Received: 15 August 2025 | Accepted: 10 October 2025

DOI: 10.1002/jeo2.70549

REVIEW PAPER

Journal of Experimental Orthopaedics WILEY

Artificial intelligence algorithms in orthopaedics: A narrative review of methods and clinical applications

Jamie Rosen[1,2] | **Jemima Russell**[2] | **Prerna Kartik**[3] | **Martinique Vella-Baldacchino**[2]

[1]Warwick Medical School, University of Warwick, Coventry, UK

[2]MSk Lab, Department of Surgery and Cancer, Imperial College London, London, UK

[3]Department of Trauma & Orthopaedics, Kettering General Hospital, Kettering, UK

Correspondence

Jamie Rosen, Warwick Medical School, University of Warwick, Medical School Building, Coventry CV4 7AL, UK.
Email: Jamie.rosen@warwick.ac.uk

Abstract

This narrative review evaluates the role of artificial intelligence (AI) algorithms in orthopaedic surgery and distinguishes itself by explaining how the main algorithmic approaches function and illustrating each with orthopaedic examples. Machine learning methods, including regression, classification and reinforcement learning, have been applied to fracture detection, prediction of revision risk and modelling of outcomes after arthroplasty and sports injury. Deep learning and convolutional neural networks have improved fracture classification, implant surveillance and segmentation of cartilage and meniscal tissue on magnetic resonance imaging. Neural networks such as FracNet and YOLO-based systems demonstrate growing capability in trauma imaging. Natural language processing has automated the extraction of operative and registry data, while large language models are emerging for diagnostic support and education. Generative artificial intelligence (GAI) have produced synthetic musculoskeletal images to expand data sets. Computer vision and image processing underpin robotic-assisted surgery and preoperative planning, and federated learning enables multicentre collaboration while protecting privacy. Each algorithm offers strengths in accuracy, efficiency or scalability, but also carries bias, transparency, computational cost and lack of external validation. This review explores how these algorithms are shaping orthopaedics, highlighting their benefits, limitations and challenges. Rigorous validation, transparent reporting and governance are essential for safe clinical use.

Level of Evidence: N/A.

KEYWORDS

algorithm, artificial intelligence, machine learning, orthopaedics, orthopaedic surgery

---

Abbreviations: AI, artificial intelligence; ANN, artificial neural network; AR, augmented reality; CNNs, convolutional neural networks; CT, computed tomography; CV, computer vision; DL, deep learning; DQN, Deep Q-Networks; FL, federated learning; GAI, generative artificial intelligence; GANS, generative adversarial networks; IP, image processing; LLM, large language models; ML, machine learning; MR, mixed reality; MRI, magnetic resonance imaging; NLP, natural language processing; NN, neural network; ResNets, residual networks; RL, reinforcement learning; RNN, recurrent neural networks; SL, supervised learning; UL, unsupervised learning; VR, virtual reality.

https://doi.org/10.1002/jeo2.70549

INTRODUCTION

The origins of artificial intelligence (AI) date back to the mid-20th century to Alan Turing, a pioneer in computer science who introduced the principle of machine learning (ML) [25]. Turing refers to a test known as the 'Imitation Game', whereby a person needs to determine whether they are communicating with a human or a machine based entirely on written conversation [96]. This paved the way for further study and development in the area of ML. A key development has been the emergence of deep learning (DL) architectures such as Residual Networks (ResNets), which have transformed medical imaging by enabling accurate feature extraction from complex data sets [34].

AI is a broad concept that refers to the capability for computer software to mimic human intelligence [45]. The term AI was formally introduced by Stanford University professor of computer science John McCarthy at Dartmouth in 1956, where he wanted to see if a machine could simulate human intelligence [8]. AI acts as an umbrella for many different subareas including machine learning (ML), deep learning (DL), neural networks (NN), natural language processing (NLP) and large language models (LLMs), computer vision (CV) and image processing (IP), federated learning (FL) and generative artificial intelligence (GAI) [53, 68]. Many of these approaches are already incorporated into orthopaedic and musculoskeletal surgery. For example, NNs are applied to fracture detection, NLP to operative notes and administrative tasks and GAI models to synthetic musculoskeletal imaging [5, 14, 105]. This review seeks to provide an overview of how these algorithms function and illustrate their applications to orthopaedic surgery, while also addressing their benefits, risks and future directions.

METHODS

Research methodology

This review was conducted as a narrative review. No protocol registration or PRISMA reporting framework was applied. A structured literature search was conducted in Ovid MEDLINE and Embase for studies published between 2021 and 2025. This timeframe was chosen to reflect recent advances in AI and the progress in clinical validation that has influenced orthopaedic practice. In addition, key journals in orthopaedics and related fields were manually screened to identify further relevant studies. To provide historical context, earlier influential works were also included. These included the history of AI, the introduction of convolutional neural networks (CNNs), the AlexNet architecture and landmark applications in clinical imaging. Such studies have provided the foundations of current AI methodologies in orthopaedics. Abstracts, preprints and animal studies were excluded.

The full search strategy is provided in the Supporting Information S1: Appendix. This included combinations of AI and algorithms with orthopaedic and related fields (trauma, fracture, arthroplasty, joint replacement), including subfields such as machine learning, deep learning and NLP. Full-text studies in English focusing on algorithm functionality, clinical applications, benefits, risks and implementation challenges in orthopaedics were included. Both original research and review articles were included to ensure a broad overview of the field. All studies were imported into Rayyan: Systematic Review Management Platform. Title, abstract and full-text screening were conducted independently and blindly by two reviewers. Included studies were then extracted into Microsoft Excel and organised thematically according to the type of AI technique. No formal critical appraisal was undertaken.

RESULTS

Key areas of AI

Machine learning (ML)

ML is a foundational subset of AI concerned with developing algorithms that analyse large data sets, perform specific tasks and improve performance through experience without explicit reprogramming [72]. It is often defined as a 'computer programme that learns from experience with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E′ [62].

In orthopaedics, ML applications include automated fracture detection and classifications on radiographs, predicting revision risk following joint replacements and a patient's length of stay in hospital [12, 21, 51]. ML has also been applied to gait and motion analysis for injury and fall risk assessments, optimising rehabilitation strategies and predicting return to sport timelines for athletes [22, 78].

Deep learning (DL)

DL is a specialised subset of ML employing multilayered artificial neural networks (ANNs) to model complex, nonlinear relationships and automatically extract hierarchical features. DL processes unstructured data, with foundational research establishing the theoretical and practical basis for its current use in orthopaedic practice [24, 54].

