

Artificial Intelligence in Musculoskeletal Medicine: A Scoping Review

Samantha Cooray,¹ Alexander Deng,² Tim Dong,³ Salah Al Ghazal Hammouche,^{4,5} Matthew Quansah,⁶ Jordan Tsigarides⁷ and Nicholas Fuggle^{6,8}

1. Infection, Immunity and Inflammation Department, University College London Great Ormond Street Institute of Child Health, London, UK; 2. Guy's and St Thomas's NHS Foundation Trust, London, UK; 3. Department of Cardiac Surgery, Bristol Heart Institute, University of Bristol, Bristol, UK; 4. Imperial College London, London, UK; 5. London Postgraduate School of Surgery, London, UK; 6. MRC Lifecourse Epidemiology Centre, University of Southampton, Southampton, UK; 7. Norwich Medical School, University of East Anglia, Norwich, UK; 8. The Alan Turing Institute, London, UK

In recent years, artificial intelligence (AI) has blossomed via the booming increase in computing power, access to large data sets and the development of novel data science techniques and architectures. This has led to a great interest in the clinical deployment of AI in healthcare. In this review, we look particularly at the field of musculoskeletal medicine and have curated the latest progress in AI and machine learning in the areas of inflammatory arthritis, vasculitis and connective tissue diseases, osteoporosis, Ehlers–Danlos syndrome, osteoarthritis and musculoskeletal surgical interventions. We also highlight the challenges and pitfalls as clinical AI moves forward.

Keywords

Arthritis, artificial intelligence, inflammatory, machine learning, musculoskeletal, osteoporosis, rheumatology

Disclosures: Samantha Cooray, Alexander Deng, Tim Dong, Salah Al Ghazal Hammouche, Matthew Quansah and Jordan Tsigarides have no potential conflicts to declare. Nicholas Fuggle has received speaker fees from Viatrix and travel bursaries from Pfizer and Eli Lilly.

Review Process: Double-blind peer review.

Compliance with ethics: This article involves a review of the literature and did not involve any studies with human or animal subjects performed by any of the authors.

Data availability: Data sharing is not applicable to this article as no datasets were generated or analysed during the writing of this article.

Digital features: This article is published with supplementary digital features. The supplementary material was peer reviewed with the manuscript. To view the material please visit here.

Authorship: All named authors meet the criteria of the International Committee of Medical Journal Editors for authorship for this manuscript, take responsibility for the integrity of the work as a whole and have given final approval for the version to be published.

Access: This article is freely accessible at touchIMMUNOLOGY.com © Touch Medical Media 2024.

Received: 15 December 2023

Accepted: 16 April 2024

Published online: 5 July 2024

Citation: *touchREVIEWS in RMD*. 2024;3(2):Online ahead of journal publication

Corresponding author: Dr Nicholas Fuggle, MRC Lifecourse Epidemiology Centre, University of Southampton, Tremona Road, Southampton SO16 6YD, UK. E: nrf@mrc.soton.ac.uk

Support: No funding was received in the publication of this article.

There is much excitement about the deployment of artificial intelligence (AI) in healthcare, and the musculoskeletal field is no exception. In this article, we introduce some of the latest developments relating to osteoarthritis (OA), osteoporosis, rheumatoid arthritis (RA) (as an example of inflammatory arthritis), connective tissue disease (CTD), Ehlers–Danlos syndrome (EDS) and musculoskeletal surgical interventions.

Artificial intelligence

AI is the ability of machines to imitate human behaviour. Machine learning (ML) is a branch of AI that uses computer algorithms to automatically learn from input data and through the identification of patterns to make accurate predictions with minimal human intervention. ML algorithms do not rely on a predetermined equation as a model but adaptively improve their performance as more data are provided.

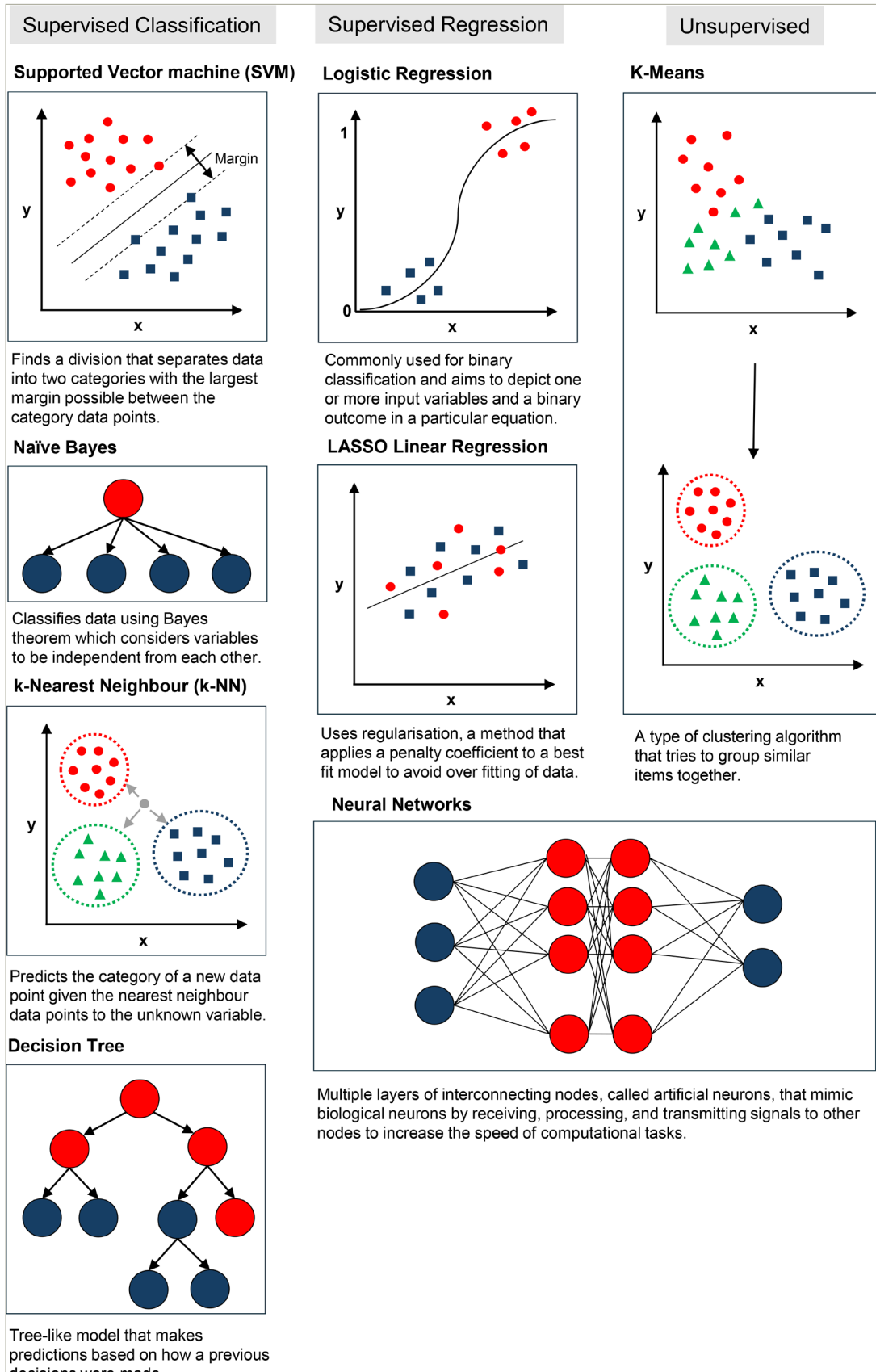
ML is increasingly being used in medical applications to independently and efficiently carry out specified tasks, such as image processing in radiology and histopathology and the identification of disease markers, and in the field of drug discovery.

The process by which ML algorithms learn is called 'training', and the training methods are classified as supervised, unsupervised or reinforcement learning. Supervised learning involves training algorithms to predict future values through learning patterns from labelled input data, for example, accurately predicting the opacification on a chest X-ray through training with images that are labelled as normal or abnormal.

Supervised learning algorithms use classification, regression and forecasting approaches. In classification models, the algorithm must draw a conclusion from the observed values and determine to which category new observations belong. In regression tasks, the ML algorithm must estimate and understand the relationships among variables. Forecasting is the process of making predictions about the future based on past and present data and is commonly used to analyse trends.

In unsupervised learning, the input data are not assigned or labelled. Instead, the algorithm infers the underlying patterns and relationships within the input data, which is useful, for example, in finding patterns in blood biomarkers that are associated with a particular disease state. Clustering is a type of unsupervised learning that involves grouping sets of similar data based on the defined criteria. It is useful for segmenting the data into groups and performing analysis on each data set to find patterns. Dimension reduction is another type of unsupervised learning that gradually reduces the number of variables being considered to find the exact information required. In reinforcement learning, the algorithm achieves a specified goal by interacting with an environment through trial and error. This type of method is commonly applied in the field of robotic surgery.

Figure 1: Schematic representations of machine learning and deep learning methods



LASSO = least absolute shrinkage and selection operator.

