International Journal of Science and Healthcare Research
[Indexed, Peer-Reviewed Journal]
Vol. 11; Issue: 3; July-Sept. 2026
Website: ijshr.com
ISSN: 2455-7587

Review Article

Current Applications of Artificial Intelligence in Orthopaedic Trauma: A Narrative Review

Ashutosh Yadav[1], Pushpa[2], Sachin Kumar[3]

[1]Senior Resident, Department of Orthopaedics, Maharishi Chyawan Government Medical College, Koriawas, Narnaul, Mahendragarh, Haryana 123001, India.
[2]Assistant Professor, Department of Anaesthesia, Maharishi Chyawan Government Medical College, Koriawas, Narnaul, Mahendragarh, Haryana 123001, India.
[3]Senior Resident, Department of Orthopaedic Surgery, Balrampur Hospital, Golaganj, Lucknow, Uttar Pradesh 226001, India.

Corresponding Author: Dr. Ashutosh Yadav

DOI: https://doi.org/10.52403/ijshr.20260330

Received: 05/05/2026; Revised: 23/07/2026; Accepted: 06/09/2026

ABSTRACT

Background: Deep learning, and in particular the convolutional neural network (CNN), has moved from proof of concept to commercial deployment in orthopaedic trauma within a single decade. Practising surgeons increasingly encounter these tools without a clear account of what the evidence does and does not support.

Objective: To synthesise current evidence on artificial intelligence across the orthopaedic trauma pathway, covering fracture detection, classification, outcome prediction, operative support and generative language applications; to appraise the methodological quality of that evidence; and to define what is required for safe adoption in India and other low- and middle-income countries.

Methods: PubMed (MEDLINE), Embase and Scopus were searched for English-language, peer-reviewed studies published between January 2015 and April 2026. Titles, abstracts and full texts were screened against pre-specified inclusion and exclusion criteria, and findings were synthesised narratively. Methodological quality was appraised against the domains of QUADAS-2 for diagnostic-accuracy studies and PROBAST for prediction-model studies.

Key findings: Pooled sensitivity and specificity for fracture detection on plain radiographs clustered around 0.87 to 0.91, comparable with specialist clinicians, and clinician sensitivity rose to 0.97 when the algorithm was used as an adjunct. Site-specific networks matched or exceeded expert performance at the hip, distal radius, scaphoid, proximal humerus, ankle, ribs and spine. Classification remained weaker (accuracy 60% to 81% for the distal radius), and outcome-prediction models gained little over conventional regression (area under the curve 0.80 versus 0.79 for 30-day mortality after hip fracture). Generative language models showed early utility in documentation and patient education but did not reach resident-level performance and produced erroneous content at clinically relevant rates. Risk of bias was high in about half of the diagnostic studies, and only about one model-development study in nine reported external validation.

Conclusion: Artificial intelligence has achieved specialist-level accuracy for fracture detection but has not yet demonstrated patient-level benefit. The priorities for India are representative multicentric datasets, external validation, CLAIM- and TRIPOD+AI-compliant reporting, and prospective clinical-impact trials rather than further internal-validation accuracy studies.

Keywords: Artificial intelligence; Deep learning; Convolutional neural network; Orthopaedic trauma; Fracture detection; Large language models.

INTRODUCTION

Artificial intelligence (AI) refers to computational systems that perform tasks normally requiring human cognition: pattern recognition, classification, prediction and decision support. Two paradigms dominate medical AI. Classical machine learning (ML) learns rules from structured tabular features such as age, sex and laboratory values. Deep learning (DL) uses multi-layer neural networks, most notably convolutional neural networks (CNNs), to learn hierarchical representations directly from raw inputs such as radiographic pixels [1,2]. Since the 2012 ImageNet breakthrough and the development of architectures such as VGG, ResNet, Inception, DenseNet and You-Only-Look-Once (YOLO), DL has reshaped medical image interpretation [1].