DL has been applied to radiological imaging including the detection of fractures on X-rays and for identifying spinal stenosis and degenerative changes on magnetic resonance imaging (MRIs) [4, 19]. Beyond fracture detection, it has been applied to pelvic fracture classification using the AO/OTA pelvic fracture classification scale highlighting how DL can support decision-making in orthopaedic trauma [55]. DL models' ability to handle large amounts of complex data enables the segmentation of meniscal and cartilage tissue on MRIs, which can support arthroscopic assessment [13]. Recently, DL models have been used to predict implant loosening by analysing radiographs [47]. This highlights the potential of DL to improve long-term arthroplasty surveillance.

Neural networks (NN)

NNs comprise interconnected input, hidden and output layers that iteratively transform data to identify complex patterns [87]. Orthopaedic NNs have demonstrated promising outcomes in detecting fractures on radiographs with limited data sets [4]. Orthopaedic NN FracNet combines supervised learning (SL) with feature fusion in order to improve fracture recognition accuracy [1].

CNNs are optimised for image analysis, capturing spatial hierarchies through their layered filters. The AlexNet architecture marked a turning point in CV, enabling widespread adoption of deep CNNs in healthcare [49]. CNNs have successfully automated the classification of distal radius fractures on radiographs, performing at the level of an orthopaedic surgeon [71, 75]. CNNs also support paediatric orthopaedic applications with models such as YOLOv8-AM improving fracture detection in children. These examples highlight how CNNs are becoming increasingly capable of enhancing the speed and accuracy of radiological interpretation in orthopaedics [7].

Natural language processing (NLP) and large language models (LLM)

NLP enables computational systems to process, interpret and generate human language, structuring information from unstructured data sources as electronic health records, discharge summaries and consultation transcripts [100, 105]. Applications within orthopaedics include extracting key fields such as operative approach, fixation method and bearing surfaces from postoperative notes following arthroplasty surgery [101]. NLP has also been used to analyse large sets of musculoskeletal radiology reports to identify fracture trends, which reduces the need for clinicians to manually check promoting more efficient research [43].

LLMs build on NLP by using DL and transformer-based architectures to generate contextually relevant, human-like text [79]. They are trained on extensive volumes and varying types of data, including open-access literature and other publicly available sources [63]. LLMs have shown the potential to support orthopaedics through the generation of educational content for both clinicians and patients [81]. They play a growing role in diagnostics and research by managing large data sets and integrating with other AI approaches to enhance accuracy [111]. Their auto-regressive design enables next-token prediction based on probabilistic modelling [106].

Computer vision (CV) and image processing (IP)

CV uses algorithms, often integrated with DL, to interpret and analyse visual data. IP is a CV subfield that employs techniques such as noise reduction, contrast enhancement and segmentation to optimise diagnostic images [23, 49]. CV has been applied to interpret pelvic x-rays and detect fractures with high accuracy and automate AO-OTA fracture classification [29]. This highlights the impact to both diagnostic and preoperative planning. IP techniques such as photon-counting computed tomography (CT) with virtual monoenergetic reconstruction reduce metal artefacts around implants, providing clinicians with better visibility of bone and soft tissue structures [86]. IP is a key feature of the Mako surgical robot which uses a preoperative CT scan and three-dimensional (3D) model reconstruction to create a 3D patient-specific plan for surgery [27].

Federated learning (FL)

FL enables collaborative model training across institutions without transferring raw patient data, thereby preserving privacy and addressing governance concerns [94, 109]. FL has been explored in spine surgery, where it shows promise for robotic telesurgery, vertebral image segmentation, privacy-preserving predictive modelling and enabling multicentre collaboration without sharing sensitive patient data [88].

Generative artificial intelligence (GAI)

GAI creates new content such as audio, code, images, text and videos based on learned data patterns [50]. GAI has demonstrated its ability to generate orthopaedic discharge documents more efficiently than clinicians whilst maintaining comparable writing quality [82]. GAI has also been used to generate synthetic radiographs and CT scans of spinal fractures [14, 20, 58]. These methods can expand training data sets and enhance diagnostic models without compromising patient confidentiality. However, outputs require rigorous expert validation before clinical application [14].

Key AI approaches

Machine learning

ML models can be split into three main types of models: supervised, unsupervised and reinforcement learning. SL is a technique where labelled data sets are used to train algorithmic models in order to predict a certain outcome, such as the presence of a disease or tumour [104]. Unsupervised learning (UL) is where unlabelled data sets are used to train models [26]. Reinforcement learning (RL) is where the algorithms learn based on trial and error and receive feedback, which allows them to learn [3]. ML Models will learn using a combination of SL, UL and RL using both human input and large data sets to ensure that models can handle new tasks [33]. Table 1 provides an overview of the mechanisms of common ML algorithms, and Table 2 presents the associated benefits and risks of their application in orthopaedics.

Neural networks

NN act as an umbrella for different types of NN models. These include but are not limited to CNNs, generative adversarial networks (GANs) and recurrent neural networks (RNNs). CNNs use the mathematical operation of convolution to recognise different patterns in images by analysing key features such as shape, colour or objects [36]. GANs are made up of two parts, with one part generating a fake image or other types of data and the other part determining if it is fake or not [2]. These two parts compete in order to make the best fake. RNNs process data and remember previous data they processed, which allows them to recognise complex patterns from sources such as text or speech [60]. Table 3 summarises the types of neural networks, their applications in orthopaedics and associated benefits and risks.

Benefits of AI in orthopaedics

The implementation of AI has already started to transform orthopaedic care. From imaging interpretation and preoperative planning to conducting administrative tasks, AI has been deployed across the surgical pathway. Recent literature highlights AI's potential to enhance diagnostic accuracy,

TABLE 1 Overview of common machine learning algorithms and their mechanism.

Algorithm	Type	Mechanism
Regression • Linear • Lasso • Random forest • Gradient boosting	SL	Continuously predicts values based on input data by establishing relationships between variables such as *X* and *Y*.
Classification • Logistic • KNN • SVM	SL	Assign data to specific categories by analysing their characteristics and features.
Naïve Bayes classifiers: • Decision trees	SL	Applies Bayes' mathematical theorem of probability to predict a data category and sort items into groups with the highest probabilities [3].
K-means clustering	SL	Separates data into groups that are most related or similar and assigns them accordingly.
Hierarchical clustering	UL	Creates a tree of different groups and iteratively grouping and assigning data based on similarity [69, 92].
Probabilistic clustering	UL	Assigns data to groups by estimating the likelihood of the data belonging to a particular group.
Q-Learning	RL	Uses an agent (computer software) that learns by updating values in a table to establish which actions yield the best results.
DQN	RL	Works like the Q-Learning algorithm but uses a NN instead of a table to establish which actions give the best results, so it can handle larger and more complex data sets.