Examples of ML algorithms are shown in *Figure 1* and described in greater depth in the glossary (online supplementary file 1).

Inflammatory arthritis

Inflammatory arthritides are characterized by joint inflammation, leading to clinical signs, such as pain, swelling, heat and erythema. This collection of rheumatological conditions encompasses those causing inflammation due to various aetiologies, including autoimmunity, crystals and reactive processes. In this review, the focus will be on RA as a vital example of inflammatory arthritis.

RA is a chronic, systemic autoimmune disease leading to inflammatory arthritis, typically involving the small joints that can result in joint destruction and severe functional impairment.¹ RA is a heterogeneous disease with a wide variation in the age of onset, degree of joint involvement and severity. The diagnosis is based on clinical criteria, and there is no definitive diagnostic test available to differentiate RA from other inflammatory joint diseases.² The aim of treatment for RA is to reduce inflammation and prevent joint destruction, and it includes corticosteroids, conventional disease-modifying anti-rheumatic drugs (DMARDs) and an ever-expanding list of targeted biological therapies.³ Patients have variable response and remission rates to different treatment strategies, and identifying responders from non-responders is difficult. However, early diagnosis and prompt initiation of treatment are important in the prevention of inflammatory joint destruction. AI and ML are increasingly being used to address the unmet needs in the diagnosis, monitoring and treatment of RA, and in this section, we review the latest developments.⁴

Machine learning in inflammatory arthritis diagnosis

AI and ML for diagnostic purposes use the data from electronic health records (EHRs), diagnostic biomarkers and imaging modalities to identify patients with RA at an early stage of the disease.

Electronic health records

The phenotypic identification of RA using data from EHRs was first carried out in 2010 using a support vector machine (SVM) model, which accurately predicted RA disease activity from the data containing specific terms (area under the curve [AUC] >0.90).⁵ The comparison of different ML methods using a 10-fold cross-validation from thousands of EHRs showed that an SVM model performed best in identifying patients with RA in <7 s (an AUC of 0.98 and a positive predictive value [PPV] of 0.94).⁶ The model compared favourably with manual methods.⁷ A random forest method and a decision tree model identified the best predictors of RA, such as local coding for RA; clinical features, such as nodules, seropositivity and arthropathy and drugs, such as prednisolone, methotrexate, sulfasalazine and leflunomide, from EHRs, with a PPV of 85.6%, specificity of 94.6% and sensitivity of 86.2%.⁸

Molecular markers

Random forest, logistic regression and other statistical models have differentiated patients with RA from controls using gene expression signatures, with an AUC of 0.96.^{9,10} Artificial neural networks (ANNs) have successfully distinguished RA from OA and psoriatic arthritis (PsA, another form of inflammatory arthritis) using serum protein profiles identified through mass spectrometry and immunoassays, although the AUC varies significantly.^{11–13}

Imaging techniques and histopathology

Supervised and unsupervised algorithms have been used, together with imaging techniques, for RA diagnosis. ANN models trained and tested on hand radiographs accurately classified RA with an AUC of 0.97, a sensitivity of 90.7% and a specificity of 92.6%.^{14,15} Larger data sets have allowed the effective differentiation of RA and OA.¹⁶ ANNs also identified synovial proliferation in joint ultrasound with an AUC of 0.863–0.886, depending on the joint area sampled.^{17,18} A systemic review and meta-analysis of computer-assisted magnetic resonance imaging (MRI) showed a good correlation with manual methods in RA diagnosis.¹⁹ Thermal imaging of diurnal joint temperatures has been studied, with quantum SVM models used to improve the accuracy.^{20,21} Scoring algorithms based on features, such as inflammatory infiltrates and hyperplasia in stained synovial tissue, have demonstrated considerable efficacy in differentiating RA, OA and PsA from normal tissue (AUC 0.87–0.98).^{22,23}

Disease progression and disease activity

Once RA is diagnosed, an ongoing assessment is vital to monitor the disease activity. An ordinal regression-based ML model developed to estimate the Clinical Disease Activity Index (CDAI) scores was trained and validated on 11,985 patient notes with CDAI scores from the OM1 RA registry (OM1® real-world data and AI for research [OM1, Boston, MA, USA]).²⁴ The model had a PPV of 0.80, a negative predictive value of 0.84 and an AUC of 0.88 when evaluating the performance using the binarized (negative CDAI score ≤10 or positive CDAI score ≥10.1) version of the outcome.

Predicting treatment response

ML techniques have shown promise in the evaluation of treatment response to DMARDs and biologics. Least absolute shrinkage and selection operator and random forest methods showed that patients with RA with a low baseline Disease Activity Score in 28 joints (DAS28)-erythrocyte sedimentation rate (ESR) score, positive anti-citrullinated protein antibody and the Health Assessment Questionnaire ≤2 were most likely to respond well to methotrexate (AUC of 0.68 and 0.79, respectively).²⁵ Conversely, a DAS28-ESR score (>3.2) at 3 months post-methotrexate treatment predicted a poor therapeutic response with different methods.²⁶ Elastic net models demonstrated a 3.6-fold better response to anti-tumour necrosis factor (TNF) biologics in males, and using genetic data, carriers of the *resistin* gene rs3219177 polymorphism had a sixfold better response than non-carriers. In a subsequent study, carriers of the *TLR9* gene rs352139 polymorphism were found to have a fivefold better response to anti-TNF than non-carriers.^{27,28} ML using biomarker signatures in blood or synovial tissue has reliably predicted responses to other biologics (e.g. rituximab, tocilizumab and sarilumab).^{29–31}

Vasculitis and connective tissue disease

Vasculitis is a complex group of rare rheumatic diseases characterized by inflammation of blood vessels that present considerable diagnostic and therapeutic challenges. Recent studies using AI techniques have attempted to support earlier diagnosis, treatment optimization and prognosis prediction.^{32–34}

The diagnosis of vasculitis relies upon the recognition of characteristic clinical features often alongside the use of imaging or biopsy. ML techniques show great promise for supporting this process, with studies using natural language processing to differentiate vasculitis from clinical notes, deep learning algorithms (such as *U-Net*) to detect features

on medical imaging and supervised learning approaches to identify predictive biomarkers detectable from blood samples.³⁵⁻⁴¹

Treatment typically includes induction through high-dose corticosteroids followed by maintenance with 'steroid-sparing' agents. However, variability in treatment response combined with common side effects makes management fraught with difficulty. ML has been used to address treatment challenges through supervised algorithms (including *random forest* and *light gradient boosting models*) to predict relapse after glucocorticoid tapering in giant cell arteritis and the development of models predicting resistance to intravenous immunoglobulin therapy in Kawasaki's disease.^{32,42,43}

The risk of morbidity following a diagnosis of vasculitis is significant, and understanding the prognostic risk factors is essential to tailor management strategies. Emerging ML-based tools show promise, especially in predicting eye diseases from clinical data (using ensemble algorithms, such as *XGBoost*) or angiography imaging and risk of renal damage in immunoglobulin A (IgA) vasculitis based on patients' clinical features.⁴⁴⁻⁴⁶

With only 26 primary research studies published over the last 20 years using AI methods in vasculitis, and the majority investigating Kawasaki's disease (n=12), there is still a dearth of available evidence despite many studies showing highly promising results.⁴⁷⁻⁵⁷

Autoimmune CTDs are systemic in nature and characterized by a variety of phenotypes. ML techniques have begun to revolutionize the study of these varied diseases, providing tools for improving diagnostic approaches, the classification of disease subtypes and predicting factors associated with treatment response and the development of disease-related complications.

CTDs present a diagnostic challenge for clinicians, and many studies using ML focus on demonstrating accurate tools to support the diagnostic process. There are encouraging studies investigating gene expression data as novel biomarkers in CTDs, but perhaps, the approaches closer to clinical implementation are those using deep learning to support and automate the interpretation of clinical data (including point-of-care tests, radiological imaging and biopsies).⁵⁸⁻⁷¹ An interesting example includes the use of deep learning to differentiate systemic sclerosis from primary Raynaud's syndrome using a low-cost digital nailfold capillaroscope, an underused tool in which AI may support more widespread use.⁶³

The spectrum of CTD presentations highlights the importance of classification, subtype definition and tailored treatment strategies. Data mining and supervised approaches have been used to discriminate subtypes of Sjögren's syndrome and lupus based on clinical data, in addition to multiple studies investigating the classification of myositis using electromyography and MRI data.⁷²⁻⁷⁶

Many of the AI-based studies on CTDs aim to predict outcomes and treatment responses to inform a personalized management approach. Multiple ML approaches have been used to predict complications and comorbidities in lupus (including nephritis), pulmonary involvement in systemic sclerosis and response to immunomodulatory treatments.⁷⁷⁻⁹⁷

Osteoporosis

Osteoporosis is characterized by reduced bone mineral density (BMD), leading to a predisposition to a fragility fracture. The diagnosis is complicated by the absence of symptoms or clinical signs and is reliant

on measures of BMD. BMD is primarily derived from dual-energy X-ray absorptiometry (DXA) as a single parameter, and it is this measurement that informs the densitometric diagnosis of osteoporosis, defined as a BMD *T*-score of ≤ -2.5 . AI techniques have been used to interrogate 'omic' data sets to identify biomarkers and therapeutic targets in osteoporosis, as they have in most major disease areas; however, the focus of this review will be on the deployment of AI on skeletal imaging (computer vision) in the field of osteoporosis.^{98,99}

The major aims of computer vision in osteoporosis are as follows:

- use DXA images to improve fracture prediction beyond BMD
- opportunistically measure BMD from routinely performed imaging
- opportunistically identify vertebral fractures from routinely performed imaging
- derive bone microarchitecture features.

