International Journal of Innovative Technologies in Social Science

e-ISSN: 2544-9435

Operating Publisher
SciFormat Publishing Inc.
ISNI: 0000 0005 1449 8214

2734 17 Avenue SW,
Calgary, Alberta, T3E0A7,
Canada
+15878858911
editorial-office@sciformat.ca

ARTICLE TITLE	ARTIFICIAL INTELLIGENCE IN THE EARLY DETECTION AND PROGRESSION PREDICTION OF OSTEOARTHRITIS: A NARRATIVE REVIEW OF CURRENT TECHNOLOGIES AND CLINICAL IMPLICATIONS
DOI	https://doi.org/10.31435/ijitss.1(49).2026.5620
RECEIVED	06 February 2026
ACCEPTED	14 March 2026
PUBLISHED	25 March 2026
LICENSE	The article is licensed under a **Creative Commons Attribution 4.0 International License.**

ARTIFICIAL INTELLIGENCE IN THE EARLY DETECTION AND PROGRESSION PREDICTION OF OSTEOARTHRITIS: A NARRATIVE REVIEW OF CURRENT TECHNOLOGIES AND CLINICAL IMPLICATIONS

Piotr Tryczyński *(Corresponding Author, Email: piter.tryczynski@gmail.com)*
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0001-8997-3225

Jakub Wrona
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0005-7722-7507

Jakub Sałak
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0005-7078-6402

Piotr Helbin
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0007-5289-2521

Aleksandra Gralec
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0001-0061-311X

Sebastian Ożga
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0003-1337-7800

Wiktoria Donocik
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0003-3801-6729

Aleksandra Spirkowicz
Medical University of Silesia in Katowice, Katowice, Poland
ORCID ID: 0009-0006-2228-4536

ABSTRACT

Background and Objectives: Knee Osteoarthritis (KOA) is a leading cause of global disability, yet current management remains reactive due to the limitations of the Kellgren-Lawrence (KL) grading system. This review aims to synthesize technological advancements from 2018–2026 to provide a framework for AI-driven precision orthopedics.
Methods: A comprehensive narrative review was conducted by searching major electronic databases for peer-reviewed research focusing on AI-driven KOA diagnostics and prognosis. Evidence was categorized into seven thematic domains, ranging from automated grading to omics-based molecular discovery.
Findings: Technical analysis reveals that multi-stage pipelines utilizing YOLOv2 and Faster R-CNN isolate joint spaces with over 98% accuracy, reducing diagnostic errors by 5–7%. Natural Language Processing (NLP) models, specifically BiLSTM architectures, identify KOA risk 24 to 36 months before radiographic confirmation with an AUC of 0.911. Multi-modal frameworks like LBTRBC-M achieve a predictive AUC of 0.913 for structural progression and improve the prognostic accuracy of resident physicians from approximately 45% to over 66%. Furthermore, Generative AI (GPT-4) delivers personalized patient guidance 14 times faster than clinicians. In molecular research, AI has identified PDK1 as a critical regulatory gene for chondrocyte autophagy.
Conclusions: AI is shifting KOA management from reactive diagnosis to proactive health forecasting. However, the wide-scale adoption of these technologies depends on robust socio-ethical governance, addressing algorithmic bias, and fostering trust through Explainable AI. Future frontiers include the development of "digital twins" and integrated omics-imaging models.

KEYWORDS

Osteoarthritis, Artificial Intelligence, Deep Learning, Early Detection, Large Language Models, Clinical Implications

CITATION

Piotr Tryczyński, Jakub Wrona, Jakub Sałak, Piotr Helbin, Aleksandra Gralec, Sebastian Ożga, Wiktoria Donocik, Aleksandra Spirkowicz. (2026) Artificial Intelligence in the Early Detection and Progression Prediction of Osteoarthritis: A Narrative Review of Current Technologies and Clinical Implications. *International Journal of Innovative Technologies in Social Science*. 1(49). doi: 10.31435/ijitss.1(49).2026.5620

1. Introduction

Knee Osteoarthritis (KOA) represents one of the most significant challenges to global public health in the 21st century, acting as a leading cause of chronic pain and functional disability. Characterized by the progressive and irreversible degradation of articular cartilage, subchondral bone remodeling, and chronic synovial inflammation, KOA imposes a profound socio-economic burden on healthcare infrastructures. As life expectancy increases and the prevalence of obesity continues to rise globally, the number of individuals suffering from severe joint degradation is projected to escalate, necessitating a fundamental shift in orthopedic workflows. Current medical management often remains reactive, primarily focusing on late-stage palliative care or invasive surgical interventions, such as total knee replacement (TKR), due to the lack of sensitive tools for early-stage detection (Tolpadi et al., 2020).

The foundational bottleneck in modern KOA management is the reliance on the Kellgren-Lawrence (KL) grading system, which has served as the gold standard for radiographic assessment for decades (Tiulpin et al., 2018). Despite its widespread adoption, the KL scale is plagued by high inter-observer variability and a significant lack of sensitivity toward early, "pre-radiographic" structural markers. Clinicians frequently encounter a clinical-radiographic mismatch, where patients report debilitating mechanical pain and functional stiffness despite radiographs showing no definitive evidence of joint space narrowing (KL grades 0–1). This diagnostic gap often results in delayed intervention, allowing the disease to progress to a stage where conservative management is no longer effective.

The emergence of Artificial Intelligence (AI), specifically Deep Learning (DL) and Convolutional Neural Networks (CNNs), offers a transformative solution to these diagnostic limitations. Between 2018 and 2026, the field has witnessed a multi-dimensional technological shift. AI is transitioning from an auxiliary tool for static image classification into a comprehensive predictive engine capable of longitudinal health forecasting. Innovative pipelines utilizing one-stage detectors like YOLOv2 and two-stage architectures such as Faster R-CNN have demonstrated the ability to isolate joint spaces with unprecedented precision, mitigating the "background noise" that often hampers human interpretation (Chen et al., 2020). By standardizing the initial localization and pre-processing stages, these technologies provide a more robust foundation for early-stage detection and personalized care.

Furthermore, the integration of Natural Language Processing (NLP) into clinical workflows has unlocked "symptomatic intelligence" hidden within unstructured clinician notes (Thanyakunsajja et al., 2025). Research indicates that AI-driven analysis of patient narratives can flag osteoarthritis risk with a lead-time of 24 to 36 months before structural damage is visible on standard radiographs. This capability is critical for social health policy, as it allows public health systems to pivot toward non-pharmacological interventions, such as preventive physical therapy and targeted weight management, thereby reducing the lifetime social cost of disability.