Abbreviations: DQN, Deep Q-Networks; KNN, K-nearest neighbours; NN, neural network; RL, reinforcement learning; SL, supervised learning; SVM, support-vector machines; UL, unsupervised learning.

TABLE 2 Summary of machine learning algorithms with associated benefits and risks in orthopaedic applications.

Algorithm	Type	Benefits	Risks
Regression: • Linear • Lasso • Random forest • Gradient boosting	SL	Quick to train, it is widely used in practice and is effective at predicting numerical values [85].	Cannot function with complex data or patterns and can be thrown off by unexpected data. Assumes a straight-line relationship between variables [85].
Classification: • Logistic • KNN • SVM	SL	Works well with different data types, can process complex patterns/relationships and is commonly implemented [85].	High computational cost, struggle interpreting complex models such as DL and can produce biased results if training data is biased [59, 85].
Naïve Bayes classifiers: • Decision trees	SL	Simple to implement, does not require significant amounts of data to be trained and can handle different data types as input [15].	Can make inaccurate predictions if trained on incorrect data, struggles to compute new data and can assign incorrect values to this, and treats all data as separate even if there is a clear relationship [15].
K-means clustering	UL	Simple to implement, it is fast and can compute relatively large data sets due to its low computational requirements; it is easy to interpret and is used widely across many fields [28, 38, 40].	Sensitive to initial values resulting in suboptimal grouping of data, a few unexpected values can skew results, and users can struggle to group data that forms nonspherical or irregular shapes [28, 38, 40].
Hierarchical clustering	UL	Reveals a clear hierarchical structure which is easy to interpret, and there is no preset number of clusters, so you can decide the number of groups later on making analysis of results easier [77, 110].	Large computational requirements, not scalable for very large data sets, and the outcome can vary based on parameters chosen, potentially leading to different cluster structures [56, 77, 110].
Probabilistic clustering	UL	Capable of analysing groups even if in an irregular shape and can handle data or values that are clear or have errors in [28].	Slow when working with large volumes of complex data, and results can be sensitive to initial values [28].
Q-Learning	RL	It can function with limited initial data and is useful for uncertain situations [80].	Can overestimate values, leading to inaccurate outcomes and poor decision-making, and can be slow when working with large and complex data sets [80].
DQN	RL	Useful in complex data analysis and recognising complex patterns [76].	Requires large data sets, requires significant computing power and training time, and can sometimes overestimate values, leading to inaccuracies [76].

Abbreviations: DL, deep learning; DQN, Deep Q-Networks; KNN, K-nearest neighbours; RL, reinforcement learning; SL, supervised learning; UL, unsupervised learning.

alleviate clinician workload and promote the use of predictive analytics for earlier intervention [32, 57].

Improved diagnostic accuracy

AI has demonstrated diagnostic accuracy that in some orthopaedic imaging tasks exceeds or performs comparably to that of clinical experts. Performance in AI models is often measured using the 'F1 score'. This metric is the harmonic mean of precision and recall and ranges from 0 (poor) to 1 (perfect) [73]. In paediatric distal radius fracture detection, CNNs outperformed experienced radiologists and orthopaedic surgeons. They achieved F1 scores of 0.90–0.93 compared with 0.77–0.86 for clinicians and area under the curves (AUCs) of 0.95–0.98 versus 0.94 for clinicians [42]. Similarly, in a study on hip fracture detection using anteroposterior pelvic radiographs, a DL ensemble achieved an F1 score of 0.942, surpassing orthopaedic surgeons (0.938) and radiologists (0.677). The model also reached a sensitivity of 0.97 and a specificity of 0.915 [11]. Whilst AI models do not consistently surpass orthopaedic surgeons in fracture detection, they often outperform nonorthopaedic clinicians [17]. This suggests there might be potential value of AI as a training tool for residents and other clinicians. A recent systematic review and meta-analysis of hip fracture detection found AI accuracy to be inferior to musculoskeletal specialists, highlighting the need for further research and model training before clinical deployment [67].

TABLE 3 Summary of types, applications in orthopaedics, benefits and risks of NNs.

Type of NN	Application in orthopaedics	Benefits	Risks
CNNs	Detection and classification of fractures, including distal radius, proximal humerus, hip, pelvis, femur and spine [39, 70]. Automated identification of hip and knee arthroplasty implants and prediction of prosthetic loosening [9, 30, 47].	Self learns what to look for when extracting, so it does not need to be told everything and is effective at pattern recognition in images and objects [6, 112].	Needs a lot of examples to be proficient and can struggle to analyse images different from what it has been trained on [6, 112].
GANs	Generating synthetic musculoskeletal radiographs to validate ML algorithms and train CAD systems [46, 74].	Can generate synthetic content which looks real, which can be useful for educational purposes, and content does not need to be labelled as GAN can determine labels [91, 98]. Protect patient privacy due to the synthetic nature of the images being generated [14, 46].	May keep generating the same content instead of new variations and can be hard to train as it can be unpredictable [91, 98].
RNNs	Gait analysis and fall prediction based on longitudinal balance monitoring [108]. Analysis of biometric and motion data in order to identify injury patterns in athletes [84].	Can handle data of varying lengths and sequences and can understand data presented in an order, making it ideal for handling text, speech or time-based data [60, 99].	Requires significant computing power and time to train and has a poor memory, so struggles to remember data past a certain timeframe [60, 99].

Abbreviations: CAD, computer-aided diagnosis; CNNs, convolutional neural networks; GANs, generative adversarial networks; ML, machine learning; NNs, neural networks; RNNs, recurrent neural networks.

Alleviates strain on clinicians and optimises resource use

Across the orthopaedic surgical pathway, AI reduces repetitive tasks that increase both clinical and administrative workloads. AI support improves fracture interpretation with a 29% reduction in missed fractures and 21% fewer false positives, without increasing reading time [10]. A systematic review and meta-analysis of commercial fracture-AI tools reported time savings of approximately 6–12 s per radiograph when AI was used to support a clinician [37]. Beyond imaging, NLP extracted key fields from postoperative total hip arthroplasty (THA) notes with 100% accuracy, enabling automated registry updates rather than manual entry [101]. ChatGPT also produced discharge documents for orthopaedic patients almost 10 times faster than physicians with comparable quality, highlighting AI's ability to reduce administrative workloads [82].