It is worth noting that, in the field of osteoporosis, the outcome of interest can be either the diagnosis of osteoporosis or the assessment of fracture risk. Although the former was vital for defining the disease and developing the field, the measures of the latter (fracture risk) are used to adjudicate treatment thresholds via fracture risk-prediction tools, including The Fracture Risk Calculator (FRAX[®]), which provides a 10-year estimation of major osteoporotic risk and hip fracture risk.¹⁰⁰ This is, in part at least, as over half of fragility fractures occur at a normal BMD and so a pure focus on the diagnosis of osteoporosis would miss a large proportion of those at risk of fracture. Indeed, fractures are associated with high morbidity and mortality, surpassing simply the densitometric diagnosis of osteoporosis.

As mentioned earlier, the DXA scan is performed to provide a measure of BMD (via levels of beam attenuation), but the rest of the DXA image (of the hips and anterior-posterior spine) is neglected unless a lateral image of the spine is acquired as a vertebral fracture assessment. These acquired, but underused, DXA images are ripe for computer vision analyses.

Indeed, a meta-analysis of seven studies that attempt to improve the diagnosis of osteoporosis via hip DXA images found a pooled sensitivity of 0.844 (95% confidence interval [CI], 0.79-0.89), a pooled specificity of 0.781 (95% CI, 0.73-0.82) and a summary of receiver operating characteristic AUC of 0.878.¹⁰¹ Although these results are encouraging, DXA imaging is limited by quality due to artefacts of soft tissue and degenerative sclerosis and, in a clinical setting, is only performed to assess BMD. Plain radiographs and computed tomography (CT) are imaging modalities that are used for a plethora of clinical indications and often contain skeletal elements. These images could all contribute to the assessment of skeletal health opportunistically.

This has led to studies showing that abdominal CT can predict lumbar spine BMD (AUC 0.96-0.97), chest CT scans (performed for lung cancer screening) can discriminate osteoporosis from normal participants (AUC 0.97) and lateral spine X-rays can identify osteoporosis, though with a lower degree of accuracy (AUC 0.85).^{102,103}

Vertebral fractures are often missed on routine imaging, and hence, models have been developed using deep learning (amidst other methods) to identify vertebral fractures from lateral spine radiographs (AUC 0.93) and sagittal reconstructions of CT images.^{104,105}

Bone is composed of two compartments, cortical and trabecular, and information regarding these compartments can provide more detailed 'bone microarchitectural' measures of bone strength and resistance to fracture than simply BMD alone. This microarchitecture can be derived from specialist scans currently largely used in research settings, including high-resolution peripheral quantitative CT (HR-pQCT). Computer vision analyses of HR-pQCT have demonstrated excellent discrimination for prevalent fractures with an AUC of 0.85–0.97 using a gradient boosting machine and statistical shape mapping.^{106–108} What is more, there are efforts, using AI, to derive some microarchitectural features (though only from the trabecular compartment) from the DXA imaging.¹⁰⁹

In conclusion, computer vision analyses have improved the identification of fracture risk on DXA images, developed the opportunistic screening of BMD and vertebral fractures and shown promise for deriving bone microarchitectural properties from clinically applicable imaging modalities.

Ehlers–Danlos syndrome

EDSs are a group of inherited disorders of connective tissue. Within this group, there are 14 recognized EDS subtypes, 13 of which are genetically characterized. The clinical presentation commonly includes features such as joint hypermobility, skin hyperextensibility, abnormal wound healing and widespread pain. Similar to connective tissue disorders, EDSs are not solely disorders of the musculoskeletal system but are complex multi-system disorders due to the role of connective tissue throughout the body. AI tools in development may eventually assist in the diagnosis of EDS, with two current examples being video assessment of joint hypermobility and the classification of rare diseases via analysis of patient-created pain drawings.

A thematic clinical feature of EDS is generalized joint hypermobility. The clinical finding of joint hypermobility is a core diagnostic feature measured through the Beighton examination completed using a goniometer. However, the use of a goniometer is uncommon in actual clinical practice, and evidence of inter-rater reliability in the Beighton score reporting is limited or conflicting.¹¹⁰ Mittal et al. have devised a method to apply computer vision to smartphone camera footage of patients, with the goal of creating a more objective, validated and scalable method for assessing hypermobility.¹¹¹ This provides an excellent case study to highlight the design of research in the field of musculoskeletal AI.

The footage will be analysed using ML algorithms from the existing open-source human pose-estimation models, which use the video input to create an output of the location and angles of joints in the body. Pose-estimation models lack training data on generalized joint hypermobility, potentially creating *out-of-distribution errors* for the pose libraries. The authors plan to address this by additional fine-tuning of the models with manually annotated examples of hyperextended joints. The method is planned to be validated compared with the ground truth of expert clinicians using the goniometer-validated assessment.

The outcomes of the pre-registered trial are awaited in 2024 (Assessing the Feasibility of a Smartphone-based, Machine Learning Visual Imaging Application for Assessment of Hyperextensibility of Peripheral Joints in Ehlers Danlos Syndrome; ClinicalTrials.gov identifier: NCT05366114), and the authors anticipate possible use cases as a triage tool for specialized EDS clinic referrals.¹¹²

Chronic pain is another thematic clinical feature of EDS. There are many aetiologies of chronic pain, including other rare diseases, and

management is informed by the underlying cause. Emmert et al. developed an algorithm trained to distinguish between patterns of chronic pain caused by different rare diseases.¹¹³ The input data to the algorithm were pain drawings. Pain drawings are made by patients who mark where they experience pain on a 2D outline of a human body to create a visual representation of their pain (instead of relying on language). The authors used an open-source platform 'Pain2D' to collect the data from patients with specific known rare disease diagnoses involving chronic pain: EDS, Guillain–Barré syndrome, facioscapulohumeral muscular dystrophy, proximal myotonic myopathy and a control group with non-specific chronic pain. An overall pain profile for each disease category was created using this platform by calculating the frequency with which each pixel in the pain drawing was marked over all instances of each diagnostic category. Pain profiles summarizing different diagnostic categories were then compared with the known cases (leave-one-out cross-validation) by a mathematical coefficient of similarity – the highest-scoring pain profile was chosen as the output classification. This method classified EDS among the five diagnostic categories with a sensitivity of 64.4% and a specificity of 88.7%. Significant limitations include the absence of pain-drawing data from other diagnostic categories of chronic pain and small data sets within the existing categories that precluded the use of deep learning by convolutional neural networks (CNNs).

It is anticipated that these study areas will be developed and built upon but provide good examples of AI study designs that have the potential to be used across musculoskeletal disorders.

Osteoarthritis

OA is a common joint disease affecting millions of people throughout the world but is plagued by a disconnect between clinical and radiographic diseases. In recent years, ML and AI techniques have shown promise for enhancing OA diagnosis, care and management.¹¹⁴ This section seeks to give readers an overview of the most recent developments in ML and AI applications in the field of OA.

Early diagnosis and disease progression detection

Early diagnosis and detection algorithms use a variety of data sources, including patient-reported outcomes and medical imaging (X-rays and MRI scans), to detect the patterns and signals linked to the disease. This can take different approaches, but one study combined clinical, imaging and demographic features using a *multi-modal feature integration method* using an L-1 normalization approach to minimize the number of irrelevant features from each modality.¹¹⁵ Early OA symptoms can be recognized so that therapies can start sooner, possibly slowing the disease's progression.

To develop predictive models for OA progression, ML algorithms have been trained on big data sets combining patient demographics, clinical data and imaging results. These models determine the chance that an illness will progress and assist physicians in selecting the best course of action, potentially improving patient outcomes via a clinical decision support system. One research presented a CNN model referred to as the *YOLOv3* tiny-based object identification approach to automatically detect and categorize knee OA according to the Kellgren–Lawrence (KL) classification scheme.¹¹⁶ The strategy used two models, one for classifying severity (progression) and the other for distinguishing between healthy and osteoarthritic knees. The CNN model sought to assist radiologists and orthopaedic surgeons by correctly identifying and classifying knee OA across various stages. Another research introduced *DenseNet169* as a model for the effective diagnosis of knee OA.¹¹⁷ It compared *DenseNet169*'s performance with five other deep learning systems and

highlighted its efficacy in finding knee OA and determining its severity through multi-classification (accuracy of 95%, sensitivity of 88.77%, specificity of 95.41% and precision of 87.08%) and binary classifications (accuracy of 93.78%, sensitivity of 91.29%, specificity of 87.57% and accuracy of 89.27%). The suggested methodology showed potential in relieving radiologists' burden, enabling early diagnosis and saving time and expense related to knee OA diagnosis. It used a weighted training approach and incorporated gradient-weighted class activation mapping, a heatmap-based algorithm that deep learning models used to visualize and decipher the significant areas of an image (using gradients in the final convolutional layer) that influence the model's prediction.