Orthopaedic trauma is an archetypal early target for medical AI. Plain radiography remains the diagnostic cornerstone and generates the highest imaging volume in most emergency departments; a fracture is a discrete visual pattern well suited to CNN learning; and missed or delayed fracture diagnosis causes substantial morbidity and is a leading source of diagnostic malpractice claims [3,4]. In a classic emergency-department audit, Guly found that 79.7% of diagnostic errors were missed fractures, most often from radiograph misinterpretation (77.8%) [5]. A tool that flags fractures in real time therefore addresses a documented failure point rather than a hypothetical one.

The volume of publication has become difficult to navigate. In a 2025 review of 217 studies, Misir reported that 52.5% appeared in 2024 alone, with detection (24.4%) and classification (12.0%) the commonest applications and the hip/femur, spine and wrist the most studied regions [6]. Several meta-analyses report pooled accuracy approaching or exceeding 90% on plain radiographs [7-9], and a parallel literature addresses outcome prediction [10], robotic fixation, report triage and rehabilitation monitoring.

Enthusiasm has nonetheless outpaced validation. Concerns persist regarding methodological transparency, external generalisability, interpretability, medico-legal accountability, equity of training data and workflow integration [11,12]. This review synthesises the evidence across the orthopaedic trauma continuum for the practising surgeon, appraises its methodological quality, and states explicitly what is and is not ready for the Indian clinical environment.

MATERIALS AND METHODS

Article identification

PubMed (MEDLINE), Embase (Elsevier) and Scopus were searched for English-language, peer-reviewed publications from 1 January 2015 to 30 April 2026. Search terms combined controlled vocabulary and free-text variants of "artificial intelligence", "machine learning", "deep learning" and "convolutional neural network" with "fracture", "orthopaedic trauma" and "musculoskeletal trauma", together with anatomical site terms. The search was supplemented by hand-searching reference lists of retrieved systematic reviews and of the 2025 review by Misir [6], and by forward citation tracking of the principal meta-analyses [7-10].

Screening approach

Records were screened in two stages by the authors: titles and abstracts for topical relevance, then full texts against the criteria below. Disagreement was resolved by discussion. Because this is a narrative rather than a systematic review, screening was not independently duplicated and no record count or PRISMA flow diagram is reported.

Inclusion criteria

Studies were eligible if they (i) applied AI, ML or DL to a clinical task within the orthopaedic trauma pathway, including detection, classification, segmentation, outcome prediction, operative planning, documentation or rehabilitation; (ii) reported quantitative performance against an explicit reference standard, or were systematic reviews or meta-analyses of such studies; and (iii) appeared in a peer-reviewed journal. Externally validated models, prospective designs and reader studies with clinician comparators were prioritised.

Exclusion criteria

Excluded were conference abstracts, preprints, editorials without original data, single-case technical demonstrations, studies of degenerative or oncological disease without a trauma component, studies in which the reference standard was not stated, and duplicate reports of the same dataset.

Rationale for study selection

Selection was purposive rather than exhaustive, designed to answer a clinical rather than a bibliometric question: what can a practising orthopaedic surgeon currently rely on? Priority went to the highest available level of synthesis for each task, the largest and most rigorously labelled datasets, and studies reporting clinician comparators. Where several studies conveyed the same message, representative ones were tabulated rather than described individually. Methodological quality was appraised narratively against QUADAS-2 and PROBAST domains [13,14]. This article does not claim systematic review status in the PRISMA 2020 sense [15]. As only published literature was synthesised, institutional ethics committee approval was not required.

OVERVIEW OF AI TECHNIQUES RELEVANT TO ORTHOPAEDIC TRAUMA

Three technical families account for almost all published applications. Supervised CNNs are trained on labelled radiographs to output a fracture probability or a localisation bounding box; ResNet, DenseNet, VGG, Inception, EfficientNet and YOLO variants predominate [6]. Classical ML methods, including regularised logistic regression, random forests, gradient-boosted trees and support vector machines, are applied mainly to tabular clinical data for outcome prediction [10]. Natural language processing (NLP), including recent generative large language models (LLMs), is applied to free-text reports and electronic health records for cohort identification, coding and triage [6].