Predictive analytics and early intervention

AI's ability to analyse large complex sets of data and accurately predict patient outcomes enables earlier interventions, highlighting its advanced diagnostic capabilities [51]. ML applied to electronic health records improves fragility fracture prediction by providing a more personalised risk assessment than traditional tools like FRAX [35]. ML models predicting outcomes after total knee arthroplasty achieved AUCs of 0.79–0.84 accurately forecasting complications and discharge status [64]. Beyond clinical outcomes, models predicted the theatre duration of hip and knee arthroplasty with mean absolute errors of 11–13 min, enhancing theatre scheduling and resource allocation [52]. These advances demonstrate the potential of AI to promote earlier intervention and to optimise patient care in orthopaedics.

Risks of AI in orthopaedics

While the adoption of AI in orthopaedics offers clear benefits, there are also significant risks that need to be addressed. These include bias arising from nonrepresentative training data sets, the 'black box phenomenon' that limits clinical insight into model decision-making and significant privacy concerns related to the use of sensitive patient data.

Bias

There is a saying amongst computer scientists that an AI model is only as good as its data. If AI models are trained on nonrepresentative data sets, such as radiographs from only one patient demographic, this bias will be reflected in their outputs [16]. In orthopaedics, this can manifest as poorer accuracy in detecting paediatric fractures when algorithms are developed primarily from adult data sets. Systematic reviews have noted limited external validation and restricted reporting of demographic diversity [89]. Research has demonstrated that bias in hip and knee image segmentation can arise from imbalances in age, body size, ethnicity and sex. This reduces the generalisability of models across diverse patient populations [90]. Therefore, it is pivotal that representative, balanced data sets are used when training orthopaedic AI models to ensure equity of care.

Transparency (black box phenomenon)

The 'black box phenomenon' refers to the difficulty users face in understanding how AI models generate their outcomes [32]. This is particularly important for tasks such as fracture detection and outcome prediction, where surgeons must be able to evaluate why a model makes a specific decision. However, the black-box nature of many models continues to hinder interpretability and limit clinical confidence [61]. A practical guide to AI implementation in orthopaedics has stressed that black-box models present a major barrier to safe clinical translation, emphasising the need for explainability in high-risk surgical settings [114]. Models predicting long-term outcomes after THA demonstrate transparency by highlighting the clinical factors that most strongly influence predictions, making them easier for clinicians to understand [41]. Finally, recent systematic reviews recommend wider adoption of explainable AI methods to mitigate black-box concerns and support safe human-centred deployment [65].

Privacy and autonomy

The use of AI in orthopaedics requires the handling and processing of large amounts of sensitive patient data which creates a significant privacy risk [48]. Despite data being anonymised, breaches or association with external data sets can still enable reidentification [107]. Respect for patient autonomy also requires transparent disclosure and explicit consent when orthopaedic data is used for AI development. This is especially important given the opacity of many algorithms [44]. Decentralised approaches such as swarm learning have been successfully applied to fracture diagnosis, enabling multicentre collaboration while keeping data local and thereby reducing the risk of breaches [102].

DISCUSSION

AI is reshaping orthopaedic workflows, with the most rapid progress in imaging interpretation, operative planning and outcome prediction. In fracture detection and arthroplasty imaging, models have demonstrated performance comparable to that of subspecialist clinicians. However, findings remain inconsistent, reflecting differences in data sets, fracture types and modelling approaches [29, 47, 57, 67]. ChatGPT-4o recently showed strong binary discrimination in knee osteoarthritis radiographs ($F1 = 0.88$) but poor grading accuracy (AI [35% correct] vs. surgeons [89.6% correct]), highlighting inconsistencies in AI radiographic classification [113]. Progress will depend on the use and validation of representative orthopaedic training data. Recent reporting frameworks TRIPOD+AI, PROBAST+AI and DECIDE−AI can help ensure such studies are transparent, reproducible and clinically meaningful [18, 66, 97].

GAI and LLMs are beginning to show utility in orthopaedics, particularly for documentation, registry curation and patient-facing education. Pilot studies suggest that LLMs can generate discharge summaries substantially faster than clinicians while maintaining comparable quality [82]. NLP has also been validated for extracting key operative details from arthroplasty notes with very high accuracy, supporting registry automation [101]. LLMs have further demonstrated value in improving the readability of orthopaedic patient education materials without compromising accuracy [81]. A recent evaluation of ChatGPT-4o and DeepSeek R1 for ACL surgery patient education found both models were highly accurate (3.9/4) and consistent (4/4) on a 4-point Likert scale. ChatGPT gave more comprehensive answers (4.0 vs. 3.2, $p < 0.001$), while DeepSeek provided clearer layperson-focused responses (3.9 vs. 3.0, $p < 0.001$), illustrating variability in LLM outputs for patient education [31].

A recent systematic review found that most orthopaedic LLM studies report efficiency gains, but clinical performance remains highly variable, with diagnostic accuracy ranging from 55% to 93% [111]. While these tools may enhance workflow and patient communication, their integration into high-stakes applications such as implant selection or intraoperative guidance will depend on explainability, transparency and close clinician oversight [41, 114].

Regulatory frameworks are now an important factor in how orthopaedic AI can be implemented. The EU AI Act (2024) introduces a risk-based approach. It places obligations on high-risk medical AI systems and on certain general-purpose models. These measures are directly relevant to decision support, imaging and robotics [95]. Platforms that generate patient-specific surgical plans demonstrate the clinical utility of AI, but they also highlight the need to clarify accountability at each stage of the surgical pathway [27]. Federated learning offers another development. It enables multi-centre collaboration in spine and fracture research while preserving privacy and avoiding direct data sharing [88]. To strengthen reporting of diagnostic accuracy studies involving AI, the new STARD-AI guideline provides items specific to ML applications [93].

Economic and workflow impact are important factors for AI adoption. ML models have been used to predict theatre duration, length of stay and implant survivorship, showing potential for resource optimisation [52]. In fracture detection, AI support reduces missed fractures and false positives without increasing reading time. Commercial tools have also shown modest time savings per radiograph, which may scale to service-level efficiency gains [10, 37]. NLP has demonstrated high accuracy in extracting registry variables from arthroplasty notes, suggesting a role in reducing manual data entry [101]. Opportunities also exist in education and training, such as simulation and synthetic data sets, but these remain underevaluated in prospective orthopaedic settings. One study found ChatGpT-4o achieved a mean score of 70.7% compared to candidates 58% on the Turkish orthopaedics and traumatology board exam. ChatGPT scored much higher in basic science (93% vs. 55%) but only slightly better in limb reconstruction (70% vs. 66%) [103]. LLMs appear capable of competitive exam-level performance in certain domains but not consistently throughout orthopaedic subspecialties.