Treatment plans for people with OA have been improved using ML and AI techniques. Algorithms can recommend personalized treatment programmes that are suited to each patient's needs by examining patient data, including medical history, genetic factors and lifestyle variables. By identifying the most efficient interventions and minimizing the process of trial and error while choosing treatments, this strategy seeks to improve patient outcomes.

One study found that AI-aided clinician diagnostic ratings have a stronger correlation with the total KL scores and Knee Injury and Osteoarthritis Outcome Scores than those by unaided physicians, increasing the consistency and precision of OA diagnosis.¹¹⁸ The KL score appears to be a limited tool on its own, with only a modest relationship between clinical severity and radiographic assessments. AI assistance has the potential to address the disconnect between radiological OA severity and clinical symptoms, resulting in more precise detection and better patient care.

The use of wearable and sensor technologies in conjunction with ML algorithms enables remote monitoring of patients with OA. These devices can monitor joint motion, levels of activity and other important factors. Clinicians can provide real-time feedback and useful insights into patient situations by analysing the acquired data, enabling proactive management and intervention. For monitoring human movement, wearable sensors provide an affordable and practical alternative to optical motion capture.¹¹⁹ Large amounts of real-world data can be provided by them, enabling remote monitoring, resolving time restrictions in the delivery of remote treatment and encouraging patient interaction. For movement analysis, accelerometers, gyroscopes and magnetometers are frequently combined into *inertial measurement units* for cost-efficiency and usability in large cohorts. For people with hip or knee OA, including those who have had joint replacement surgery, wearable sensors hold great promise for improving research and care.

Another study validated a remote patient monitoring system for 25 patients with OA undergoing total knee arthroplasty using a wearable knee sleeve and smartphone applications.¹²⁰ The sleeve accurately measured mobility, range of motion, patient-reported outcome measures, opioid compliance and home exercise compliance. Such studies demonstrate the potential of AI-augmented remote patient monitoring for evaluating hip and knee arthroplasty recovery and rehabilitation.

In OA management, persistent challenges remain around the quality and accessibility of data. Obtaining comprehensive data sets spanning diverse patient profiles, multiple data modalities and longitudinal records remains a hurdle, essential for training reliable ML models. Another critical challenge lies in the interpretability and explainability of these models; understanding the rationale behind their diagnoses or treatment suggestions is crucial for building trust among healthcare professionals and patients. Additionally, ensuring the ethical use of sensitive medical

information while maintaining patient privacy poses an ongoing concern. Moreover, integrating these advanced technologies safely into clinical workflows and healthcare systems presents a significant challenge for wider adoption and practical implementation in real-world settings.

Musculoskeletal surgical interventions

The sphere of surgical intervention has notably benefitted from the introduction of AI-aided technologies, which present the potential to amplify several aspects of surgical care.¹²¹⁻¹²⁸ The purview of these tools encompasses diagnosis and planning to intraoperative guidance.¹²¹⁻¹²⁸ A salient advantage rests in their capacity to aid clinicians in the early detection of musculoskeletal disorders, such as OA and soft tissue pathologies.^{123,127,129,130} Furthermore, the competency of AI in the precise identification of fractures and distinguishing between diverse types of spondylitis parallels that of experienced radiologists potentially enabling faster management.¹³¹⁻¹³³

In the planning phase, AI resources are able to deliver patient-specific preoperative plans, thereby minimizing the necessity for surgeon-delivered modifications to the surgical plan.^{121,123,134} These resources can also construct 3D models of patient anatomy, providing orthopaedic surgeons with a holistic and detailed visual representation of the surgical area.^{124,125} In the operating theatre, AI's role expands to offering real-time feedback and assistance in navigating complex anatomical structures.^{124,126,135,136}

Postoperative care has similarly been affected by the advent of AI, aiding in the identification of implant positioning and type, forecasting potential implant failure and assisting in planning for future surgical interventions.^{122,137,138} AI-assisted tools have manifested as an efficacious solution to the common difficulties in postoperative CT and MRI associated with metal-related image artefact and distortion.¹³⁹ AI can also help in determining follow-up recommendation plans based on analysing the radiological report.¹⁴⁰

Challenges

The field of musculoskeletal AI is filled with optimism, but there are substantial challenges that must be addressed to enable safe, unbiased and equitable deployment in clinical practice.¹⁴¹⁻¹⁴³

Most current ML methods use supervised methods using large data sets, which need manual checks and large amounts of data. This means that the accuracy of the models relies on the quality of the input data. Quality control measures are, therefore, required, and these need time-consuming human input.

The integration of AI algorithms into the clinical workflow must be performed carefully and thoughtfully and be congruent with the existing information technology infrastructure.¹⁴¹⁻¹⁴³ The replication of AI studies on 'real-life' data remains problematic due to the sporadic availability of training data and code and must be addressed moving forward and is sometimes neglected in favour of model performance.¹⁴²

The data for future AI research must be suitably anonymized with consideration of patient privacy and ethical considerations.¹⁴² The patients' preference for human experts' diagnoses accentuates the importance of cultivating appropriate levels of trust in AI applications. Additionally, considerable challenges persist in data availability and model validation, especially for rare musculoskeletal disorders and diseases (including vasculitis and CTD). The regulation in AI is developing and is vital to ensure that these tools are created, maintained and used appropriately in clinical practice.^{142,144,145}

Conclusions

AI within the musculoskeletal field harbours immense benefits, and as detailed earlier, a great deal of work has been performed and is ongoing. As research and development progress, AI has the potential to become an increasingly vital component of musculoskeletal care. However, the

above challenges must be considered of utmost importance as the field moves forward.¹⁴⁶ While AI embodies the promise of enhanced patient care, cognizance of the related challenges and ethical considerations is indispensable for wider acceptance and implementation of AI in the musculoskeletal field. □