Reported metrics include sensitivity, specificity, accuracy, predictive values, F1 score and the area under the receiver operating characteristic curve (AUC); localisation tasks add intersection over union and mean average precision. Prediction-model studies should also report calibration, not discrimination alone, a requirement frequently unmet. CLAIM and TRIPOD+AI are the contemporary reporting standards [16,17], and QUADAS-2 and PROBAST the corresponding risk-of-bias tools [13,14]. Figure 1 maps these technique families onto the trauma pathway and structures the sections that follow.

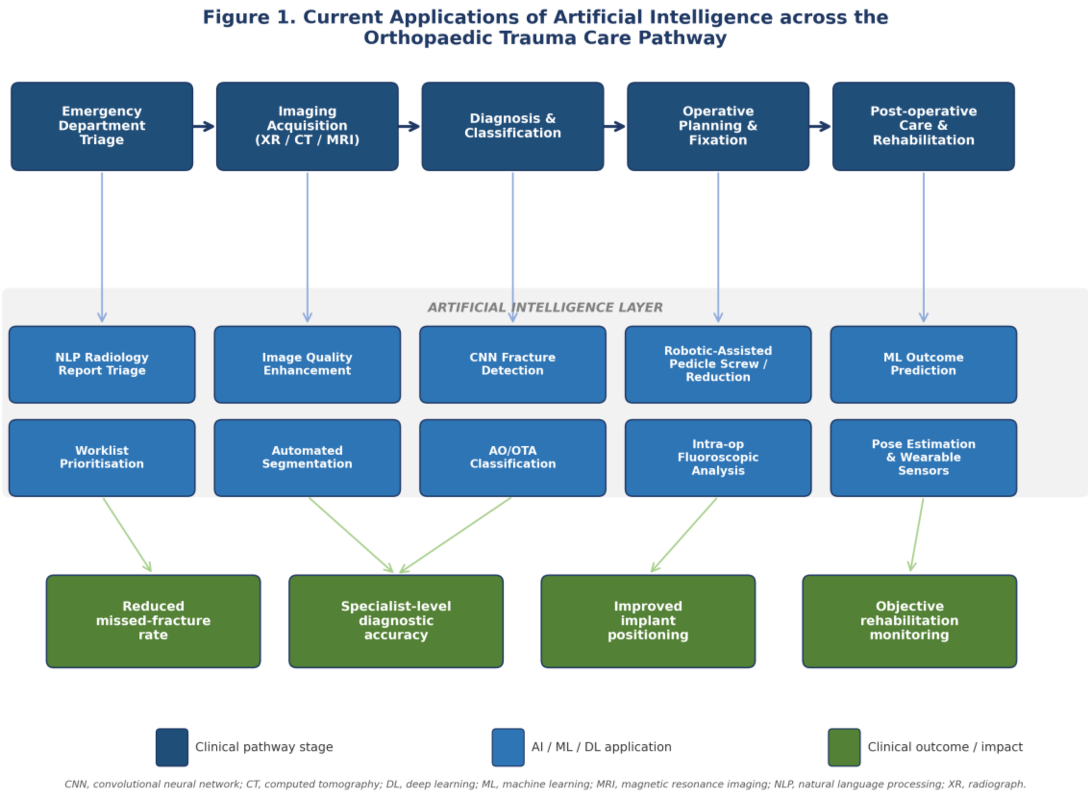

Figure 1: Applications of artificial intelligence (AI) across the orthopaedic trauma care pathway. The middle layer shows the AI tasks deployed at each clinical stage; the lower row summarises the resulting clinical impacts. CNN, convolutional neural network; CT, computed tomography; DL, deep learning; ML, machine learning; MRI, magnetic resonance imaging; NLP, natural language processing; XR, radiograph.