As these technologies move from research laboratories to clinical implementation, the discourse has shifted toward socio-ethical governance—a key priority for the International Journal of Innovative Technologies and Social Science (IJITSS). The deployment of AI in social infrastructure raises complex questions regarding algorithmic bias, the "liability gap" in automated decision-making, and the democratization of healthcare. While high-complexity multi-modal models combining MRI radiomics and biochemical biomarkers offer the highest predictive ceiling, the scalability of lightweight radiographic tools remains essential for broad clinical adoption, particularly in resource-limited community health centers.

This narrative review provides a systematic and comprehensive synthesis of the evidence from core publications spanning 2018 to 2026. The findings are categorized into seven critical thematic domains: (1) Automated Diagnostic Grading and Image Pre-processing; (2) Multi-modal Structural Progression Prediction; (3) Symptomatic Progression and Pain Forecasting; (4) Natural Language Processing (NLP) of Clinical Narratives; (5) Generative AI in Patient Education and Visual Forecasting; (6) Omics and Molecular Biomarker Discovery; and (7) Systematic Benchmarking and Socio-Ethical Governance. By evaluating the intersection of technical innovation and clinical practice, this review aims to provide a robust framework for the future of AI-driven precision orthopedics and its impact on global social health.

2. Methodology

2.1. Study Design

This study is conducted as a comprehensive narrative review, focusing on the intersection of artificial intelligence (AI) and knee osteoarthritis (KOA) diagnostics and prognosis. Given the rapid evolution of deep learning architectures and large language models (LLMs) between 2018 and 2026, a narrative approach was selected to synthesize heterogeneous data types—ranging from radiographic imaging metrics to molecular omics and ethical governance frameworks. The review aims to provide a qualitative and quantitative synthesis of state-of-the-art technologies that are currently redefining orthopedic clinical workflows.

2.2. Search Strategy and Information Sources

A systematic literature search was performed across major electronic databases, including PubMed/MEDLINE, Scopus, Google Scholar, and IEEE Xplore. The search period was strictly limited to publications released between January 2018 and early 2026 to capture the most current technological shifts, particularly the emergence of generative AI and spatial transcriptomics.

The search utilized a combination of Medical Subject Headings (MeSH) terms and specific keywords, including: "Knee Osteoarthritis," "Deep Learning," "Convolutional Neural Networks," "Large Language Models," "Radiomics," "Disease Progression Prediction," and "Socio-ethical Governance". To ensure comprehensive coverage, "snowballing" techniques were applied, wherein the reference lists of core publications and meta-analyses were manually screened for additional relevant studies.

2.3. Eligibility Criteria and Study Selection

To maintain a high standard of evidentiary rigor, specific inclusion and exclusion criteria were established:

- **Inclusion Criteria:** (1) Peer-reviewed original research, systematic reviews, or meta-analyses published in English; (2) Studies focusing on AI-driven detection, grading, or progression prediction of KOA;

(3) Research providing validated performance metrics such as AUC, Sensitivity, Specificity, or Kappa coefficients; (4) Papers discussing the socio-ethical or legal implications of medical AI.

• **Exclusion Criteria:** (1) Case reports with small sample sizes (n<10); (2) Studies focusing solely on non-human (animal) models; (3) Conference abstracts without full-text availability; (4) Proprietary white papers lacking peer-review validation.

2.4. Data Extraction and Thematic Synthesis

The primary data extraction process focused on capturing technical architectures (e.g., YOLOv2, Faster R-CNN, BiLSTM), dataset characteristics (e.g., OAI, MOST, FNIH), and clinical outcomes. Following the initial extraction, the synthesized evidence was categorized into seven thematic domains:

1. Automated Diagnostic Grading and Image Pre-processing;
2. Multi-modal Structural Progression Prediction;
3. Symptomatic Progression and Pain Forecasting;
4. Natural Language Processing (NLP) of Clinical Narratives;
5. Generative AI in Patient Education and Visual Forecasting;
6. Omics and Molecular Biomarker Discovery;
7. Systematic Benchmarking and Socio-Ethical Governance.

Each domain was analyzed not only for its technical precision but also for its clinical scalability and socio-economic impact, aligning with the interdisciplinary scope of the International Journal of Innovative Technologies and Social Science (IJITSS).

3. Results

The systematic synthesis of the evidence from the reviewed core publications (2018–2026) reveals a multi-dimensional technological shift in the management of Knee Osteoarthritis (KOA). The findings are categorized into seven thematic domains: (1) Automated Diagnostic Grading and Image Pre-processing; (2) Multi-modal Structural Progression Prediction; (3) Symptomatic Progression and Pain Forecasting; (4) Natural Language Processing (NLP) of Clinical Narratives; (5) Generative AI in Patient Education and Visual Forecasting; (6) Omics and Molecular Biomarker Discovery; and (7) Systematic Benchmarking and Socio-Ethical Governance.

3.1. Automated Diagnostic Grading and Radiographic Analysis

The automation of the Kellgren-Lawrence (KL) grading system remains the foundational application of AI in orthopedics, with current research focusing on reducing inter-observer variability and enhancing feature extraction robustness.

3.1.1. Technical Nuances of Multi-Stage Pipelines and Localization Accuracy

The transition from semi-automatic to fully automatic knee osteoarthritis (KOA) grading is fundamentally predicated on the precision of the initial localization stage. The synthesis of results from Chen et al. (2019) and Sheik Abdullah and Rajasekaran (2022) confirms that isolating the joint space from full-size radiographs is the most effective method to mitigate "background noise" and enhance the signal-to-noise ratio for downstream classification models.

One-Stage Detection: The YOLOv2 Implementation

As documented by Chen et al. (2019), the use of a customized one-stage YOLOv2 (You Only Look Once) network revolutionized the processing of high-resolution X-ray images.

• **Input Resolution and Anchors:** The researchers processed images at a resolution of 1024 × 1024 pixels. To account for the specific anatomical proportions of the knee joint in anterior-posterior (AP) views, they customized the anchor boxes within the YOLOv2 framework. This adaptation was crucial to ensure that the bounding boxes tightly encapsulated the tibial and femoral margins without including excessive soft tissue.

• **Detection Performance:** This customized YOLOv2 implementation achieved a recall of 92.2% under a Jaccard index (IoU) threshold of 0.75, with a mean IoU of 0.858. Furthermore, the model successfully localized 99.9% of knee joints when applying a more relaxed IoU threshold of 0.50. From a clinical perspective, this high reliability in automated cropping ensures that the grading phase (KL 0–4) is not contaminated by irrelevant skeletal features (such as the mid-shaft of the femur or tibia), which otherwise could mislead the convolutional neural network (CNN) during feature extraction.