Recent evidence suggests that rigorous validation, transparent reporting and lifecycle governance are as important as algorithmic advances for safe integration of AI in orthopaedics. Collaboration between engineers and surgeons, evaluation against workflow-relevant outcomes and postdeployment monitoring should become minimum standards. Equally important is preparing both surgeons and patients. Surgeons will need practical training in how to use and interpret these tools safely. Patients should receive clear explanations of how these tools function and how their data is handled. A recent study of 360 orthopaedic surgeons found that most expect AI to substantially impact practice within 5–10 years, predominantly as a complementary tool rather than a replacement [83]. Trust in AI will depend on validation, transparency and governance.

Limitations

This review has several limitations. This review was restricted to English-language studies published between 2021 and 2025. While selected seminal works published before this period were included to provide historical and methodological context, it remains possible that other relevant earlier studies were not captured. Language or publication bias may also have been introduced. Formal bias assessment tools were not used which limits evaluation of study quality. The rapidly evolving nature of AI means that some findings may become outdated. Finally, the search was limited to two databases, so additional relevant studies may not have been captured.

Future research directions

Future research in the use of AI in orthopaedics needs to go beyond retrospective studies. There is a lack of solid prospective research that directly compares the performance of AI models with surgeons, especially in the areas of fracture detection and outcome prediction. As most of the studies included in this paper were conducted at single centres, their findings are unlikely to be representative of broader hospital settings or diverse patient populations. Integrating AI with trauma and implant registries could enhance long-term tracking. Economic evaluations, such as the impact on theatre time or hospital stays, are rare but essential for judging real-world value. Simulation and synthetic data sets also show promise for training, though they have not yet been properly tested in real clinical settings.

CONCLUSION

AI is reshaping orthopaedics. It not only improves imaging interpretation and predictive modelling but also helps ease clinical and administrative burdens. While studies have shown better diagnostic accuracy and more efficient workflows, this progress is not without challenges. Bias remains a concern. Transparency is often limited, and stronger rules and oversight are needed for safe implementation. AI should be viewed as a tool to complement a surgeon and support decision-making when integrated appropriately. With proper testing, responsible rollout and collaboration between experts, it has the potential to raise the standard of patient care in orthopaedics and beyond. Achieving this will require ongoing evaluation and shared accountability.

AUTHOR CONTRIBUTIONS

Jamie Rosen, Jemima Russell, Prerna Kartik and Martinique Vella-Baldacchino each made substantial contributions to this work, including study design, data collection, analysis and interpretation. Jamie Rosen drafted the initial manuscript, and all authors (Jamie Rosen, Jemima Russell, Prerna Kartik and Martinique Vella-Baldacchino) were involved in revising the manuscript and gave final approval for the version to be published. All authors had access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

ACKNOWLEDGEMENTS

The authors have no funding to report.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ETHICS STATEMENT

The authors have nothing to report.

ORCID

Jamie Rosen https://orcid.org/0009-0005-7505-6871
Jemima Russell https://orcid.org/0009-0009-0014-4390
Prerna Kartik https://orcid.org/0000-0003-1855-8881
Martinique Vella-Baldacchino https://orcid.org/0000-0002-3284-7236

REFERENCES

1. A. Alwzwazy H, Alzubaidi L, Zhao Z, Gu Y. FracNet: an end-to-end deep learning framework for bone fracture detection. Pattern Recognit Lett. 2025;190:1–7.
2. Alajaji SA, Khoury ZH, Elgharib M, Saeed M, Ahmed ARH, Khan MB, et al. Generative adversarial networks in digital histopathology: current applications, limitations, ethical considerations, and future directions. Mod Pathol. 2024;37:100369.
3. Alanazi A. Using machine learning for healthcare challenges and opportunities. Inform Med Unlocked. 2022;30:100924.
4. Aldhyani T, Ahmed ZAT, Alsharbi BM, Ahmad S, Al-Adhaileh MH, Kamal AH, et al. Diagnosis and detection of bone fracture in radiographic images using deep learning approaches. Front Med (Lausanne). 2025;11:1506686.
5. Alsoof D, McDonald CL, Kuris EO, Daniels AH. Machine learning for the orthopaedic surgeon. J Bone Joint Surg. 2022;104:1586–94.
6. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021;8(1):53.
7. Amirouche F, Prosper AM, Mzeihem M. Enhancing pediatric distal radius fracture detection: optimizing YOLOv8 with advanced AI and machine learning techniques. BMC Med Imaging. 2025;25:316.
8. Andresen SL. John McCarthy: father of AI. IEEE Intell Syst. 2002;17:84–5.
9. Antonson NW, Buckner BC, Konigsberg BS, Hartman CW, Garvin KL, Kildow BJ. Novel technique for the identification of hip implants using artificial intelligence. J Arthroplasty. 2024;39:1178–83.
10. Bachmann R, Gunes G, Hangaard S, Nexmann A, Lisouski P, Boesen M, et al. Improving traumatic fracture detection on radiographs with artificial intelligence support: a multi-reader study. BJR|Open. 2023;6:tzae011.
11. Beyaz S, Yayli SB, Kılıç E, Kılıç K. Comparison of artificial intelligence algorithm for the diagnosis of hip fracture on plain radiography with decision-making physicians: a validation study. Acta Orthop Traumatol Turc. 2024;58:4–9.
12. Blackman B, Vivekanantha P, Mughal R, Pareek A, Bozzo A, Samuelsson K, et al. Machine learning models predicting risk of revision or secondary knee injury after anterior cruciate ligament reconstruction demonstrate variable discriminatory and accuracy performance: a systematic review. BMC Musculoskelet Disord. 2025;26:16.
13. Botnari A, Kadar M, Puia DR, Patrascu JM, Patrascu Jr, JM. Automated segmentation of knee menisci using U-net deep learning model: preliminary results. Maedica. 2024;19:690.
14. Carroll AN, Storms LA, Malempati C, Shanavas RV, Badarudeen S. Generative artificial intelligence and prompt engineering: a primer for orthopaedic surgeons. JBJS Rev. 2024;12:00122.
15. Chen H, Hu S, Hua R, Zhao X. Improved naive Bayes classification algorithm for traffic risk management. EURASIP J Adv Signal Process. 2021;2021:30.
16. Chen K, Stotter C, Klestil T, Nehrer S. Artificial intelligence in orthopedic radiography analysis: a narrative review. Diagnostics. 2022;12:2235.
17. Christoforides E, Rust B, Mulroy D, Oswald S, Muralidhar R. Artificial intelligence vs. physician expertise in appendicular skeleton fracture detection: a scoping review. J Am Osteopath Acad Orthop. 2024;5(1):ZP987.
18. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378.
19. Cui Y, Zhu J, Duan Z, Liao Z, Wang S, Liu W. Artificial intelligence in spinal imaging: current status and future directions. Int J Environ Res Public Health. 2022;19:11708.
20. D N S, Pai RM, Bhat SN, Pai M M M. Assessment of perceived realism in AI-generated synthetic spine fracture CT images. Technol Health Care. 2025;33:931–44.
21. Devana SK, Shah AA, Lee C, Roney AR, van der Schaar M, SooHoo NF. A novel, potentially universal machine learning algorithm to predict complications in total knee arthroplasty. Arthroplast Today. 2021;10:135–43.
22. Dindorf C, Horst F, Slijepčević D, Dumphart B, Dully J, Zeppelzauer M, et al. Machine learning in biomechanics: key applications and limitations in walking, running and sports movements. In: Blondin MJ, Fister Jr. I, Pardalos PM, editors. Artificial intelligence, optimization, and data sciences in sports [Internet]. Cham: Springer Nature Switzerland; 2025. p. 91–148. Available from: https://doi.org/10.1007/978-3-031-76047-1_4
23. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, et al. Deep learning-enabled medical computer vision. npj Digit Med. 2021;4:5.
24. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115–8.
25. Farhadi F, Barnes MR, Sugito HR, Sin JM, Henderson ER, Levy JJ. Applications of artificial intelligence in orthopaedic surgery. Front Med Technol. 2022;4:995526.
26. Fiste O, Liontos M, Zagouri F, Stamatakos G, Dimopoulos MA. Machine learning applications in gynecological cancer: a critical review. Crit Rev Oncol Hematol. 2022;179:103808.
27. Fontalis A, Kayani B, Thompson JW, Plastow R, Haddad FS. Robotic total hip arthroplasty: past, present and future. Orthop Trauma. 2022;36:6–13.
28. Gao CX, Dwyer D, Zhu Y, Smith CL, Du L, Filia KM, et al. An overview of clustering methods with guidelines for application in mental health research. Psychiatry Res. 2023;327:115265.
29. Gao Y, Soh NYT, Liu N, Lim G, Ting D, Cheng LT-E, et al. Application of a deep learning algorithm in the detection of hip fractures. iScience. 2023;26:107350.
30. Gong Z, Fu Y, He M, Fu X. Automated identification of hip arthroplasty implants using artificial intelligence. Sci Rep. 2022;12:12179.
31. Gültekin O, Inoue J, Yilmaz B, Cerci MH, Kilinc BE, Yilmaz H, et al. Evaluating DeepResearch and DeepThink in anterior cruciate ligament surgery patient education: ChatGPT-4o excels in comprehensiveness, DeepSeek R1 leads in clarity and