AI FOR FRACTURE DETECTION

Detection on conventional radiographs is by some distance the most mature application, and three meta-analyses frame the overall picture (Table 1). Kuo and colleagues, pooling 42 studies, reported AI sensitivity and specificity of 0.91 (95% CI 0.84-0.95 and 0.81-0.95 respectively), statistically comparable with clinicians at external validation; the clinically important finding was that clinician sensitivity rose to 0.97 when AI was used as an adjunct rather than as a replacement [7]. Yang and colleagues reported closely similar pooled estimates independently [8], and Jung and colleagues, across 66 studies, confirmed that imaging-based models substantially outperform tabular and hybrid approaches [9]. These estimates must be read against a single sobering figure: 22 of the 42 studies (52%) in Kuo's meta-analysis were at high risk of bias on QUADAS-2 appraisal [7], a caveat first raised by Langerhuizen and colleagues [18].

Table 1: Major systematic reviews and meta-analyses on artificial intelligence in orthopaedic trauma.

First author, year (journal)	Scope	Studies (n)	Key findings
Langerhuizen, 2019 (Clin Orthop Relat Res) [18]	Detection and classification in trauma imaging	10	Feasibility established; reference-standard rigour and external validity flagged
Yang, 2020 (Clin Radiol) [8]	Deep learning in fracture detection	14 (9 pooled)	Pooled sensitivity 0.87, specificity 0.91
Kuo, 2022 (Radiology) [7]	AI in fracture detection	42	Pooled sensitivity and specificity 0.91; AI-assisted clinician sensitivity 0.97; 52% high risk of bias

First author, year (journal)	Scope	Studies (n)	Key findings
Lex, 2023 (JAMA Netw Open) [10]	Hip fracture detection and outcome prediction	39	Sensitivity 89.3%, specificity 87.5%; non-inferior to clinicians; 30-day mortality AUC 0.80 vs 0.79
Jung, 2024 (PLOS Digit Health) [9]	Detection across image and tabular modalities	66	Imaging-based models outperform tabular and hybrid; marked heterogeneity
Mohammed, 2024 (Int J Comput Intell Syst) [19]	Taxonomic review of AI in orthopaedic disease detection	85	Six-category taxonomy; trustworthiness concerns emphasised
Husarek, 2024 (Sci Rep) [20]	Commercial fracture-detection products	17	Favourable pooled accuracy; industry-funding bias; sparse independent validation
Oude Nijhuis, 2024 (Eur J Trauma Emerg Surg) [21]	AI for distal radius fractures	15	Detection AUC 0.87-0.99; classification accuracy only 60-81%
Misir, 2025 (Injury) [6]	Comprehensive review of AI in orthopaedic trauma	217	52.5% published in 2024; hip/femur, spine and wrist most studied
Yu, 2022 (Radiol Artif Intell) [22]	External validation of DL radiology algorithms	Not pooled	External validation uncommon; 5-10 pp accuracy decrement when performed

Abbreviations: *AI, artificial intelligence; AUC, area under the receiver operating characteristic curve; DL, deep learning; pp, percentage points.*

Hip and proximal femur

The hip is the most-studied site; performance figures for representative studies at every site are tabulated in Table 2 rather than repeated here. Urakawa and colleagues trained a VGG-16 CNN on 3,346 proximal-femoral images and outperformed five comparator orthopaedic surgeons on every metric (accuracy 95.5% versus 92.2%) [23]. Cheng and colleagues added gradient-weighted class activation mapping to a DenseNet-121, reaching AUC 0.98 while supplying the heat-map localisation that supports clinician trust [24]. Krogue and colleagues advanced the task from binary detection to functional subclassification of femoral-neck and intertrochanteric fractures [25]. Mutasa and colleagues showed that data augmentation can partly compensate for limited training sets [26], and Yamada and colleagues demonstrated the incremental value of integrating two radiographic views [27].