• **Computational Efficiency:** Being a one-stage detector, YOLOv2 offered a significant advantage in inference speed, a result that supports the scalability of AI tools in high-throughput hospital environments.

Two-Stage Detection: Faster R-CNN for Joint Space Width (JSW)

In contrast to the single-pass approach of YOLOv2, Sheik Abdullah and Rajasekaran (2022) utilized the Faster R-CNN (Region-based Convolutional Neural Network) architecture. Their objective was even more specialized: the localization of the Joint Space Width (JSW) region.

- **Architecture Synergy:** The Faster R-CNN was coupled with a ResNet-50 backbone for initial feature extraction. This two-stage process—where a Region Proposal Network (RPN) first identifies potential joint areas before a final classification/refining stage—proved exceptionally effective for digital X-ray images of patients over 50 years old.
- **Feature Extraction Consistency:** By specifically targeting the JSW region, this model provided a refined ROI (Region of Interest) that focuses on the primary site of cartilage degradation. This level of localization precision is vital for the detection of early-stage OA (KL grade 1), where the narrowing of the joint space is often too subtle for standard, full-image classifiers to detect.

Impact of Localization on Diagnostic Stability

The comparative analysis of these results indicates that the primary socio-clinical benefit of multi-stage pipelines is the reduction of inter-observer variability.

- **Standardization:** Automated localization provides a standardized cropping mechanism that remains consistent across different X-ray machines and imaging protocols.
- **Error Mitigation:** Both Chen et al. (2019) and Sheik Abdullah (2022) concluded that failing to implement a robust localization stage leads to a drop in classification accuracy of approximately 5–7%, particularly in "Doubtful" (KL 1) cases. In a social healthcare context, this translates to thousands of potential misdiagnoses that can be avoided through these innovative detection technologies.

3.1.2. The Innovation of Ordinal Loss and Siamese Networks Conventional classification treat KL grades (0–4) as independent categories, but the reviewed literature highlights a shift toward acknowledging the disease's progressive nature.

- **Adjustable Ordinal Loss:** Chen et al. (2019) demonstrated that a VGG-16 network combined with a novel ordinal loss function achieved a quadratic weighted Kappa of 0.852. The fine-tuned VGG-19 model using the proposed ordinal loss reached a peak classification accuracy of 70.4% on automatically detected joints and 69.6% on manually annotated joints, significantly outperforming standard Cross-Entropy (CE) models. This mathematical innovation penalizes "distant" misclassifications (e.g., KL 0 vs. KL 4) more heavily than "adjacent" ones (e.g., KL 1 vs. KL 2).
- **Siamese Architectures:** Tiulpin et al. (2018) introduced the Deep Siamese CNN, which mimics comparative clinical logic by analyzing both compartments of the joint simultaneously. Their model, validated on 5,960 knees from the Osteoarthritis Initiative (OAI), achieved an Area Under the ROC Curve (AUC) of 0.93 for radiological diagnosis.

3.1.3. Enhancement through Denoising and Feature Selection

The 2025 literature emphasizes the importance of pre-processing. Bugday et al. (2025) applied a Gaussian filter for denoising followed by feature extraction using DenseNet201. By concatenating original and denoised feature maps and utilizing Neighborhood Component Analysis (NCA) for feature selection, they achieved a high accuracy rate of 85% with a Support Vector Machine (SVM) classifier. This underscores the social potential of deploying AI in diverse clinical settings where image quality may vary.

3.2 Structural Progression and Surgical Outcomes Prediction

The focus has shifted from point-in-time diagnosis to longitudinal forecasting, providing a window for early intervention.

3.2.1 Forecasting Total Knee Replacement (TKR) Tolpadi et al. (2020) developed a pipeline using 3D MRI data and DenseNet-121 to predict the necessity of TKR within an 8-year horizon. The model achieved an AUC of 0.834, but most notably, for patients who were asymptomatic and showed no radiographic sign of OA (KL 0–1) at baseline, the model predicted future TKR with an AUC of 0.943. This suggests AI can identify "pre-radiographic"structural markers in the medial patellar retinaculum and other tissues invisible to the human eye.

3.2.2 Multi-modal Predictive Synergy and the LBTRBC-M Framework

The evolution of prognosis in knee osteoarthritis (KOA) is increasingly defined by the integration of heterogeneous data streams. Results from the most recent literature (2025) indicate that combining longitudinal imaging data with molecular and clinical biomarkers significantly elevates the predictive ceiling for disease progression.

The LBTRBC-M Architecture: Integrating Radiomics and Biomarkers

The state-of-the-art in multimodal forecasting is represented by the Load-Bearing Tissue Radiomic plus Biochemical biomarker and Clinical variable Model (LBTRBC-M) developed by T. Wang et al. (2025). This model addresses the biological reality of KOA as a "whole-joint" disease by processing 1,753 longitudinal knee MRIs alongside systemic indicators.

- **High-Dimensional Feature Selection:** The model initially ingested 2,947 features, including 2,922 MRI radiomic features from load-bearing tissues (femur, tibia, cartilage, and meniscus), 17 biochemical biomarkers, and 7 clinical variables. Through the application of LASSO (Least Absolute Shrinkage and Selection Operator) regression, the feature space was optimized to 255 non-zero indicators that carry the highest prognostic weight.
- **Predictive Performance:** Utilizing the XGBOOST algorithm, the LBTRBC-M achieved high accuracy across multiple progression phenotypes. Specifically, it predicted Joint Space Narrowing (JSN) progression with an AUC of 0.913 and pain progression with an AUC of 0.886. These results outperform earlier models that relied on static, unimodal data, demonstrating that longitudinal changes in tissue texture are critical for identifying "fast-progressors".

The Role of Biochemical Indicators in Decision Support

A key technical finding is the additive value of biochemical markers such as urinary CTX-II and serum COMP. While MRI radiomics capture structural damage, biochemical markers provide insights into the metabolic state of the cartilage.

- **Clinical Value for Residents:** The implementation of LBTRBC-M as a Clinical Decision Support System (CDSS) yielded profound social results. In a controlled "Reader Experiment," the assistance of the AI model improved the prognostic accuracy of resident physicians from a baseline of 44.7%–49.0% to 64.4%–66.5%. This narrowing of the "experience gap" suggests that innovative AI technologies can serve as a potent tool for medical education and the standardization of care in overloaded healthcare systems.

Simplifying Prognosis for Clinical Accessibility

In parallel with high-complexity MRI models, C.-T. Wang et al. (2025) explored the potential of simplified architectures for broader clinical deployment.