readability of orthopaedic information. Knee Surg Sports Traumatol Arthrosc. 2025;33:3025–31.

32. Han F, Huang X, Wang X, Chen YF, Lu C, Li S, et al. Artificial intelligence in orthopedic surgery: current applications, challenges, and future directions. MedComm. 2025; 6:e70260.
33. Harrer S. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. EBioMedicine. 2023;90:104512.
34. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR); 2016 Jun 27–30; Las Vegas, NV, USA. IEEE; 2016. p. 770–8. https://doi.org/10.48550/arXiv.1512.03385
35. Hong N, Whittier DE, Glüer C-C, Leslie WD. The potential role for artificial intelligence in fracture risk prediction. Lancet Diabetes Endocrinol. 2024;12:596–600.
36. Hu W, Li X, Li C, Li R, Jiang T, Sun H, et al. A state-of-the-art survey of artificial neural networks for whole-slide image analysis: from popular convolutional neural networks to potential visual transformers. Comput Biol Med. 2023; 161:107034.
37. Husarek J, Hess S, Razaeian S, Ruder TD, Sehmisch S, Müller M, et al. Artificial intelligence in commercial fracture detection products: a systematic review and meta-analysis of diagnostic test accuracy. Sci Rep. 2024;14:23053.
38. Ikotun AM, Ezugwu AE, Abualigah L, Abuhaija B, Heming J. K-means clustering algorithms: a comprehensive review, variants analysis, and advances in the era of big data. Inf Sci (N Y). 2023;622:178–210.
39. Inoue T, Maki S, Furuya T, Mikami Y, Mizutani M, Takada I, et al. Automated fracture screening using an object detection algorithm on whole-body trauma computed tomography. Sci Rep. 2022;12:16549.
40. Jahangoshai Rezaee M, Eshkevari M, Saberi M, Hussain O. GBK-means clustering algorithm: an improvement to the K-means algorithm based on the bargaining game. Knowl Based Syst. 2021;213:106672.
41. Jang SJ, Fontana MA, Kunze KN, Anderson CG, Sculco TP, Mayman DJ, et al. An interpretable machine learning model for predicting 10-year total hip arthroplasty risk. J Arthroplasty. 2023;38:S44–50.e6.
42. Janisch M, Apfaltrer G, Hržić F, Castellani C, Mittl B, Singer G, et al. Pediatric radius torus fractures in x-rays—how computer vision could render lateral projections obsolete. Front Pediatr. 2022;10:1005099.
43. Jungmann F, Kämpgen B, Hahn F, Wagner D, Mildenberger P, Düber C, et al. Natural language processing of radiology reports to investigate the effects of the COVID-19 pandemic on the incidence and age distribution of fractures. Skeletal Radiol. 2022;51:375–80.
44. Kaya Bicer E, Fangerau H, Sur H. Artifical intelligence use in orthopedics: an ethical point of view. EFORT Open Rev. 2023; 8:592–6.
45. Khoriati A-A, Shahid Z, Fok M, Frank RM, Voss A, D'Hooghe P, et al. Artificial intelligence and the orthopaedic surgeon: a review of the literature and potential applications for future practice: current concepts. J ISAKOS. 2024;9: 227–33.
46. Khosravi B, Rouzrokh P, Mickley JP, Faghani S, Larson AN, Garner HW, et al. Creating high fidelity synthetic pelvis radiographs using generative adversarial networks: unlocking the potential of deep learning models without patient privacy concerns. J Arthroplasty. 2023;38:2037–43.e1.
47. Kim M-S, Kim J-J, Kang K-H, Lee J-H, In Y. Detection of prosthetic loosening in hip and knee arthroplasty using machine learning: a systematic review and meta-analysis. Medicina. 2023;59:782.
48. Kiwinda LV, Kocher SD, Bryniarski AR, Pean CA. Bioethical considerations of deploying artificial intelligence in clinical orthopedic settings: a narrative review. HSS J. 2025;21: 274–82.
49. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60:84–90.
50. Kunze KN. Generative artificial intelligence and musculoskeletal health care. HSS J. 2025;21:248–56.
51. Kutbi M. Artificial intelligence-based applications for bone fracture detection using medical images: a systematic review. Diagnostics. 2024;14:1879.
52. Langenberger B, Schrednitzki D, Halder A, Busse R, Pross C. Leveraging machine learning for duration of surgery prediction in knee and hip arthroplasty—a development and validation study. BMC Med Inform Decis Mak. 2025;25:106.
53. Leal JA. Artificial intelligence in orthopaedic and trauma surgery education: applications, ethics, and future perspectives. J Am Acad Orthop Surg Glob Res Rev. 2025;9:e25.00174.
54. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521:436–44.
55. Lee SH, Jeon J, Lee GJ, Park JY, Kim YJ, Kim KG. Automated Association for Osteosynthesis Foundation and Orthopedic Trauma Association classification of pelvic fractures on pelvic radiographs using deep learning. Sci Rep. 2024;14:20548.
56. Loftus TJ, Shickel B, Balch JA, Tighe PJ, Abbott KL, Fazzone B, et al. Phenotype clustering in health care: a narrative review for clinicians. Front Artif Intell. 2022;5:842306.
57. Longo UG, Lalli A, Nicodemi G, Pisani MG, De Sire A, D'Hooghe P, et al. Artificial intelligence demonstrates potential to enhance orthopaedic imaging across multiple modalities: a systematic review. J Exp Orthop. 2025;12:e70259.
58. McNulty JR, Kho L, Case AL, Slater D, Abzug JM, Russell SA. Synthetic medical imaging generation with generative adversarial networks for plain radiographs. Appl Sci. 2024;14:6831.
59. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. ACM Comput Surv. 2022;54:1–35.
60. Mienye ID, Swart TG, Obaido G. Recurrent neural networks: a comprehensive review of architectures, variants, and applications. Information 2024. 2024;15:517.
61. Misir A, Yuce A. AI in orthopedic research: a comprehensive review. J Orthop Res. 2025;43:1508–27.
62. Mitchell TM. The discipline of machine learning [Internet]. Vol. 9. Carnegie Mellon University, School of Computer Science, Machine Learning; 2006. p. 1–7. Available from: https://www.cs.cmu.edu/afs/cs/usr/mitchell/ftp/pubs/MachineLearningTR.pdf
63. Mo K, Lin R, Dunn E, Girgis G, Fang W, Walsh J, et al. Systematic review on large language models in orthopaedic surgery. J Clin Med. 2025;14:5876.
64. Mohammed H, Huang Y, Memtsoudis S, Parks M, Huang Y, Ma Y. Utilization of machine learning methods for predicting surgical outcomes after total knee arthroplasty. PLoS One. 2022;17:e0263897.
65. Mohammed TJ, Xinying C, Alnoor A, Khaw KW, Albahri AS, Teoh WL, et al. A systematic review of artificial intelligence in orthopaedic disease detection: a taxonomy for analysis and trustworthiness evaluation. Int J Comput Intell Syst. 2024;17:303.
66. Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ. 2025; 388:e082505.
67. Nowroozi A, Salehi MA, Shobeiri P, Agahi S, Momtazmanesh S, Kaviani P, et al. Artificial intelligence diagnostic accuracy in fracture detection from plain radiographs and comparing it with