The definitive synthesis remains that of Lex and colleagues, covering 39 studies, 39,598 radiographs and 714,939 patient records [10]. Two conclusions carry practical weight. For diagnosis, the pooled odds ratio for diagnostic error of ML versus clinicians was 0.79 (95% CI 0.48-1.31), which establishes non-inferiority but not superiority. For post-operative outcome prediction, the median AUC for 30-day mortality was 0.80 against approximately 0.79 for conventional scoring, an increment that does not obviously justify the accompanying loss of interpretability [10].

Wrist, distal radius and scaphoid

The distal radius is the commonest upper-limb fracture site and is frequently first assessed by non-specialists, making it a high-value target for decision support. Reported AUCs cluster between 0.90 and 0.99 (Table 2). Suzuki and colleagues approached the performance ceiling by combining two views, matching or exceeding three hand surgeons [28]. Thian and colleagues illustrate the opposite caution: sensitivity of 98.1% was purchased at a specificity of 72.9%, an operating point that would generate a substantial false-positive workload in a busy casualty [29]. Gan and colleagues found CNN performance superior to radiologists and equivalent to orthopaedists [30], while Oude Nijhuis and colleagues released an open-source model, a rare contribution to reproducibility [31]. Most importantly, Lindsey and colleagues showed in a controlled reader study that decision support significantly improved clinician sensitivity, shifting the evidence from algorithmic accuracy toward human-machine performance [32].

Scaphoid fracture, where a missed diagnosis commonly progresses to non-union, is of particular clinical relevance. Hendrix and colleagues developed and validated a CNN reaching musculoskeletal-radiologist-level performance on multi-view radiographs [33,34], and Ozkaya and colleagues reported concordant results in an independent evaluation [35].

Proximal humerus, ankle and axial skeleton

Chung and colleagues trained a ResNet-152 for detection and Neer classification of proximal humerus fractures, exceeding comparator specialists on binary accuracy [36]. Prijs and colleagues reported one of the few externally validated orthopaedic models, retaining discrimination for ankle fracture detection, classification and localisation on an out-of-distribution test set [37], and Olczak and colleagues extended this to AO/OTA 2018 malleolar classification [38]. In polytrauma, where rib fractures are notoriously underdiagnosed, Weikert and colleagues validated a DL algorithm on whole-body trauma computed tomography (CT) with high per-patient sensitivity [39]; Husarek and colleagues have since synthesised CT-based detection more broadly [40]. Hong and colleagues demonstrated simultaneous detection of vertebral fracture and osteoporosis from lateral spine radiographs, enabling opportunistic screening [41], an application with plausible public-health value in India.

Paediatric and multi-region models

Paediatric application must contend with physeal anatomy and growth-plate variability. Dupuis and colleagues externally validated a commercial algorithm in children and found preserved but lower performance than in adults, which argues against deploying adult-trained models in paediatric casualty without age-stratified validation [42]. Regnard and colleagues assessed a commercial algorithm across limbs, pelvis, focal bone lesions and elbow effusions with measurable improvement in reader performance [43], and Anderson and colleagues showed that DL assistance closed the accuracy gap between less-experienced and specialist clinicians, the largest gains accruing to residents and non-musculoskeletal radiologists [44]; Cohen and colleagues reported non-inferiority for wrist fractures [45]. Two workflow studies report acceptable turnaround in prospective routine use [11], and fewer missed fractures without an increase in reading time [12].

Table 2: Representative deep-learning studies for fracture detection across major anatomical regions.