- Momtazmanesh S, Nowroozi A, Rezaei N. Artificial intelligence in rheumatoid arthritis: Current status and future perspectives: A state-of-the-art review. *Rheumatol Ther*. 2022;9:1249–304. DOI: 10.1007/s40744-022-00475-4.
- Gravallese EM, Firestein GS. Rheumatoid arthritis – Common origins, divergent mechanisms. *N Engl J Med*. 2023;388:529–42. DOI: 10.1056/NEJMr2103726.
- Prasad P, Verma SS, Rathi A. Rheumatoid arthritis: Advances in treatment strategies. *Mol Cell Biochem*. 2023;478:69–88. DOI: 10.1007/s11010-022-04492-3.
- Kedra J, Davergne T, Braithwaite B, et al. Machine learning approaches to improve disease management of patients with rheumatoid arthritis: Review and future directions. *Expert Rev Clin Immunol*. 2021;17:1311–21. DOI: 10.1080/1744666X.2022.2017773.
- Carroll RJ, Eyley AE, Denny JC. Naive electronic health record phenotype identification for rheumatoid arthritis. *AMIA Annu Symp Proc*. 2011;2011:189–96.
- Maarseveen TD, Meinderink T, Reinders MJT, et al. Machine learning electronic health record identification of patients with rheumatoid arthritis: Algorithm pipeline development and validation study. *JMIR Med Inform*. 2020;8:e23930. DOI: 10.2196/23930.
- Maarseveen TD, Maurits MP, Niemantsverdriet E, et al. Handwork vs machine: A comparison of rheumatoid arthritis patient populations as identified from EHR free-text by diagnosis extraction through machine-learning or traditional criteria-based chart review. *Arthritis Res Ther*. 2021;23:174. DOI: 10.1186/s13075-021-02553-4.
- Zhou S-M, Fernandez-Gutierrez F, Kennedy J, et al. Defining disease phenotypes in primary care electronic health records by a machine learning approach: A case study in identifying rheumatoid arthritis. *PLoS One*. 2016;11:e0154515. DOI: 10.1371/journal.pone.0154515.
- Liu J, Chen N, A 9 mRNAs-based diagnostic signature for rheumatoid arthritis by integrating bioinformatic analysis and machine-learning. *J Orthop Surg Res*. 2021;16:44. DOI: 10.1186/s13018-020-02180-w.
- Long NP, Park S, Anh NH, et al. Efficacy of integrating a novel 16-gene biomarker panel and intelligence classifiers for differential diagnosis of rheumatoid arthritis and osteoarthritis. *J Clin Med*. 2019;8:50. DOI: 10.3390/jcm8010050.
- Chocholova E, Bertok T, Jane E, et al. Glycomics meets artificial intelligence – Potential of glycan analysis for identification of seropositive and seronegative rheumatoid arthritis patients revealed. *Clin Chim Acta*. 2018;481:49–55. DOI: 10.1016/j.cca.2018.02.031.
- Heard BJ, Rosvold JM, Fritzler MJ, et al. A computational method to differentiate normal individuals, osteoarthritis and rheumatoid arthritis patients using serum biomarkers. *J R Soc Interface*. 2014;11:20140428. DOI: 10.1098/rsif.2014.0428.
- Mc Ardle A, Kwasnik A, Szentpetery A, et al. Identification and evaluation of serum protein biomarkers that differentiate psoriatic arthritis from rheumatoid arthritis. *Arthritis Rheumatol*. 2022;74:81–91. DOI: 10.1002/art.41899.
- Mate GS, Kureshi AK, Singh BK. An efficient CNN for hand X-ray classification of rheumatoid arthritis. *J Healthc Eng*. 2021;2021:6712785. DOI: 10.1155/2021/6712785.
- Üreten K, Erbay H, Maraş HH. Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network. *Clin Rheumatol*. 2020;39:969–74. DOI: 10.1007/s10067-019-04487-4.
- Üreten K, Maraş HH. Automated classification of rheumatoid arthritis, osteoarthritis, and normal hand radiographs with deep learning methods. *J Digit Imaging*. 2022;35:193–9. DOI: 10.1007/s10278-021-00564-w.
- Andersen JKH, Pedersen JS, Laursen MS, et al. Neural networks for automatic scoring of arthritis disease activity on ultrasound images. *RMD Open*. 2019;5:e000891. DOI: 10.1136/rmdopen-2018-000891.
- Wu M, Wu H, Wu L, et al. A deep learning classification of metacarpophalangeal joints synovial proliferation in rheumatoid arthritis by ultrasound images. *J Clin Ultrasound*. 2022;50:296–301. DOI: 10.1002/jcu.23143.
- Haj-Mirzaian A, Kubassova O, Boesen M, et al. Computer-assisted image analysis in assessment of peripheral joint MRI in inflammatory arthritis: A systematic review and meta-analysis. *ACR Open Rheumatol*. 2022;4:721–34. DOI: 10.1002/acr2.11450.
- Ahalya RK, Snehalatha U, Dhanraj V. Automated segmentation and classification of hand thermal images in rheumatoid arthritis using machine learning algorithms: A comparison with quantum machine learning technique. *J Therm Biol*. 2023;111:103404. DOI: 10.1016/j.jtherbio.2022.103404.
- Bardhan S, Bhowmik MK. 2-stage classification of knee joint thermograms for rheumatoid arthritis prediction in subclinical inflammation. *Australas Phys Eng Sci Med*. 2019;42:259–77. DOI: 10.1007/s13246-019-00726-9.
- Slansky E, Li J, Häupl T, et al. Quantitative determination of the diagnostic accuracy of the synovitis score and its components. *Histopathology*. 2010;57:436–43. DOI: 10.1111/j.1365-2559.2010.03641.x.
- Mehta B, Goodman S, DiCarlo E, et al. Machine learning identification of thresholds to discriminate osteoarthritis and rheumatoid arthritis synovial inflammation. *Arthritis Res Ther*. 2023;25:31. DOI: 10.1186/s13075-023-03008-8.
- Spencer AK, Bandaria J, Leavy MB, et al. Validation of a machine learning approach to estimate Clinical Disease Activity Index scores for rheumatoid arthritis. *RMD Open*. 2021;7:e001781. DOI: 10.1136/rmdopen-2021-001781.
- Duong SQ, Crowson CS, Athreya A, et al. Clinical predictors of response to methotrexate in patients with rheumatoid arthritis: A machine learning approach using clinical trial data. *Arthritis Res Ther*. 2022;24:162. DOI: 10.1186/s13075-022-02851-5.
- Gosselt HR, Verhoeven MMA, Bulatovic-Čalasan M, et al. Complex machine-learning algorithms and multivariable logistic regression on par in the prediction of insufficient clinical response to methotrexate in rheumatoid arthritis. *J Pers Med*. 2021;11:44. DOI: 10.3390/jpm11010044.
- Kim W, Jin Oh S, Thi Trinh N, et al. Effects of RETN polymorphisms on treatment response in rheumatoid arthritis patients receiving TNF- α inhibitors and utilization of machine-learning algorithms. *Int Immunopharmacol*. 2022;111:109094. DOI: 10.1016/j.intimp.2022.109094.
- Kim W, Kim TH, Oh SJ, et al. Association of TLR 9 gene polymorphisms with remission in patients with rheumatoid arthritis receiving TNF- α inhibitors and development of machine learning models. *Sci Rep*. 2021;11:20169. DOI: 10.1038/s41598-021-99625-x.
- Johansson FD, Collins JE, Yau V, et al. Predicting response to tocilizumab monotherapy in rheumatoid arthritis: A real-world data analysis using machine learning. *J Rheumatol*. 2021;48:1364–70. DOI: 10.3899/jrheum.201626.
- Rehberg M, Giegerich C, Praestgaard A, et al. Identification of a rule to predict response to sarilumab in patients with rheumatoid arthritis using machine learning and clinical trial data. *Rheumatol Ther*. 2021;8:1661–75. DOI: 10.1007/s40744-021-00361-5.
- Rivellese F, Surace AEA, Goldmann K, et al. Rituximab versus tocilizumab in rheumatoid arthritis: Synovial biopsy-based biomarker analysis of the phase 4 R4RA randomized trial. *Nat Med*. 2022;28:1256–68. DOI: 10.1038/s41591-022-01789-0.
- Sunaga Y, Watanabe A, Katsumata N, et al. A simple scoring model based on machine learning predicts intravenous immunoglobulin resistance in Kawasaki disease. *Clin Rheumatol*. 2023;42:1351–61. DOI: 10.1007/s10067-023-06502-1.
- Duff LM, Scarsbrook AF, Ravikumar N, et al. An automated method for artificial intelligence assisted diagnosis of active aortitis using radiomic analysis of FDG PET-CT images. *Biomolecules*. 2023;13:343. DOI: 10.3390/biom13020343.
- Coticini E, d'Alessandro M, Bergantini L, et al. KL-6 in ANCA-associated vasculitis patients with and without ILD: A machine learning approach. *Biology (Basel)*. 2022;11:94. DOI: 10.3390/biology11010094.
- Doan S, Maehara CK, Chaparro JD, et al. Building a natural language processing tool to identify patients with high clinical suspicion for Kawasaki disease from emergency department notes. *Acad Emerg Med*. 2016;23:628–36. DOI: 10.1111/acem.12925.
- Novovoy M, Dionne A, McCrindle BW, et al. Deep learning-based approach to automatically assess coronary distensibility following Kawasaki disease. *Pediatr Cardiol*. 2022;43:807–15. DOI: 10.1007/s00246-021-02790-z.
- Güler I, Ubeylil ED. Detection of ophthalmic arterial Doppler signals with Behcet disease using multilayer perceptron neural network. *Comput Biol Med*. 2005;35:121–32. DOI: 10.1016/j.combiomed.2003.12.007.
- Kim JM, Kang JG, Kim S, et al. Deep-learning system for real-time differentiation between Crohn's disease, intestinal Behçet's disease, and intestinal tuberculosis. *J Gastroenterol Hepatol*. 2021;36:2141–8. DOI: 10.1111/jgh.15433.
- Lee H, Eun Y, Hwang JY, et al. Explainable deep learning algorithm for distinguishing incomplete Kawasaki disease by coronary artery lesions on echocardiographic imaging. *Comput Methods Programs Biomed*. 2022;223:106970. DOI: 10.1016/j.cmpb.2022.106970.
- Roncato C, Perez L, Brochet-Guégan A, et al. Colour Doppler ultrasound of temporal arteries for the diagnosis of giant cell arteritis: A multicentre deep learning study. *Clin Exp Rheumatol*. 2020;38:120–5.
- Portman MA, Margaret CA, Barnes G, et al. An artificial intelligence derived blood test to diagnose Kawasaki disease. *Hosp Pediatr*. 2023;13:201–10. DOI: 10.1542/hpeds.2022-006868.
- Venerito V, Emmi G, Cantarini L, et al. Validity of machine learning in predicting giant cell arteritis flare after glucocorticoids tapering. *Front Immunol*. 2022;13:860877. DOI: 10.3389/fimmu.2022.860877.
- Wang T, Liu G, Lin H. A machine learning approach to predict intravenous immunoglobulin resistance in Kawasaki disease patients: A study based on a southeast China population. *PLoS One*. 2020;15:e0237321. DOI: 10.1371/journal.pone.0237321.
- Hammam N, Bakhiet A, El-Latif EA, et al. Development of machine learning models for detection of vision threatening Behçet's disease (BD) using Egyptian College of Rheumatology (ECR)-BD cohort. *BMC Med Inform Decis Mak*. 2023;23:37. DOI: 10.1186/s12911-023-02130-6.
- Keino H, Wakitani T, Sunayama W, et al. Quantitative analysis of retinal vascular leakage in retinal vasculitis using machine learning. *Appl Sci*. 2022;12:12751. DOI: 10.3390/app122412751.
- Wang J, Chu H, Pan Y. Prediction of renal damage in children with IgA vasculitis based on machine learning. *Medicine*. 2022;101:e31135. DOI: 10.1097/MD.0000000000003135.
- Zimmermann M, Klaus M, Wong MN, et al. Deep learning-based molecular morphometrics for kidney biopsies. *JCI Insight*. 2021;6:e144779. DOI: 10.1172/jci.insight.144779.
- Standardization of Uveitis Nomenclature (SUN) Working Group. Classification criteria for Behçet disease uveitis. *Am J Ophthalmol*. 2021;228:80–8. DOI: 10.1016/j.ajo.2021.03.058.
- Senusi AA, Liu J, Bevec D, et al. Why are Behçet's disease patients always exhausted? *Clin Exp Rheumatol*. 2018;36:53–62.
- Chupp G, Baudy P, Carstens D, et al. Utilizing machine learning with claims data to diagnose and quantify the prevalence of eosinophilic granulomatosis with polyangiitis. *J Allergy Clin Immunol*. 2024;153:AB215. DOI: 10.1016/j.jaci.2023.11.694.
- Linder R, Orth I, Hagen EC, et al. Differentiation between Wegener's granulomatosis and microscopic polyangiitis by an artificial neural network and by traditional methods. *J Rheumatol*. 2011;38:1039–47. DOI: 10.3899/jrheum.100814.
- Xu E, Nemati S, Tremoulet AH. A deep convolutional neural network for Kawasaki disease diagnosis. *Sci Rep*. 2022;12:11438. DOI: 10.1038/s41598-022-15495-x.
- Zhang M, Ke B, Zhuo H, Guo B. Diagnostic model based on bioinformatics and machine learning to distinguish Kawasaki disease using multiple datasets. *BMC Pediatr*. 2022;22:512. DOI: 10.1186/s12887-022-03557-y.
- Lam JY, Shimizu C, Tremoulet AH, et al. A machine-learning algorithm for diagnosis of multisystem inflammatory syndrome in children and Kawasaki disease in the USA: A retrospective model development and validation study. *Lancet Digit Health*. 2022;4:e717–26. DOI: 10.1016/S2589-7500(22)00149-2.
- Wang H, Tan X, Huang Z, et al. Mining incomplete clinical data for the early assessment of Kawasaki disease based on feature clustering and convolutional neural networks. *Artif Intell Med*. 2020;105:101859. DOI: 10.1016/j.artmed.2020.101859.
- Tsai C-M, Lin C-H, Kuo H-C, et al. Use of machine learning to differentiate children with Kawasaki disease from other febrile children in a pediatric emergency department. *JAMA Netw Open*. 2023;6:e237489. DOI: 10.1001/jamanetworkopen.2023.7489.
- Ghosh P, Katkar GD, Shimizu C, et al. An artificial intelligence-guided signature reveals the shared host immune response in MIS-C and Kawasaki disease. *Nat Commun*. 2022;13:2687. DOI: 10.1038/s41467-022-30357-w.
- Martorell-Marugán J, Chierici M, Jurman G, et al. Differential diagnosis of systemic lupus erythematosus and Sjögren's syndrome using machine learning and multi-omics data. *Comput Biol Med*. 2023;152:106373. DOI: 10.1016/j.combiomed.2022.106373.
- Wu J, Yang W, Li H. An artificial neural network model based on autophagy-related genes in childhood systemic lupus erythematosus. *Hereditas*. 2022;159:34. DOI: 10.1186/s41065-022-00248-7.
- Yang K, Wang Q, Wu L, et al. Development and verification of a combined diagnostic model for primary Sjögren's syndrome by integrated bioinformatics analysis and machine learning. *Sci Rep*. 2023;13:8641. DOI: 10.1038/s41598-023-35864-4.
- Zhong Y, Zhang W, Hong X, et al. Screening biomarkers for systemic lupus erythematosus based on machine learning and exploring their expression correlations with the ratios of various immune cells. *Front Immunol*. 2022;13:873787. DOI: 10.3389/fimmu.2022.873787.
- Zhou L, Wang H, Zhang H, et al. Diagnostic markers and potential therapeutic agents for Sjögren's syndrome screened through multiple machine learning and molecular docking. *Clin Exp Immunol*. 2023;212:224–38. DOI: 10.1093/cei/uxad037.
- Bharathi PG, Berks M, Dinsdale G, et al. A deep learning system for quantitative assessment of microvascular abnormalities in nailfold capillary images. *Rheumatology*. 2023;62:2325–9. DOI: 10.1093/rheumatology/kead026.
- Correia C, Mawe S, Lofgren S, et al. High-throughput quantitative histology in systemic sclerosis skin disease using