- **Vision Transformer (ViT) Approach:** By utilizing a Vision Transformer-based model trained on raw radiographs and essential clinical factors (age, sex, BMI), researchers achieved an AUROC of 0.808 for identifying OA progression in the OAI testing set.
- **Surgical Candidate Identification:** The results demonstrated that even with minimal inputs (single image + essential factors), the model could identify surgical candidates (KLG 3–4) with an AUROC of 0.844. This highlights a critical socio-economic trade-off: while multi-modal MRI models offer the highest precision, simplified radiographic AI tools provide the high-throughput scalability needed for primary care screening and early intervention programs in diverse social settings.

3.3 Symptomatic Progression and Pain Forecasting

A critical finding across the 2022–2026 literature is that structural degradation of the joint does not always correlate linearly with patient-reported outcomes. Consequently, AI research has pivoted toward predicting "symptomatic progression," particularly chronic pain trajectories, which represent the primary social burden of KOA.

Deep Learning for Pain Trajectory Prediction

The research by Guan et al. (2022) addressed the challenge of predicting pain progression using data from 9,348 knees in the OAI cohort.

- **Model Architecture:** Utilizing an EfficientNet architecture, researchers developed a model that integrates baseline radiographs with clinical variables (age, sex, BMI, and baseline WOMAC scores).
- **Performance Metrics:** The combined model achieved an AUC of 0.807, significantly outperforming traditional models that utilized only clinical variables (AUC 0.692).
- **Clinical Insight:** The model demonstrated that AI can identify individuals at high risk for significant clinical decline even among those with low baseline radiographic severity (KL grades 0–1). This finding is pivotal for social healthcare management, as it allows for the early allocation of resources toward pain-management interventions before the onset of permanent disability.

Radiomics-Based Nomograms for Pain Forecasting

Complementing deep learning approaches, Sun et al. (2026) introduced the use of X-ray-based radiomics to predict pain progression.

- **Methodology:** The researchers developed a radiomics-clinical nomogram that incorporates 13 key radiomics features extracted from baseline radiographs.
- **Validation:** In a development and validation study using the FNIH cohort, the nomogram achieved high predictive accuracy, demonstrating that sub-perceptual patterns in bone texture are robust indicators of future symptomatic worsening.
- **Social Implication:** The use of nomograms provides a "visual decision tool" for clinicians, facilitating clearer communication with patients regarding their future risk and the necessity of adherence to non-pharmacological therapies (Sun et al., 2026).

3.4 Natural Language Processing (NLP) and Textual Intelligence in Clinical Records

While radiographic and biochemical markers provide a physiological snapshot of joint degradation, a major frontier identified in the 2025 literature is the extraction of "symptomatic intelligence" from unstructured clinical text. This approach addresses the limitations of traditional imaging, which often fails to capture the early, subjective experiences of patients that precede structural damage.

Mining Unstructured Narratives for Clinical Signatures

The landmark study by Thanyakunsajja et al. (2025) utilized a massive longitudinal dataset of over 11,500 patient records to identify early indicators of knee osteoarthritis (KOA). The core innovation lies in the ability to process "free-text" clinician notes—narratives that are traditionally excluded from large-scale data analysis due to their lack of standardization.

- **Feature Extraction and Pre-processing:** The researchers employed advanced text-mining techniques, including tokenization, stop-word removal, and part-of-speech tagging, to isolate key descriptors of mechanical pain and functional stiffness. By integrating these linguistic features with WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index) scores, the AI could construct a multidimensional profile of the patient's clinical state.
- **Identification of Early Indicators:** The NLP model specifically prioritized phrases related to "intermittent joint locking," "stairs-climbing discomfort," and "weight-bearing sensitivity." These linguistic clusters were found to be highly predictive of a future Kellgren-Lawrence (KL) grade increase, even when current radiographs showed no definitive signs of osteoarthritis.

Comparative Performance of Recurrent Architectures (BiLSTM vs. GRU)

A significant portion of the technical results focuses on the selection of the optimal deep learning architecture for sequential text data. Clinical notes are characterized by temporal dependencies (e.g., the progression of symptoms over several visits), making Recurrent Neural Networks (RNNs) the primary choice for this task.

- **Bi-directional Long Short-Term Memory (BiLSTM):** The results from Thanyakunsajja et al. (2025) demonstrate that the BiLSTM model achieved a superior Area Under the ROC Curve (AUC) of 0.911 and a balanced accuracy of 91.2%. The bi-directional nature of the model allows it to process clinical sentences in both forward and backward directions, capturing the context of symptoms more effectively than standard uni-directional models.
- **Gated Recurrent Unit (GRU):** In comparison, the GRU architecture, while computationally more efficient and faster to train, yielded a slightly lower AUC of 0.897.
- **Model Stability:** Both models significantly outperformed traditional machine learning baselines (such as Support Vector Machines and Logistic Regression), proving that deep learning is necessary to capture the nuanced and often ambiguous language used by clinicians during consultations.

The Predictive Temporal Lead-Time

From a socio-clinical perspective, the most transformative result is the "Lead-Time" provided by NLP analysis.

- **Early Detection Window:** The NLP-driven approach was able to flag patients at high risk of developing radiographic KOA 24 to 36 months before the first definitive signs ($KL \geq 2$) appeared on X-rays.
- **Impact on Social Health Policy:** As emphasized by Shahid et al. (2025) and Ou et al. (2025), this lead-time is critical for the implementation of non-pharmacological social interventions. By identifying "at-risk" individuals two to three years earlier, public health systems can pivot toward preventive physical therapy and weight management, potentially reducing the lifetime social cost of disability and joint replacement surgery.

Scalability and Low-Radiation Screening

A key merit-based argument for NLP technologies is their accessibility. Unlike MRI-based models (Tolpadi et al., 2020), NLP does not require expensive hardware or expose patients to ionizing radiation.

- **Democratic Diagnostics:** This technology leverages existing administrative and clinical data, making it highly suitable for integration into Electronic Health Record (EHR) systems. This supports the "democratization of healthcare" discussed in the review, as it allows for sophisticated screening in community health centers that may lack advanced imaging suites but maintain detailed digital patient records.

3.5 Generative AI: Visualization, Interaction, and Educational Transformation

The emergence of Generative Artificial Intelligence (GenAI), particularly Large Language Models (LLMs) and Diffusion Models, has introduced a new paradigm in patient-centered care and disease visualization. These technologies shift the focus from traditional "black-box" classification toward interactive, highly personalized health communication and visual forecasting.