Anatomical site	First author, year (journal)	Architecture / dataset	Reported performance
Intertrochanteric hip	Urakawa, 2019 (Skeletal Radiol) [23]	VGG-16; 3,346 images	Accuracy 95.5% vs 92.2% for 5 surgeons
Hip (pelvic radiograph)	Cheng, 2019 (Eur Radiol) [24]	DenseNet-121 + Grad-CAM; 25,505 radiographs	AUC 0.98 with heat-map localisation
Hip (multi-class)	Krogue, 2020 (Radiol Artif Intell) [25]	DenseNet detection; 3,026 hips	Binary accuracy 93.7%; functional subclassification
Femoral neck	Mutasa, 2020 (J Digit Imaging) [26]	CNN with GAN and DRR augmentation; 9,063 images	Detection AUC 0.92; Garden classification AUC 0.96
Distal radius	Gan, 2019 (Acta Orthop) [30]	Inception-v4 + Faster R-CNN; 2,340 radiographs	AUC 0.96; superior to radiologists
Distal radius	Thian, 2019 (Radiol Artif Intell) [29]	Inception-ResNet Faster R-CNN; 7,356 studies	Sensitivity 98.1% but specificity 72.9%
Distal radius	Suzuki, 2022 (J Digit Imaging) [28]	DCNN on two views	Accuracy 99.3%, AUC 0.993
Wrist (decision support)	Lindsey, 2018 (Proc Natl Acad Sci USA) [32]	U-Net + DCNN; 135,409 radiographs	Clinician sensitivity significantly improved

Anatomical site	First author, year (journal)	Architecture / dataset	Reported performance
Scaphoid	Hendrix, 2021 (Radiol Artif Intell) [33]	CNN; multi-view hand and wrist	Musculoskeletal-radiologist-level discrimination
Proximal humerus	Chung, 2018 (Acta Orthop) [36]	ResNet-152; shoulder radiographs	Accuracy 96% vs 93% for specialists
Ankle	Prijs, 2023 (Eur J Trauma Emerg Surg) [37]	CNN with external validation	Discrimination retained out of distribution
Ankle (AO/OTA)	Olczak, 2021 (Acta Orthop) [38]	CNN; AO/OTA 2018 malleolar	High correct-classification rate
Rib (trauma CT)	Weikert, 2020 (Korean J Radiol) [39]	DL; whole-body trauma CT	High per-patient sensitivity
Vertebra (lateral radiograph)	Hong, 2023 (J Bone Miner Res) [41]	DL spine radiograph model	Concurrent fracture and osteoporosis detection
Paediatric	Dupuis, 2022 (Diagn Interv Imaging) [42]	Commercial DL; external validation in children	Preserved but lower than adult performance
Multi-region limb/pelvis	Regnard, 2022 (Eur J Radiol) [43]	BoneView commercial DL	Improved reader detection of fractures and dislocations

Abbreviations: *AUC, area under the receiver operating characteristic curve; CNN, convolutional neural network; CT, computed tomography; DCNN, deep convolutional neural network; DL, deep learning; DRR, digitally reconstructed radiograph; GAN, generative adversarial network.*

AI FOR FRACTURE CLASSIFICATION

If detection is the field's success story, classification is its reality check. Classification supports AO/OTA coding, operative decision-making and outcome stratification, but performs consistently worse than detection because the reference standard itself carries substantial inter-observer disagreement. Olczak and colleagues achieved a high rate of correct AO/OTA-2018 malleolar classification [38], and Chung and colleagues reported Neer classification concordant with specialists [36]. The clearest quantification of the gap comes from Oude Nijhuis and colleagues: across 15 distal radius studies, classification accuracy was 60-81% and AUC 0.59-0.84, against 0.87-0.99 for detection in the same series [21]. Prediction of loss of reduction after cast immobilisation, the most clinically actionable task in this domain, remains an open question [21].

AI FOR OUTCOME PREDICTION

Prognostic modelling has attracted comparable enthusiasm and delivered less. Published models address predominantly hip-fracture mortality, length of stay, delirium, 30-day complications, discharge destination and functional recovery. The synthesis by Lex and colleagues found only modest incremental discrimination over conventional regression, with wide heterogeneity in input variables and model complexity [10]. Beyond the hip, models exist for distal-radius functional outcomes, non-union risk after long-bone trauma and neurological recovery in spinal trauma. The recurring lesson is that well-specified clinical scores remain competitive benchmarks, and that most apparent ML gains derive from high-dimensional data integration rather than from the algorithm itself [10]. Responsible deployment therefore requires calibration assessment, external validation and PROBAST-compliant appraisal [14].