- computer vision. *Arthritis Res Ther.* 2020;22:48. DOI: 10.1186/s13075-020-2127-0.
65. Fabry V, Mamalet F, Laforet A, et al. A deep learning tool without muscle-by-muscle grading to differentiate myositis from facio-scapulo-humeral dystrophy using MRI. *Diagn Interv Imaging.* 2022;103:353–9. DOI: 10.1016/j.diii.2022.01.012.
 66. Garaiman A, Nooralahzadeh F, Mihai C, et al. Vision transformer assisting rheumatologists in screening for capillaroscopy changes in systemic sclerosis: An artificial intelligence model. *Rheumatology.* 2023;62:2492–500. DOI: 10.1093/rheumatology/keac541.
 67. Kise Y, Ikeda H, Fujii T, et al. Preliminary study on the application of deep learning system to diagnosis of Sjögren's syndrome on CT images. *Dentomaxillofac Radiol.* 2019;48:20190019. DOI: 10.1259/dmfr.20190019.
 68. Kise Y, Shimizu M, Ikeda H, et al. Usefulness of a deep learning system for diagnosing Sjögren's syndrome using ultrasonography images. *Dentomaxillofac Radiol.* 2020;49:20190348. DOI: 10.1259/dmfr.20190348.
 69. Nitkunanantharajah S, Haedicke K, Moore TB, et al. Three-dimensional optoacoustic imaging of nailfold capillaries in systemic sclerosis and its potential for disease differentiation using deep learning. *Sci Rep.* 2020;10:16444. DOI: 10.1038/s41598-020-73319-2.
 70. Vukicevic AM, Radovic M, Zabotti A, et al. Deep learning segmentation of primary Sjögren's syndrome affected salivary glands from ultrasonography images. *Comput Biol Med.* 2021;129:104154. DOI: 10.1016/j.compbimed.2020.104154.
 71. Yuan Y, Quan T, Song Y, et al. Noise-immune extreme ensemble learning for early diagnosis of neuropsychiatric systemic lupus erythematosus. *IEEE J Biomed Health Inform.* 2022;26:3495–506. DOI: 10.1109/JBHI.2022.3164937.
 72. Baldini C, Ferro F, Luciano N, et al. Artificial neural networks help to identify disease subsets and to predict lymphoma in primary Sjögren's syndrome. *Clin Exp Rheumatol.* 2018;36:137–44.
 73. Wang M, Liang Y, Hu Z, et al. Lupus nephritis diagnosis using enhanced moth flame algorithm with support vector machines. *Comput Biol Med.* 2022;145:105435. DOI: 10.1016/j.compbimed.2022.105435.
 74. Nagawa K, Suzuki M, Yamamoto Y, et al. Texture analysis of muscle MRI: Machine learning-based classifications in idiopathic inflammatory myopathies. *Sci Rep.* 2021;11:9821. DOI: 10.1038/s41598-021-89311-3.
 75. Tanmeemaat MR, Kefalas M, Geraetsdts VJ, et al. Distinguishing normal, neuropathic and myopathic EMG with an automated machine learning approach. *Clin Neurophysiol.* 2023;146:49–54. DOI: 10.1016/j.clinph.2022.11.019.
 76. Wang F, Zhou S, Hou B, et al. Assessment of idiopathic inflammatory myopathy using a deep learning method for muscle T2 mapping segmentation. *Eur Radiol.* 2023;33:2350–7. DOI: 10.1007/s00330-022-09254-9.
 77. Alves P, Bandaria J, Leavy MB, et al. Validation of a machine learning approach to estimate systemic lupus erythematosus disease activity index score categories and application in a real-world dataset. *RMD Open.* 2021;7:e001586. DOI: 10.1136/rmdopen-2021-001586.
 78. Ceccarelli F, Scandrone M, Perricone C, et al. Biomarkers of erosive arthritis in systemic lupus erythematosus: Application of machine learning models. *PLoS One.* 2018;13:e0207926. DOI: 10.1371/journal.pone.0207926.
 79. Ceccarelli F, Scandrone M, Perricone C, et al. Prediction of chronic damage in systemic lupus erythematosus by using machine-learning models. *PLoS One.* 2017;12:e0174200. DOI: 10.1371/journal.pone.0174200.
 80. Chen Y, Huang S, Chen T, et al. Machine learning for prediction and risk stratification of lupus nephritis renal flare. *Am J Nephrol.* 2021;52:152–60. DOI: 10.1159/000513566.
 81. Choi MY, Chen J, Clarke AE, et al. Machine learning identifies clusters of longitudinal autoantibody profiles predictive of systemic lupus erythematosus disease outcomes. *Ann Rheum Dis.* 2023;82:927–36. DOI: 10.1136/ard-2022-223808.
 82. Fazzari MJ, Guerra MM, Salmon J, et al. Adverse pregnancy outcomes in women with systemic lupus erythematosus: Can we improve predictions with machine learning? *Lupus Sci Med.* 2022;9:e000769. DOI: 10.1136/lupus-2022-000769.
 83. Hu Z, Wu L, Lin Z, et al. Prevalence and associated factors of electrocardiogram abnormalities in patients with systemic lupus erythematosus: A machine learning study. *Arthritis Care Res.* 2022;74:1640–8. DOI: 10.1002/acr.24612.
 84. Jorge AM, Smith D, Wu Z, et al. Exploration of machine learning methods to predict systemic lupus erythematosus hospitalizations. *Lupus.* 2022;31:1296–305. DOI: 10.1177/09612033221114805.
 85. Kegerreis B, Catalina MD, Bachali P, et al. Machine learning approaches to predict lupus disease activity from gene expression data. *Sci Rep.* 2019;9:9617. DOI: 10.1038/s41598-019-45989-0.
 86. Liu C, Zhou Y, Zhou Y, et al. Identification of crucial genes for predicting the risk of atherosclerosis with system lupus erythematosus based on comprehensive bioinformatics analysis and machine learning. *Comput Biol Med.* 2023;152:106388. DOI: 10.1016/j.compbimed.2022.106388.
 87. Ma J-H, Feng Z, Wu J-Y, et al. Learning from imbalanced fetal outcomes of systemic lupus erythematosus in artificial neural networks. *BMC Med Inform Decis Mak.* 2021;21:127. DOI: 10.1186/s12911-021-01486-x.
 88. Reddy BK, Delen D. Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology. *Comput Biol Med.* 2018;101:199–209. DOI: 10.1016/j.compbimed.2018.08.029.
 89. Rumetschofer T, Inglesse F, de Bresser J, et al. Tract-based white matter hyperintensity patterns in patients with systemic lupus erythematosus using an unsupervised machine learning approach. *Sci Rep.* 2022;12:21376. DOI: 10.1038/s41598-022-25990-w.
 90. Stojanowski J, Konieczny A, Ryzdyńska K, et al. Artificial neural network – An effective tool for predicting the lupus nephritis outcome. *BMC Nephrol.* 2022;23:381. DOI: 10.1186/s12882-022-02978-2.
 91. Zhao Y, Smith D, Jorge A. Comparing two machine learning approaches in predicting lupus hospitalization using longitudinal data. *Sci Rep.* 2022;12:16424. DOI: 10.1038/s41598-022-20845-w.
 92. Bauer Y, de Bernard S, Hickey P, et al. Identifying early pulmonary arterial hypertension biomarkers in systemic sclerosis: Machine learning on proteomics from the DETECT cohort. *Eur Respir J.* 2021;57:2002591. DOI: 10.1183/13993003.02591-2020.
 93. Murdaca G, Caprioli S, Tonacci A, et al. A machine learning application to predict early lung involvement in scleroderma: A feasibility evaluation. *Diagnosics.* 2021;11:1880. DOI: 10.3390/diagnosics11101880.
 94. Andrade DSM, Ribeiro LM, Lopes AJ, et al. Machine learning associated with respiratory oscillometry: A computer-aided diagnosis system for the detection of respiratory abnormalities in systemic sclerosis. *Biomed Eng Online.* 2021;20:31. DOI: 10.1186/s12938-021-00865-9.
 95. Ebata S, Oba K, Kashiwabara K, et al. Predictors of rituximab effect on modified Rodnan skin score in systemic sclerosis: A machine-learning analysis of the Desires trial. *Rheumatology.* 2022;61:4364–73. DOI: 10.1093/rheumatology/keac023.
 96. Mehta BK, Espinoza ME, Franks JM, et al. Machine-learning classification identifies patients with early systemic sclerosis as abatacept responders via CD28 pathway modulation. *JCI Insight.* 2022;7:e155282. DOI: 10.1172/jci.insight.155282.
 97. Wang D-D, Li Y-F, Zhang C, et al. Predicting the effect of sirolimus on disease activity in patients with systemic lupus erythematosus using machine learning. *J Clin Pharm Ther.* 2022;47:1845–50. DOI: 10.1111/jcpt.13778.
 98. Liu D, Hu Z, Tang Z, et al. Identification of biomarkers associated with oxidative stress-related genes in postmenopausal osteoporosis. *Cell Mol Biol.* 2023;69:186–92. DOI: 10.14715/cmb/2023.69.6.28.
 99. Huang X, Ma J, Wei Y, et al. Identification of biomarkers associated with diagnosis of postmenopausal osteoporosis patients based on bioinformatics and machine learning. *Front Genet.* 2023;14:1198417. DOI: 10.3389/fgene.2023.1198417.
 100. University of Sheffield. The Fracture Risk Assessment Tool - FRAX Available at: <https://frax.shef.ac.uk/FRAX/tool.aspx?country=9> (Date last accessed: 29 May 2024).