clinicians: a systematic review and meta-analysis. Clin Radiol. 2024;79:579–88.
68. Oettl FC, Zsidai B, Oeding JF, Hirschmann MT, Feldt R, Fendrich D, et al. Artificial intelligence-assisted analysis of musculoskeletal imaging—a narrative review of the current state of machine learning models. Knee Surg Sports Traumatol Arthrosc. 2025;33:3032–8.
69. Oikonomou EK, Khera R. Machine learning in precision diabetes care and cardiovascular risk prediction. Cardiovasc Diabetol. 2023;22:259.
70. Oliveira e Carmo L, van den Merkhof A, Olczak J, Gordon M, Jutte PC, Jaarsma RL, et al. An increasing number of convolutional neural networks for fracture recognition and classification in orthopaedics. Bone Joint Open. 2021;2:879–85.
71. Oude Nijhuis KD, Prijs J, Barvelink B, van Luit H, Zhao Y, Liao Z, et al. Open-source convolutional neural network to classify distal radial fractures according to the AO/OTA classification on plain radiographs. Eur J Trauma Emerg Surg. 2025;51:261.
72. Padash S, Mickley JP, Vera Garcia DV, Nugen F, Khosravi B, Erickson BJ, et al. An overview of machine learning in orthopedic surgery: an educational paper. J Arthroplasty. 2023;38: 1938–42.
73. Powers D. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. J Mach Learn Technol. 2011;2:37–63.
74. Pundir BS, Kumar R, Gupta M, Obaid AJ. Synthetic image generation for improved detection of bone defects using GAN. In: 2024 15th international conference on computing communication and networking technologies (ICCCNT) [Internet]. IEEE; 2024. p. 1–5. Available from: https://ieeexplore.ieee.org/document/10723913/
75. Ramadanov N, John P, Hable R, Schreyer AG, Shabo S, Prill R, et al. Artificial intelligence-guided distal radius fracture detection on plain radiographs in comparison with human raters. J Orthop Surg. 2025;20:468.
76. Ramesh S, Sathyavarapu NSB, Sharma SJ, Nippun V, Khanna NK M. Comparative analysis of Q-learning, SARSA, and deep Q-network for microgrid energy management. Sci Rep. 2025;15:694.
77. Ran X, Xi Y, Lu Y, Wang X, Lu Z. Comprehensive survey on hierarchical clustering algorithms and the recent developments. Artif Intell Rev. 2023;56:8219–64.
78. Rezapour M, Seymour RB, Sims SH, Karunakar MA, Habet N, Gurcan MN. Employing machine learning to enhance fracture recovery insights through gait analysis. J Orthop Res. 2024; 42:1748–61.
79. Rizk PA, Gonzalez MR, Galoaa BM, Girgis AG, Van Der Linden L, Chang CY, et al. Machine learning–assisted decision making in orthopaedic oncology. JBJS Rev. 2024; 12:00057.
80. Rolf B, Jackson I, Müller M, Lang S, Reggelin T, Ivanov D. A review on reinforcement learning algorithms and applications in supply chain management. Int J Prod Res. 2023;61: 7151–79.
81. Romoff M, Brunette M, Peterson MK, Hashmi SZ, Kim MS. The role of large language models in improving the readability of orthopaedic spine patient educational material. J Orthop Surg. 2025;20:531.
82. Rosenberg GS, Magnéli M, Barle N, Kontakis MG, Müller AM, Wittauer M, et al. ChatGPT-4 generates orthopedic discharge documents faster than humans maintaining comparable quality: a pilot study of 6 cases. Acta Orthop. 2024;95:152–6.
83. Rupp MC, Moser LB, Hess S, Angele P, Aurich M, Dyrna F, et al. Orthopaedic surgeons display a positive outlook towards artificial intelligence: a survey among members of the AGA Society for Arthroscopy and Joint Surgery. J Exp Orthop. 2024;11:e12080.
84. Sadr MM, Khani M, Tootkaleh SM. Predicting athletic injuries with deep Learning: evaluating CNNs and RNNs for enhanced performance and Safety. Biomed Signal Process Control. 2025;105:107692.
85. Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput Sci. 2021;2:160.
86. Schreck J, Laukamp KR, Niehoff JH, Michael AE, Boriesosdick J, Wöltjen MM, et al. Metal artifact reduction in patients with total hip replacements: evaluation of clinical photon counting CT using virtual monoenergetic images. Eur Radiol. 2023;33:9286–95.
87. Sequí-Sabater JM, Benavent D. Artificial intelligence in rheumatology research: what is it good for? RMD Open. 2025; 11:e004309.
88. Shahzad H, Veliky C, Le H, Qureshi S, Phillips FM, Javidan Y, et al. Preserving privacy in big data research: the role of federated learning in spine surgery. Eur Spine J. 2024;33: 4076–81.
89. Shelmerdine SC, White RD, Liu H, Arthurs OJ, Sebire NJ. Artificial intelligence for radiological paediatric fracture assessment: a systematic review. Insights Imaging. 2022; 13:94.
90. Siddiqui IA, Littlefield N, Carlson LA, Gong M, Chhabra A, Menezes Z, et al. Fair AI-powered orthopedic image segmentation: addressing bias and promoting equitable healthcare. Sci Rep. 2024;14:16105.
91. Simion AM, Radu Ș, Florea AM. A review of generative adversarial networks for computer vision tasks. Electronics. 2024;13:713.
92. Singh P, Goyal L, Mallick DC, Surani SR, Kaushik N, Chandramohan D, et al. Artificial intelligence in nephrology: clinical applications and challenges. Kidney Med. 2025; 7:100927.
93. Sounderajah V, Guni A, Liu X, Collins GS, Karthikesalingam A, Markar SR, et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. Nat Med. 2025;31:3283–9.
94. Teo ZL, Jin L, Li S, Miao D, Zhang X, Ng WY, et al. Federated machine learning in healthcare: a systematic review on clinical applications and technical architecture. Cell Rep Med. 2024; 5:101419.
95. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), Official Journal L 2024/1689; 2024. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng
96. Turing AM. Computing machinery and intelligence. Mind. 1950;59:433–60.
97. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. BMJ. 2022;377:e070904.
98. Wang J, Zhu H, Liu H, Ma Z. Lossy point cloud geometry compression via end-to-end learning. IEEE Trans Circuits Syst Video Technol. 2021;31:4909–23.
99. Waqas M, Humphries UW. A critical review of RNN and LSTM variants in hydrological time series predictions. MethodsX. 2024;13:102946.
100. Wu H, Wang M, Wu J, Francis F, Chang Y-H, Shavick A, et al. A survey on clinical natural language processing in the United Kingdom from 2007 to 2022. npj Digit Med. 2022;5:186.
101. Wyles CC, Fu S, Odum SL, Rowe T, Habet NA, Berry DJ, et al. External validation of natural language processing algorithms to extract common data elements in THA operative notes. J Arthroplasty. 2023;38:2081–4.