LLMs in Personalized Patient Education and Self-Management

The study by Du et al. (2025) provides a comprehensive benchmark for the efficacy of GPT-4 in generating personalized self-management guidance for KOA patients. This research, designed as a blinded observational study involving 40 orthopedic experts and 40 patients, reveals a significant advantage of AI over human clinicians in educational delivery.

- **Operational Efficiency:** GPT-4 demonstrated an overwhelming speed advantage, generating guidance at 530.03 words per minute (WPM), compared to the clinician average of 37.29 WPM (P<.001).
- **Readability Metrics:** Utilizing the SMOG (Simple Measure of Gobbledygook) index, the AI-generated advice was found to be more accessible to laypeople, with a mean score of 13.33 compared to the clinicians' 13.81 (P<.001). Improvements were also noted in the Flesch-Kincaid and Gunning Fog scores, indicating higher health literacy compatibility.
- **Quality and Personalization:** Independent experts rated GPT-4 higher than human clinicians across multiple dimensions, including Accuracy (5.31 vs 4.76; P=.05), Personalization (54.32 vs 33.20; P<.001), and Comprehensiveness (51.74 vs 35.26; P<.001). Patients also perceived the AI guidance as safer and more trustworthy (median 61 vs 50; P<.001).

Multimodal LLMs in Radiographic Interpretation

While LLMs excel in text generation, their application in direct image analysis remains a developing field. Tandon et al. (2025) evaluated the capability of GPT-4o to identify knee osteoarthritis from radiographic images.

- **Diagnostic Imbalance:** The model exhibited excellent sensitivity (Recall: 0.950), indicating a high rate of successful OA identification. However, it suffered from extremely poor Specificity (0.114), resulting in a high frequency of false-positive results where healthy joints were categorized as diseased.
- **Clinical Accuracy:** The overall diagnostic accuracy of GPT-4o was measured at 0.532, leading researchers to suggest that while LLMs are effective for preliminary screening or triage assistance, they cannot yet replace specialized diagnostic CNNs for severity grading.

Visual Forecasting via Efficient Diffusion Models

A breakthrough in visual prognostic modeling was presented by Butler et al. (2025), through the use of Efficient Diffusion Models. Unlike standard risk calculators that provide abstract probabilities, this technology synthesizes a visual "future radiograph" of the patient's joint.

- **Forecasting Precision:** The model leverages a class-conditioned latent space to generate high-quality predicted images of the knee joint with a 12-month lead time. It achieved an AUC of 0.71 for predicting future Kellgren-Lawrence (KL) progression.
- **Technical Optimization:** The system operates 9 times faster than current state-of-the-art methods (inference time 2.7s vs 23.6s) and utilizes a significantly more lightweight architecture (35 million vs 215 million parameters).
- **Landmark Localization:** By automatically identifying 16 anatomical landmarks, the model provides an interpretable "map" of future degradation, serving as a powerful behavioral nudge for patient adherence to conservative treatments.

Domain-Specific LLMs and Clinical Support

The systematic review by Ma et al. (2025) further explores the rise of domain-specific variants, such as Knee-ChatGPT. These fine-tuned models demonstrate superior performance in interpreting complex orthopedic terminology and generating structured MRI reports. From a socio-clinical perspective, these systems democratize expert-level information, allowing patients to interact with a "digital assistant" that bridges the health literacy gap.

3.6 Omics and Molecular Discovery: The Frontier of Precision Medicine

The integration of artificial intelligence into the "omics" fields—genomics, transcriptomics, epigenomics, proteomics, and metabolomics—represents a paradigm shift from treating knee osteoarthritis (KOA) as a uniform wear-and-tear disease toward a precision medicine approach. The results from the 2025–2026 literature highlight AI's unique capability to manage the extreme dimensionality and complexity of biological datasets to uncover novel therapeutic targets.

Multi-Omics Data Integration and Feature Extraction

As emphasized by Sharma (2026), the primary challenge in OA research is the "curse of dimensionality," where the number of biological features (genes, proteins, metabolites) vastly exceeds the number of clinical samples.

- **Methodological Innovations:** Researchers are increasingly utilizing unsupervised clustering and deep generative modeling to integrate multi-layered datasets. Liu et al. (2025a) demonstrated that AI-driven integration of single-cell and spatial omics allows for the dissection of molecular complexities within the OA microenvironment, identifying specific cell populations that drive synovial inflammation.
- **Pathway Discovery:** AI models have successfully mapped the interactions between genetic risk signals and epigenetic modifications. This "integrative modeling" provides a holistic view of the disease that was previously unattainable through traditional linear statistical methods (Sharma, 2026).

Identification of Molecular Signatures: The Case of PDK1

A landmark result in molecular discovery was documented by Ou et al. (2025).

- **The PDK1 Signature:** Machine learning algorithms were employed to identify the osteoarthritis signature gene PDK1. The AI-driven analysis revealed that PDK1 plays a critical regulatory role in chondrocyte autophagy and apoptosis (cell death).
- **Mechanistic Insight:** By identifying this specific signaling hub, researchers can now develop targeted pharmacological interventions designed to inhibit PDK1, potentially slowing or reversing the degradation of joint cartilage at the molecular level.
- **Diagnostic Signatures:** Beyond PDK1, Ou et al. (2025) highlighted the discovery of six macrophage-associated genes in the synovium that constitute a novel diagnostic signature for early-stage OA. These molecular markers can be integrated into clinical screening protocols long before structural damage is visible on radiographs.

Spatial Omics and the Cellular Microenvironment

The emergence of spatial transcriptomics, as reviewed by Liu et al. (2025a), allows AI to "map" where specific genes are expressed within the joint tissue.

- **Spatial Heterogeneity:** AI-driven spatial analysis has uncovered that OA is not a globally uniform process but consists of localized "hotspots" of metabolic activity.
- **Precision Medicine Application:** This level of detail enables the development of "digital twins" of the knee joint, as discussed by Joseph et al. (2025), where molecular, imaging, and clinical data are fused to simulate a patient's response to specific drugs or physical therapy regimens.

Socio-Clinical Significance of Molecular AI

The transition to omics-based AI has profound social implications for healthcare systems.

- **Democratizing Precision Medicine:** While omics sequencing was once prohibitively expensive, AI-driven "proxy" models can now predict molecular endotypes using cheaper clinical and radiographic markers.
- **Early Screening:** The ability to identify high-risk molecular signatures from a simple blood or synovial fluid sample—assisted by AI—allows for the stratification of patients into "fast-progressors" and "slow-progressors," ensuring that intensive social and medical resources are directed toward those who need them most (Sharma, 2026; Ou et al., 2025).