 101. Rahim F, Zaki Zadeh A, Javanmardi P, et al. Machine learning algorithms for diagnosis of hip bone osteoporosis: A systematic review and meta-analysis study. *Biomed Eng Online.* 2023;22:68. DOI: 10.1186/s12938-023-01132-9.
 102. Yasaka K, Akai H, Kunimatsu A, et al. Prediction of bone mineral density from computed tomography: Application of deep learning with a convolutional neural network. *Eur Radiol.* 2020;30:3549–57. DOI: 10.1007/s00330-020-06677-0.
 103. Yang J, Liao M, Wang Y, et al. Opportunistic osteoporosis screening using chest CT with artificial intelligence. *Osteoporos Int.* 2022;33:2547–61. DOI: 10.1007/s00198-022-06491-y.
 104. Hong N, Cho SW, Shin S, et al. Deep-learning-based detection of vertebral fracture and osteoporosis using lateral spine X-ray radiography. *J Bone Miner Res.* 2023;38:887–95. DOI: 10.1002/jbmr.4814.
 105. Roux C, Rozes A, Reizine D, et al. Fully automated opportunistic screening of vertebral fractures and osteoporosis on more than 150 000 routine computed tomography scans. *Rheumatology.* 2022;61:3269–78. DOI: 10.1093/rheumatology/keab878.
 106. Lu S, Fuggle NR, Westbury LD, et al. Machine learning applied to HR-pQCT images improves fracture discrimination provided by DXA and clinical risk factors. *Bone.* 2023;168:116653. DOI: 10.1016/j.bone.2022.116653.
 107. Atkinson EJ, Therneau TM, Melton LJ, et al. Assessing fracture risk using gradient boosting machine (GBM) models. *J Bone Miner Res.* 2012;27:1397–404. DOI: 10.1002/jbmr.1577.
 108. Treece GM, Gee AH, Tonkin C, et al. Predicting hip fracture type with cortical bone mapping (CBM) in the osteoporotic fractures in men (MROS) study. *J Bone Miner Res.* 2015;30:2067–77. DOI: 10.1002/jbmr.2552.
 109. Xiao P, Zhang T, Dong XN, et al. Prediction of trabecular bone architectural features by deep learning models using simulated DXA images. *Bone Rep.* 2020;13:100295. DOI: 10.1016/j.bonr.2020.100295.
 110. Juul-Kristensen B, Schmedling K, Rombaut L, et al. Measurement properties of clinical assessment methods for classifying generalized joint hypermobility – A systematic review. *Am J Med Genet Part C.* 2017;175:116–47. DOI: 10.1002/ajmg.c.31540.
 111. Mittal N, Sabo A, Deshpande A, et al. Feasibility of video-based joint hypermobility assessment in individuals with suspected Ehlers-Danlos syndromes/generalised hypermobility spectrum disorders: A single-site observational study protocol. *BMJ Open.* 2022;12:e068098. DOI: 10.1136/bmjopen-2022-068098.
 112. ClinicalTrials.gov. Vision-based Assessment of Joint Extensibility in Ehlers Danlos Syndrome. [ClinicalTrials.gov/study/NCT05366114](https://clinicaltrials.gov/study/NCT05366114). Available at: <https://clinicaltrials.gov/study/NCT05366114> (Date last accessed: 29 May 2024).
 113. Emmert D, Szczygien N, Bender TTA, et al. A diagnostic support system based on pain drawings: Binary and K-disease classification of EDS, GBS, FSHD, PROMM, and a control group with Pain2D. *Orphanet J Rare Dis.* 2023;18:70. DOI: 10.1186/s13023-023-02663-z.
 114. Jayakumar P, Moore MG, Furlough KA, et al. Comparison of an artificial intelligence-enabled patient decision aid vs educational material on decision quality, shared decision-making, patient experience, and functional outcomes in adults with knee osteoarthritis: A randomized clinical trial. *JAMA Netw Open.* 2021;4:e2037107. DOI: 10.1001/jamanetworkopen.2020.37107.
 115. Liu L, Chang J, Zhang P, et al. A joint multi-modal learning method for early-stage knee osteoarthritis disease classification. *Heliyon.* 2023;9:e15461. DOI: 10.1016/j.heliyon.2023.e15461.
 116. Pongsakonprattikul N, Angthong C, Kittichai V, et al. Artificial intelligence assistance in radiographic detection and classification of knee osteoarthritis and its severity: A cross-sectional diagnostic study. *Eur Rev Med Pharmacol Sci.* 2022;26:1549–58. DOI: 10.26355/eurrev_202203_28220.
 117. El-Ghany SA, Elmog M, El-Aziz AAA. A fully automatic fine tuned deep learning model for knee osteoarthritis detection and progression analysis. *Egypt Inform J.* 2023;24:229–40. DOI: 10.1016/j.eij.2023.03.005.
 118. Neubauer M, Moser L, Neugebauer J, et al. Artificial-intelligence-aided radiographic diagnostic of knee osteoarthritis leads to a higher association of clinical findings with diagnostic ratings. *J Clin Med.* 2023;12:744. DOI: 10.3390/jcm12030744.
 119. Rose MJ, Costello KE, Eigenbrod S, et al. Inertial measurement units and application for remote health care in hip and knee osteoarthritis. *JMIR Rehabil Assist Technol.* 2022;9:e33521. DOI: 10.2196/33521.
 120. Polissett TS, Jain S, Pang M, et al. Concerns surrounding application of artificial intelligence in hip and knee arthroplasty: A review of literature and recommendations for meaningful adoption. *Bone Joint J.* 2022;104-B:1292–303. DOI: 10.1302/0301-620X.104B12-BJJ-2022-0922.R1.
 121. Huo J, Huang G, Han D, et al. Value of 3D preoperative planning for primary total hip arthroplasty based on artificial intelligence technology. *J Orthop Surg Res.* 2021;16:156. DOI: 10.1186/s13018-021-02294-9.
 122. Kunze KN, Jiang SJ, Li TY, et al. Artificial intelligence for automated identification of total shoulder arthroplasty implants. *J Shoulder Elbow Surg.* 2023;32:2115–22. DOI: 10.1016/j.jse.2023.03.028.
 123. Lambrechts A, Wirix-Speetjens R, Maes F, et al. Artificial intelligence based patient-specific preoperative planning algorithm for total knee arthroplasty. *Front Robot AI.* 2022;9:840282. DOI: 10.3389/frobt.2022.840282.
 124. Longo UG, De Salvatore S, Candela V, et al. Augmented reality, virtual reality and artificial intelligence in orthopedic surgery: A systematic review. *Appl Sci.* 2021;11:3253. DOI: 10.3390/app11073253.
 125. Matthews JH, Shields JS. The clinical application of augmented reality in orthopaedics: Where do we stand? *Curr Rev Musculoskelet Med.* 2021;14:316–9. DOI: 10.1007/s12178-021-09713-8.
 126. Moussaoui G, Zakaria AS, Negrean C, et al. Accuracy of Clarius, handheld wireless point-of-care ultrasound, in evaluating prostate morphology and volume compared to radical prostatectomy specimen weight: Is there a difference between transabdominal vs transrectal approach? *J Endourol.* 2021;35:1300–6. DOI: 10.1089/end.2020.0874.
 127. Smolle MA, Goetz C, Maurer D, et al. Artificial intelligence-based computer-aided system for knee osteoarthritis assessment increases experienced orthopaedic surgeons' agreement rate and accuracy. *Knee Surg Sports Traumatol Arthrosc.* 2023;31:1053–62. DOI: 10.1007/s00167-022-07220-y.
 128. Sultan H, Owais M, Choi J, et al. Artificial intelligence-based solution in personalized computer-aided arthroscopy of shoulder prostheses. *J Pers Med.* 2022;12:109. DOI: 10.3390/jpm12010109.
 129. Gytopoulos S, Lin D, Knoll F, et al. Artificial intelligence in musculoskeletal imaging: Current status and future directions. *AJR Am J Roentgenol.* 2019;213:506–13. DOI: 10.2214/AJR.19.21117.
 130. Kunze KN, Rossi DM, White GM, et al. Diagnostic performance of artificial intelligence for detection of anterior cruciate ligament and meniscus tears: A systematic review. *Arthrosc J Arthrosc Related Surg.* 2021;37:771–81. DOI: 10.1016/j.arthro.2020.09.012.
 131. Kim K, Kim S, Lee YH, et al. Performance of the deep convolutional neural network based magnetic resonance image scoring algorithm for differentiating between tuberculous and pyogenic spondylitis. *Sci Rep.* 2018;8:13124. DOI: 10.1038/s41598-018-31486-3.
 132. Yu JS, Yu SM, Erdal BS, et al. Detection and localisation of hip fractures on anteroposterior radiographs with artificial intelligence: Proof of concept. *Clin Radiol.* 2020;75:237. DOI: 10.1016/j.crad.2019.10.022.
 133. Jimenez-Pastor A, Alberich-Bayarri A, Fos-Guarinos B, et al. Automated vertebral localization and identification by decision forests and image-based refinement on real-world CT data. *Radiol Med.* 2020;125:48–56. DOI: 10.1007/s11547-019-01079-9.
 134. Ding X, Zhang B, Li W, et al. Value of preoperative three-dimensional planning software (AI-HIP) in primary total hip arthroplasty: A retrospective study. *J Int Med Res.* 2021;49:3000605211058874. DOI: 10.1177/03000605211058874.
 135. Carson T, Ghoshal G, Cornwall GB, et al. Artificial intelligence-enabled, real-time intraoperative ultrasound imaging of neural structures within the Psoas: validation in a porcine spine

