.2/61551

374. Greenwood M. A Report on the Natural Duration of Cancer. A Report on the Natural Duration of Cancer. 1926;(33).

375. Jia X, Baig MM, Mirza F, GholamHosseini H. A Cox-Based Risk Prediction Model for Early Detection of Cardiovascular Disease: Identification of Key Risk Factors for the Development of a 10-Year CVD Risk Prediction. Adv Prev Med [Internet]. 2019 Apr 9 [cited 2024 Mar 5];2019:1–11. Available from: /pmc/articles/PMC6481149/

376. DeLong E, DeLong D, Biometrics DCP, 1988 undefined. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. JSTOR.

377. Bacon K, Lavalley MP, Jafarzadeh SR, Felson D. Does cartilage loss cause pain in osteoarthritis and if so, how much? Ann Rheum Dis [Internet]. 2020 Aug 1 [cited 2021 Feb 25];79(8):1105–10. Available from: https://pubmed.ncbi.nlm.nih.gov/32381567/

378. Eckstein F, Guermazi A, Gold G, Duryea J, Hellio Le Graverand MP, Wirth W, et al. Imaging of cartilage and bone: Promises and pitfalls in clinical trials of osteoarthritis. Vol. 22, Osteoarthritis and Cartilage. W.B. Saunders Ltd; 2014. p. 1516–32.

379. Han Y, Sunwoo L, Ye JC. K-Space Deep Learning for Accelerated MRI. IEEE Trans Med Imaging. 2020 Feb 1;39(2):377–86.

380. Lluch E, Torres R, Nijs J, van Oosterwijck J. Evidence for central sensitization in patients with osteoarthritis pain: A systematic literature review. European Journal of Pain (United Kingdom) [Internet]. 2014 Nov 1 [cited 2022 Jun 26];18(10):1367–75. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/j.1532-2149.2014.499.x

381. Iannaccone F, Dixon S, physician AKP, 2017 undefined. A review of long-term pain relief after genicular nerve radiofrequency ablation in chronic knee osteoarthritis. painphysicianjournal.com [Internet]. [cited 2022 Jun 26]; Available from: https://www.painphysicianjournal.com/current/pdf?article=NDMxOQ%3D%3D&journal=104

382. Beard DJ, Davies LJ, Cook JA, MacLennan G, Price A, Kent S, et al. The clinical and cost-effectiveness of total versus partial knee replacement in patients with medial compartment osteoarthritis (TOPKAT): 5-year outcomes of a randomised controlled trial. The Lancet. 2019 Aug 31;394(10200):746–56.

383. Sun C, Zhang X, Lee WG, Tu Y, Li H, Cai X, et al. Infrapatellar fat pad resection or preservation during total knee arthroplasty: a meta-analysis of randomized controlled trials. J Orthop Surg Res [Internet]. 2020 Aug 5 [cited 2022 Jun 26];15(1):1–9. Available from: https://josr-online.biomedcentral.com/articles/10.1186/s13018-020-01823-2

384. Nairn LN, Subramaniam M, Ekhtiari S, Axelrod DE, Grant JA, Khan M. Safety and early results of Subchondroplasty® for the treatment of bone marrow lesions in osteoarthritis: a systematic review. Knee Surgery, Sports Traumatology, Arthroscopy [Internet]. 2021 Nov 1 [cited 2022 Jun 26];29(11):3599–607. Available from: https://link.springer.com/article/10.1007/s00167-020-06294-w

385. Jüni P, Hari R, Rutjes AWS, Fischer R, Silletta MG, Reichenbach S, et al. Intra-articular corticosteroid for knee osteoarthritis. Cochrane Database Syst Rev [Internet]. 2015 Oct 22 [cited 2021 Dec 18];2015(10). Available from: https://pubmed.ncbi.nlm.nih.gov/26490760/

386. Radojčić MR, Arden NK, Yang X, Strauss VY, Birrell F, Cooper C, et al. Pain trajectory defines knee osteoarthritis subgroups: a prospective observational study. Pain [Internet]. 2020 Dec 1 [cited 2021 Oct 27];161(12):2841. Available from: /pmc/articles/PMC7654950/

387. Physiotherapy IAAJ of, 2009 undefined. Western ontario and mcMaster universities osteoarthritis index (WOMAC). core.ac.uk [Internet]. 2009 [cited 2021 Nov 7]; Available from: https://core.ac.uk/download/pdf/82202672.pdf

388. Vogl TP, Mangis JK, Rigler AK, Zink WT, Alkon DL. Accelerating the convergence of the back-propagation method. Biol Cybern [Internet]. 1988 Sep 1 [cited 2022 Jan 10];59(4–5):257–63. Available from: https://link.springer.com/article/10.1007/BF00332914

389. Guan B, Liu F, Mizaian AH, Demehri S, Samsonov A, Guermazi A, et al. Deep learning approach to predict pain progression in knee osteoarthritis. Skeletal Radiol [Internet]. 2022 Feb 1 [cited 2022 Jun 30];51(2):363–73. Available from: https://link.springer.com/article/10.1007/s00256-021-03773-0

390. Chang GH, Felson DT, Qiu S, Guermazi A, Capellini TD, Kolachalama VB. Assessment of knee pain from MR imaging using a convolutional Siamese network. Eur Radiol [Internet]. 2020 Jun 1 [cited 2022 Jun 30];30(6):3538. Available from: /pmc/articles/PMC7786238/

391. Zhong H, Miller DJ, Urish KL. T2 map signal variation predicts symptomatic osteoarthritis progression: data from the Osteoarthritis Initiative. Skeletal Radiol. 2016 Jul 1;45(7):909–13.