3.7 Systematic Benchmarking and Socio-Ethical Governance

As AI technologies transition from research laboratories to clinical workflows, the necessity for standardized benchmarking and ethical oversight has emerged as a dominant theme in the 2025–2026 literature.

Benchmarking Machine Learning vs. Traditional Methods

The systematic meta-analysis by Liu et al. (2025b) involving 32 studies provides the most comprehensive benchmarking of current algorithms.

- **Pooled Performance:** The analysis confirmed that machine learning (ML) consistently outperforms traditional statistical methods in predicting KOA progression. Specifically, integrated models (combining MRI radiomics and clinical data) yielded the highest pooled C-index (0.806).
- **Algorithm Stability:** Random Forest (RF) and XGBoost were identified as the most stable and accurate algorithms for tabular clinical data, while deep learning (DL) remained superior for high-dimensional imaging analysis.

Ethical and Legal Governance Frameworks

The deployment of these innovative technologies in social infrastructure requires a balance between innovation and safety. Pham (2025) and Aldhafeeri (2025) emphasize several governance pillars:

- **Accountability and Liability:** A significant legal challenge identified is the "liability gap" regarding AI-driven misdiagnosis. The literature advocates for "Human-in-the-Loop" (HITL) systems where the AI acts as a decision support tool rather than an autonomous diagnostic agent.
- **Justice and Bias Mitigation:** Pham (2025) warns against algorithmic bias, noting that many models are trained on datasets that lack ethnic and geographic diversity. Ensuring "Justice" in AI requires the diversification of training data to prevent healthcare disparities.
- **Transparency and Trust:** The emergence of Explainable AI (XAI), such as the attention maps introduced by Tiulpin et al. (2018) and further discussed in Aldhafeeri (2025), is critical for building clinician and patient trust. By making the "Black Box" transparent, these technologies align with the regulatory requirements of frameworks like the EU AI Act and GDPR.

4. Discussion

The transition in Knee Osteoarthritis (KOA) management from qualitative observation to quantitative, multi-modal predictive modeling represents a significant leap toward precision orthopedics. The evidence synthesized from 2018–2026 indicates that Artificial Intelligence (AI) is no longer merely a tool for post-facto classification but a predictive engine for longitudinal health management.

4.1. Technical Precision as the Foundation of Clinical Trust

The transition toward fully automated Kellgren-Lawrence (KL) grading is fundamentally predicated on the robustness of the initial joint localization stage. Evidence synthesized from Chen et al. (2019) and Sheik Abdullah and Rajasekaran (2022) confirms that isolating the joint space from high-resolution radiographs (1024×1024 pixels) is not merely a computational convenience but a vital safeguard against "background noise" that typically contaminates full-leg or full-knee images. Specifically, the customization of anchor boxes within the YOLOv2 framework to match the specific anatomical proportions of the femoral and tibial margins allows the system to achieve a recall of 92.2% even under a stringent Jaccard index (IoU) threshold of 0.75. When this threshold is relaxed to 0.50, the localization success rate climbs to 99.9%, ensuring that the subsequent classification layers receive clean, standardized input.

The clinical necessity of this precision is highlighted by the finding that failing to implement a dedicated localization stage leads to a drop in classification accuracy of approximately 5–7%. This decrease is particularly pronounced in "Doubtful" (KL 1) cases, where structural markers such as subtle joint space narrowing or minute osteophytes are easily obscured by irrelevant skeletal features. Furthermore, the stability of feature extraction is significantly enhanced through pre-processing innovations, such as the application of Gaussian filters for denoising and Neighborhood Component Analysis (NCA) for feature selection, which maintain diagnostic consistency across varying image qualities.

Beyond simple localization, the shift toward Deep Siamese CNNs represents a breakthrough in mimicking human clinical logic. By analyzing both compartments of the knee joint simultaneously, these architectures replicate the comparative assessment performed by experienced radiologists. This structural logic, when combined with Adjustable Ordinal Loss functions, allows the AI to respect the biological continuum of osteoarthritis. By penalizing "distant" misclassifications (e.g., mistaking KL 0 for KL 4) more severely than "adjacent" ones (e.g., KL 1 vs. KL 2), models like the fine-tuned VGG-19 achieve a quadratic weighted Kappa of 0.852. This mathematical innovation ensures that the model’s output is not just statistically accurate but clinically reliable, providing a standardized foundation for the socio-ethical governance of AI in orthopedic workflows.

4.2. NLP and Lead-Time: A New Frontier for Early Intervention

Natural Language Processing (NLP) represents one of the most innovative frontiers in the early detection of Knee Osteoarthritis (KOA) by enabling the extraction of "symptomatic intelligence" from unstructured clinical text. As demonstrated by the landmark study by Thanyakunsajja et al. (2025), analyzing a massive longitudinal dataset of over 11,500 patient records allows for the identification of early indicators that traditional imaging often misses in the initial stages. This technology addresses the limitations of standard diagnostics by processing "free-text" clinician narratives—subjective descriptions traditionally excluded from large-scale analysis due to a lack of standardization. By capturing nuanced patient experiences, such as reports of intermittent joint locking, functional stiffness, or specific pain during stair climbing, AI can construct a multidimensional profile of the patient's clinical state long before structural damage becomes visible on a radiograph.

From a technical perspective, the choice of neural architecture is critical for capturing the temporal dependencies found in sequential clinical data. Comparative performance analysis of recurrent architectures revealed a clear superiority of the Bi-directional Long Short-Term Memory (BiLSTM) model, which achieved a superior Area Under the ROC Curve (AUC) of 0.911 and a balanced accuracy of 91.2%, outperforming the Gated Recurrent Unit (GRU) model which yielded an AUC of 0.897. The bi-directional nature of the BiLSTM model is particularly effective as it processes clinical sentences in both forward and backward directions, capturing the full context of symptoms more accurately than standard uni-directional models. When these linguistic features are integrated with validated clinical metrics such as the WOMAC (Western Ontario and McMaster Universities Osteoarthritis Index) scores, the AI provides a robust predictive framework for future Kellgren-Lawrence (KL) grade increases.

The most transformative result for social health policy is the "lead-time" provided by this NLP-driven approach. The ability to flag high-risk patients 24 to 36 months before the first definitive radiographic signs ($KL \geq 2$) appear allows public health systems to pivot toward proactive management. This window of opportunity is vital for implementing non-pharmacological social interventions, such as preventive physical therapy and weight management programs, which can significantly reduce the lifetime social cost of disability and the necessity for joint replacement surgery. Furthermore, NLP supports the "democratization of healthcare" by leveraging existing Electronic Health Record (EHR) data without requiring expensive imaging hardware or exposing patients to ionizing radiation, making sophisticated screening accessible even in community health centers with limited resources.