- model. *Spine (Phila Pa 1976)*. 2021;46:E146–52. DOI: 10.1097/BRS.0000000000003704.
136. Viderman D, Dossov M, Seitenov S, Lee MH. Artificial intelligence in ultrasound-guided regional anesthesia: A scoping review. *Front Med*. 2022;9:994805. DOI: 10.3389/fmed.2022.994805.
137. Sultan H, Owais M, Park C, et al. Artificial intelligence-based recognition of different types of shoulder implants in X-ray scans based on dense residual ensemble-network for personalized medicine. *J Pers Med*. 2021;11:482. DOI: 10.3390/jpm11060482.
138. Borjali A, Chen AF, Bedair HS, et al. Comparing the performance of a deep convolutional neural network with orthopedic surgeons on the identification of total hip prosthesis design from plain radiographs. *Med Phys*. 2021;48:2327–36. DOI: 10.1002/mp.14705.
139. Zhang Y, Yu H. Convolutional neural network based metal artifact reduction in X-ray computed tomography. *IEEE Trans Med Imaging*. 2018;37:1370–81. DOI: 10.1109/TMI.2018.2823083.
140. Carrodegua E, Lacson R, Swanson W, et al. Use of machine learning to identify follow-up recommendations in radiology reports. *J Am Coll Radiol*. 2019;16:336–43. DOI: 10.1016/j.jacr.2018.10.020.
141. Lee YH. Efficiency improvement in a busy radiology practice: Determination of musculoskeletal magnetic resonance imaging protocol using deep-learning convolutional neural networks. *J Digit Imaging*. 2018;31:604–10. DOI: 10.1007/s10278-018-0066-y.
142. Hirschmann A, Cyriac J, Stieltjes B, et al. Artificial intelligence in musculoskeletal imaging: Review of current literature, challenges, and trends. *Semin Musculoskelet Radiol*. 2019;23:304–11. DOI: 10.1055/s-0039-1684024.
143. Trivedi H, Mesterhazy J, Laguna B, et al. Automatic determination of the need for intravenous contrast in musculoskeletal MRI examinations using IBM Watson's natural language processing algorithm. *J Digit Imaging*. 2018;31:245–51. DOI: 10.1007/s10278-017-0021-3.
144. McKee M, Wouters OJ. The challenges of regulating artificial intelligence in healthcare comment on clinical decision support and new regulatory frameworks for medical devices: Are we ready for it? A viewpoint paper. *Int J Health Policy Manag*. 2023;12:7261. DOI: 10.34172/ijhpm.2022.7261.
145. Groot OQ, Bongers MER, Ogink PT, et al. Does artificial intelligence outperform natural intelligence in interpreting musculoskeletal radiological studies? A systematic review. *Clin Orthop Relat Res*. 2020;478:2751–64. DOI: 10.1097/CORR.0000000000001360.
146. Naik N, Hameed BMZ, Shetty DK, et al. Legal and ethical consideration in artificial intelligence in healthcare: Who takes responsibility?. *Front Surg*. 2022;9:862322. DOI: 10.3389/fsurg.2022.862322.