392. Urish KL, Keffalas MG, Durkin JR, Miller DJ, Chu CR, Mosher TJ. T2 Texture Index of Cartilage Can Predict Early Symptomatic OA Progression: Data from the Osteoarthritis Initiative. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society [Internet]. 2013 Oct [cited 2022 Jun 30];21(10):1550. Available from: /pmc/articles/PMC3779506/

393. Pollard TCB, Gwilym SE, Carr AJ. The assessment of early osteoarthritis. Journal of Bone and Joint Surgery - Series B [Internet]. 2008 Apr 1 [cited 2022 Jun 7];90(4):411–21. Available from: https://online.boneandjoint.org.uk/doi/abs/10.1302/0301-620X.90B4.20284

394. Hunter DJ, Zhang Y, Niu J, Goggins J, Amin S, LaValley MP, et al. Increase in bone marrow lesions associated with cartilage loss: a longitudinal magnetic resonance imaging study of knee osteoarthritis. Arthritis Rheum [Internet]. 2006 May [cited 2022 Jun 7];54(5):1529–35. Available from: https://pubmed.ncbi.nlm.nih.gov/16646037/

395. Zhai G, Blizzard L, Srikanth V, Ding C, Cooley H, Cicuttini F, et al. Correlates of knee pain in older adults: Tasmanian Older Adult Cohort Study. Arthritis Rheum [Internet]. 2006 Apr 15 [cited 2022 Jun 3];55(2):264–71. Available from: https://pubmed.ncbi.nlm.nih.gov/16583417/

396. Dore D, Quinn S, Ding C, Winzenberg T, Zhai G, Cicuttini F, et al. Natural history and clinical significance of MRI-detected bone marrow lesions at the knee: a prospective study in community dwelling older adults. Arthritis Res Ther [Internet]. 2010 Dec 29 [cited 2022 Jun 3];12(6). Available from: https://pubmed-ncbi-nlm-nih-gov.nottingham.idm.oclc.org/21190555/

397. Zhang Y, Nevitt M, Niu J, Lewis C, Torner J, Guermazi A, et al. Fluctuation of Knee Pain and Changes in Bone Marrow Lesions, Effusions and Synovitis on MRI: The MOST Study. Arthritis Rheum [Internet]. 2011 Mar [cited 2022 Jun 3];63(3):691–9. Available from: /pmc/articles/PMC3056156/

398. Jones G. Sources of pain in osteoarthritis: Implications for therapy. Int J Clin Rheumtol. 2013 Jun;8(3):335–45.

399. Roemer FW, Guermazi A, Collins JE, Losina E, Nevitt MC, Lynch JA, et al. Semi-quantitative MRI biomarkers of knee osteoarthritis progression in the FNIH biomarkers consortium cohort - Methodologic aspects and definition of change. BMC Musculoskelet Disord [Internet]. 2016 [cited 2019 Feb 26];17(1):466. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27832771

400. Greve D, Neuroimage BF, 2009 undefined. Accurate and robust brain image alignment using boundary-based registration. Elsevier [Internet]. [cited 2022 Jun 27]; Available from: https://www.sciencedirect.com/science/article/pii/S1053811909006752?casa_token=TorQWMoO-K8AAAAA:5AZGCILeORYLRwzcFosKZCiRNkQApEnxENNzyVcZySnKtfxfMYid50sQtc5HUFy1JFXrxpQ

401. Jenkinson M, Bannister P, Brady M, Neuroimage SS, 2002 undefined. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Elsevier [Internet]. 2002 [cited 2022 Jun 27]; Available from: https://www.sciencedirect.com/science/article/pii/S1053811902911328

402. Jenkinson M, analysis SSM image, 2001 undefined. A global optimisation method for robust affine registration of brain images. Elsevier [Internet]. [cited 2022 Jun 27]; Available from: https://www.sciencedirect.com/science/article/pii/S1361841501000366?casa_token=mij-yr-aHdAAAAAA:7D0XlWLILXQ98-QfC_rT-AaHjjHmoD-mCIx-fpBonhUR_qhDoe3yDwqpSpuj1opcMKGLJYo

403. Perrot S. Osteoarthritis pain. Best Pract Res Clin Rheumatol. 2015 Feb 1;29(1):90–7.

404. Culvenor AG, Øiestad BE, Hart HF, Stefanik JJ, Guermazi A, Crossley KM. Prevalence of knee osteoarthritis features on magnetic resonance imaging in asymptomatic uninjured adults: a systematic review and meta-analysis. Br J Sports Med [Internet]. 2019 Oct 1 [cited 2022 Jun 30];53(20):1268. Available from: /pmc/articles/PMC6837253/