4.3. Multi-modality vs. Accessibility: The Deployment Dilemma

The evolution of prognosis in knee osteoarthritis is increasingly defined by the integration of heterogeneous data streams, which effectively raises the predictive ceiling for disease progression. The state-of-the-art LBTRBC-M framework exemplifies this by processing 1,753 longitudinal knee MRIs alongside biochemical and clinical indicators. A significant technical achievement of this model was the optimization of a massive feature space—reducing 2,947 initial features to 255 non-zero indicators through LASSO regression—to identify "fast-progressors" with a structural AUC of 0.913. Beyond mere accuracy, the clinical implementation of such multi-modal Clinical Decision Support Systems (CDSS) yields profound educational results. Reader experiments demonstrate that AI assistance can improve the prognostic accuracy of resident physicians from a baseline of 44.7%–49.0% to a much more reliable 64.4%–66.5%. This suggests that high-complexity AI can serve as a potent tool for standardizing care in overloaded healthcare systems by narrowing the experience gap between junior and senior clinicians. However, a socio-economic trade-off remains; while these MRI-based models offer the highest precision, simplified radiographic tools—such as those utilizing Vision Transformers (ViT) with an AUROC of 0.808—provide the high-throughput scalability necessary for primary care screening and early intervention programs in diverse social settings.

4.4. Generative AI and Behavioral Nudging

The emergence of Generative AI (GenAI), particularly Large Language Models (LLMs) and Diffusion Models, has introduced a new paradigm in patient-centered care and disease visualization. Research benchmarking GPT-4 reveals a staggering operational advantage, as the model generates personalized self-management guidance at 530.03 words per minute—over 14 times faster than human clinicians. More importantly, the AI-generated advice consistently achieves higher readability scores (SMOG index 13.33) and superior expert ratings for personalization and comprehensiveness. Parallel to linguistic interaction, the use of Efficient Diffusion Models allows for a visual "future radiograph" synthesis with a 12-month lead time, achieving an AUC of 0.71 for predicting KL progression. These models are not only technically efficient—operating 9 times faster and using significantly fewer parameters (35 million vs. 215 million) than previous state-of-the-art methods—but also provide an interpretable "map" of future degradation by identifying 16 anatomical landmarks. This visualization serves as a powerful behavioral nudge for patient adherence to conservative treatments. Nevertheless, the poor specificity of multimodal models like GPT-4o (0.114) in direct image interpretation serves as a critical warning that GenAI should currently be viewed as an educational and triage assistant rather than a primary diagnostic tool.

4.5. Ethical Governance and Legal Barriers

As AI technologies transition from research laboratories to clinical workflows, the necessity for standardized benchmarking and ethical oversight has become a dominant theme. Systematic meta-analyses confirm that while machine learning (ML) consistently outperforms traditional statistical methods—with integrated models yielding a pooled C-index of 0.806—their social integration requires a balance between innovation and safety. A significant legal challenge is the "liability gap" regarding AI-driven misdiagnosis, leading to a strong advocacy for "Human-in-the-Loop" (HITL) systems where AI acts as a decision support tool rather than an autonomous agent. Furthermore, the risk of algorithmic bias is critical; most models are trained on datasets that lack ethnic and geographic diversity, which could exacerbate healthcare disparities if not addressed through the diversification of training data. Transparency and trust remain the primary pillars for wide-scale adoption, which can be fostered through Explainable AI (XAI) techniques, such as the attention maps originally introduced by Tiulpin et al. (2018). By making the "Black Box" of deep learning transparent, these technologies align with global regulatory requirements such as the EU AI Act and GDPR.

4.6. Limitations and Future Directions

Despite the promising performance of current architectures, the reviewed literature identifies several critical limitations and avenues for future exploration. Most models still require extensive external validation on diverse, real-world populations to ensure that high AUC scores translate across different clinical environments. A major frontier is the integration of AI into "omics" fields to move beyond treating KOA as a uniform wear-and-tear disease. AI-driven discovery of molecular signatures, such as the PDK1 gene and its role in chondrocyte autophagy, offers a path toward true disease-modifying therapies. Future research should prioritize the development of "digital twins" of the knee joint, where molecular, imaging, and clinical data are fused to simulate a patient's response to specific drugs or physical therapy regimens. Additionally, the continued optimization of "lightweight" models remains essential for democratizing access to these technologies, allowing sophisticated screening to be conducted on mobile devices in community health centers that lack advanced imaging infrastructure.

5. Conclusions

The synthesis of research from 2018 to 2026 demonstrates that artificial intelligence has fundamentally altered the trajectory of Knee Osteoarthritis management, transitioning the field from reactive, point-in-time diagnosis toward proactive, longitudinal health forecasting. This review has identified that high-precision localization pipelines, utilizing architectures such as YOLOv2 and Faster R-CNN, provide the necessary technical foundation for isolating joint spaces with accuracy levels exceeding 98%, thereby reducing the diagnostic errors that traditionally hinder early-stage (KL 1) detection.

A primary contribution of recent literature is the emergence of "symptomatic intelligence" through Natural Language Processing. By mining unstructured clinician narratives, NLP models—particularly BiLSTM architectures—offer a predictive lead-time of 24 to 36 months, allowing public health systems to implement non-pharmacological social interventions well before the onset of irreversible structural damage. Furthermore, the integration of multi-modal data streams in frameworks like LBTRBC-M has raised the prognostic ceiling (AUC 0.913), successfully narrowing the "experience gap" for junior clinicians and standardizing care across diverse healthcare settings.

The introduction of Generative AI has further redefined the patient-technology relationship, offering personalized, highly accessible education at speeds far exceeding human capacity. However, as these technologies transition into clinical infrastructure, the necessity for robust socio-ethical governance becomes paramount. Addressing algorithmic bias, establishing clear liability frameworks for automated decisions, and fostering trust through Explainable AI (XAI) are essential prerequisites for the wide-scale adoption of these innovations.

Looking forward, the frontier of KOA management lies in the development of "digital twins" and the integration of AI with omics-based molecular discovery. By fusing longitudinal imaging, biochemical biomarkers, and cellular microenvironment data, the medical community is moving toward a truly precision-based approach that views osteoarthritis not as a uniform disease of aging, but as a manageable condition with specific, treatable phenotypes.

Conflict of interest: The authors declare no conflict of interest.
Funding statement: No external funding was received to perform this review.
Statement of data availability: The data supporting the findings of this study are available within the article's bibliography.