102. Xie Y, Wang X, Yang H, Zhang J, Wang H, Yan Z, et al. Swarm learning network for privacy-preserving and collaborative deep learning assisted diagnosis of fracture: a multi-center diagnostic study. Front Med. 2025;12:1534117.
103. Yağar H, Gümüşoğlu E, Asfuroğlu ZM. Assessing the performance of ChatGPT-4o on the Turkish Orthopedics and Traumatology Board Examination. Joint Dis Relat Surg. 2025;36:304–10.
104. Yang Y, Liu Y, Chen Y, Luo D, Xu K, Zhang L. Artificial intelligence for predicting treatment responses in autoimmune rheumatic diseases: advancements, challenges, and future perspectives. Front Immunol. 2024;15:1477130.
105. Yao JJ, Aggarwal M, Lopez RD, Namdari S. Large language models in orthopaedics. J Bone Joint Surg. 2024;106:1411–8.
106. Yao JJ, Lopez RD, Rizk AA, Aggarwal M, Namdari S. Evaluation of a popular large language model in orthopedic literature review: comparison to previously published reviews. Arch Bone Joint Surg. 2025;13:460–9.
107. Youssef Y, De Wet D, Back DA, Scherer J. Digitalization in orthopaedics: a narrative review. Front Surg. 2024;10:1325423.
108. Yu C-H, Yeh C-C, Lu Y-F, Lu Y-L, Wang T-M, Lin FY-S, et al. Recurrent neural network methods for extracting dynamic balance variables during gait from a single inertial measurement unit. Sensors. 2023;23:9040.
109. Yurdem B, Kuzlu M, Gullu MK, Catak FO, Tabassum M. Federated learning: overview, strategies, applications, tools and future directions. Heliyon. 2024;10:e38137.
110. Zhang C, Huang W, Niu T, Liu Z, Li G, Cao D. Review of clustering technology and its application in coordinating vehicle subsystems. Automot Innov. 2023;6:89–115.
111. Zhang C, Liu S, Zhou X, Zhou S, Tian Y, Wang S, et al. Examining the role of large language models in orthopedics: systematic review. J Med Internet Res. 2024;26:e59607.
112. Zhao X, Wang L, Zhang Y, Han X, Deveci M, Parmar M. A review of convolutional neural networks in computer vision. Artif Intell Rev. 2024;57:99.
113. Zhu J, Jiang Y, Chen D, Lu Y, Huang Y, Lin Y, et al. High identification and positive-negative discrimination but limited detailed grading accuracy of ChatGPT-4o in knee osteoarthritis radiographs. Knee Surg Sports Traumatol Arthrosc. 2025;33:1911–9.
114. Zsidai B, Hilkert AS, Kaarre J, Narup E, Senorski EH, Grassi A, et al. A practical guide to the implementation of AI in orthopaedic research—part 1: opportunities in clinical application and overcoming existing challenges. J Exp Orthop. 2023;10:00683-z.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Rosen J, Russell J, Kartik P, Vella-Baldacchino M. Artificial intelligence algorithms in orthopaedics: a narrative review of methods and clinical applications. J Exp Orthop. 2025;12:e70549. https://doi.org/10.1002/jeo2.70549