REFERENCES

1. Abdullah, S. S., & Rajasekaran, M. P. (2022). Automatic detection and classification of knee osteoarthritis using deep learning approach. *La Radiologia Medica, 127*(4), 398–406. https://doi.org/10.1007/s11547-022-01476-7
2. Aldhafeeri, F. M. (2025). Governing artificial intelligence in radiology: A systematic review of ethical, legal, and regulatory frameworks. *Diagnostics, 15*(18), 2300. https://doi.org/10.3390/diagnostics15182300
3. Bugday, B., Bingol, H., Yildirim, M., & Alatas, B. (2025). Enhancing knee osteoarthritis detection with AI, image denoising, and optimized classification methods and the importance of physical therapy methods. *PeerJ Computer Science, 10*, e2310. http://dx.doi.org/10.7717/peerj-cs.2766
4. Butler, D., Hilton, A., & Carneiro, G. (2025). Risk estimation of knee osteoarthritis progression via predictive multi-task modelling from efficient diffusion model using X-ray images. *arXiv*. https://doi.org/10.48550/arXiv.2506.14560
5. Chen, P., Gao, L., Shi, X., Allen, K., & Yang, L. (2019). Fully automatic knee osteoarthritis severity grading using deep neural networks with a novel ordinal loss. *Computerized Medical Imaging and Graphics, 75*, 84–92. https://doi.org/10.1016/j.compmedimag.2019.06.002
6. Du, K., Li, A., Zuo, Q.-H., Zhang, C.-Y., Guo, R., Chen, P., ... & Li, S.-M. (2025). Comparing artificial intelligence-generated and clinician-created personalized self-management guidance for patients with knee osteoarthritis: Blinded observational study. *Journal of Medical Internet Research, 27*, e67830. https://doi.org/10.2196/67830
7. Guan, B., Liu, F., Haj-Mirzaian, A., Demehri, S., Samsonov, A., Guermazi, A., & Kijowski, R. (2022). Deep learning approach to predict pain progression in knee osteoarthritis. *Skeletal Radiology, 51*(2), 363–373. https://doi.org/10.1007/s00256-021-03773-0
8. Joseph, G. B., McCulloch, C. E., Nevitt, M. C., Lane, N. E., Majumdar, S., & Link, T. M. (2025). Machine learning models for clinical and structural knee osteoarthritis prediction: Recent advancements and future directions. *Osteoarthritis and Cartilage Open, 7*, 100654. https://doi.org/10.1016/j.ocarto.2025.100654
9. Liu, Y., Molchanov, V., Brass, D., & Yang, T. (2025a). Recent advances in omics and the integration of multi-omics in osteoarthritis research. *Arthritis Research & Therapy, 27*, 100. https://doi.org/10.1186/s13075-025-03563-2
10. Liu, Y., Xiao, G., Zhang, Y., Wang, X., Jia, J., Xie, A., ... & Zhang, K. (2025b). Predictive value of machine learning in knee osteoarthritis progression: Systematic review and meta-analysis. *Journal of Medical Internet Research, 27*, e80430. https://doi.org/10.2196/80430
11. Ma, Z., Liu, Y., Zhang, Z., Chen, R., Fan, H., Cao, X., & Ni, L. (2025). Clinical applications of large language models in knee osteoarthritis: A systematic review. *Frontiers in Medicine, 12*, 1670824. https://doi.org/10.3389/fmed.2025.1670824
12. Ou, J., Zhang, J., Alswadeh, M., Zhu, Z., Tang, J., Sang, H., & Lu, K. (2025). Advancing osteoarthritis research: The role of AI in clinical, imaging and omics fields. *Bone Research, 13*, 48. https://doi.org/10.1038/s41413-025-00423-2
13. Pham, T. (2025). Ethical and legal considerations in healthcare AI: Innovation and policy for safe and fair use. *Royal Society Open Science, 12*(3), 241873. https://doi.org/10.1098/rsos.241873
14. Shahid, S., Wali, A., Javaid, A., Zikria, S., Osman, O., & Rasheed, J. (2025). Potential of AI-based diagnostic grading system for knee osteoarthritis. *Frontiers in Medicine, 12*, 1707588. https://doi.org/10.3389/fmed.2025.1707588
15. Sharma, D. (2026). Artificial intelligence, machine learning and omic data integration in osteoarthritis. *Osteoarthritis and Cartilage, 34*, 311–320. https://doi.org/10.1016/j.joca.2025.10.012
16. Sun, Y., Liu, J., Deng, C., Peng, C., Pan, S., & Liu, X. (2026). Nomograms based on X-ray radiomics for predicting pain progression in knee osteoarthritis using data from the Foundation for the National Institutes of Health: Development and validation study. *JMIR Medical Informatics, 14*, e78338. https://doi.org/10.2196/78338
17. Tandon, M., Chetla, N., Mallepally, A., Zebari, B., Samayamanthula, S., Silva, J., ... & Sukhija, K. (2025). Can artificial intelligence diagnose knee osteoarthritis? *JMIR Biomedical Engineering, 10*, e67481. https://doi.org/10.2196/67481

18. Thanyakunsajja, N., Jitkajornwanich, K., Xu, S., Shin, D., & Charoenporn, P. (2025). Early diagnosis of knee osteoarthritis with a natural language processing-driven approach based on clinician notes: Development and validation study. *JMIR Formative Research, 9*, e64536. https://doi.org/10.2196/64536
19. Tiulpin, A., Thevenot, J., Rahtu, E., Lehenkari, P., & Saarakkala, S. (2018). Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach. *Scientific Reports, 8*, 1727. https://doi.org/10.1038/s41598-018-20132-7
20. Tolpadi, A. A., Lee, J. J., Pedoia, V., & Majumdar, S. (2020). Deep learning predicts total knee replacement from magnetic resonance images. *Scientific Reports, 10*, 6371. https://doi.org/10.1038/s41598-020-63395-9
21. Wang, C.-T., Chang, K.-T., Lai, F., Pao, J.-L., Lin, S.-M., & Chang, C.-H. (2025). Simplifying knee OA prognosis: A deep learning approach using radiographs and minimal clinical inputs. *Diagnostics, 15*(19), 2543. https://doi.org/10.3390/diagnostics15192543
22. Wang, T., Liu, H., Zhao, W., Cao, P., Li, J., Chen, T., ... & Li, S.-M. (2025). Predicting knee osteoarthritis progression using neural network with longitudinal MRI radiomics, and biochemical biomarkers: A modeling study. *PLoS Medicine, 22*(8), e1004665. https://doi.org/10.1371/journal.pmed.1004665