METHODOLOGICAL QUALITY OF THE EVIDENCE BASE

Headline accuracy figures in this field are systematically optimistic, for structural rather than incidental reasons. Appraised against QUADAS-2 and PROBAST domains, the same weaknesses recur across diagnostic and prognostic studies alike (Table 3). Patient selection is usually retrospective and single-centre, often with case-control sampling and exclusion of technically imperfect radiographs, which are precisely the images on which clinicians most need help. Reference standards are frequently derived from radiology reports rather than CT, operative findings or follow-up, so the model learns to reproduce the reporting radiologist rather than the underlying truth. Index-test thresholds are commonly selected after the fact, and prediction-model studies report discrimination while omitting calibration.

Two quantitative anchors define the scale of the problem. Kuo and colleagues judged 52% of the 42 studies in their meta-analysis to be at high risk of bias [7], and Oliveira e Carmo and colleagues, screening 1,349 records, found that of 36 studies developing a CNN for fracture detection or classification only four (11%) reported any external validation [46]; Yu and colleagues reached the same conclusion across radiology more broadly [22]. The practical implication is that a reported internal AUC of 0.95 should be discounted rather than accepted at face value, the decrement on out-of-distribution data being of the order of 5 to 10 percentage points [22].

Table 3: Recurrent risk-of-bias concerns in orthopaedic trauma AI studies, mapped to QUADAS-2 and PROBAST domains.

Domain (tool)	Recurrent problem	Illustrative evidence
Patient selection (QUADAS-2)	Retrospective single-centre or case-control sampling; equivocal and technically poor radiographs excluded, inflating apparent accuracy	52% of 42 studies at high risk of bias [7]
Index test (QUADAS-2)	Thresholds chosen after analysis; readers unblinded in adjunct studies; best-performing model iteration reported selectively	Design limitations flagged from the outset [18]
Reference standard (QUADAS-2)	Labels from radiology reports rather than CT, operative findings or follow-up; inter-observer disagreement rarely quantified	Classification accuracy 60-81% vs detection AUC 0.87-0.99 [21]
Flow and timing (QUADAS-2)	Excluded images incompletely accounted for; partial verification of the reference standard	Heterogeneity across 66 studies [9]
Predictors and outcome (PROBAST)	Data-driven predictor selection without clinical rationale; outcome definitions differ across registries	Wide heterogeneity in inputs and complexity [10]
Analysis (PROBAST)	Discrimination reported without calibration; low events-per-variable; overfitting unaddressed	Median AUC 0.80 vs 0.79 for regression [10]
Applicability (both)	Training data skewed away from South Asian populations; external validation rare and performance falls when performed	4 of 36 CNN studies (11%) externally validated [22,46]

Abbreviations: *AI, artificial intelligence; AUC, area under the receiver operating characteristic curve; CNN, convolutional neural network; CT, computed tomography; PROBAST, Prediction model Risk of Bias Assessment Tool; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies 2.*

PRE-OPERATIVE PLANNING, ROBOTICS AND INTRA-OPERATIVE APPLICATIONS

Beyond diagnosis, AI is being applied to the operative episode itself. Automated fragment segmentation and virtual reduction have progressed beyond feasibility at selected sites: Jeon and colleagues clinically validated a deep-learning segmentation and simulation model for Neer 3- and 4-part proximal humerus fractures, which reproduced post-operative anatomy more closely than manual reduction (Dice coefficient 0.78 versus 0.69, $P < 0.001$) in less working time [47]. In spinal trauma, commercial robotic platforms integrate AI-guided trajectory planning with intra-operative navigation, with comparative studies reporting improved pedicle-screw accuracy relative to freehand technique [48].

Post-operative surveillance is more mature. Tan and colleagues developed classification and object-detection models for implant cut-out on post-operative pelvic radiographs using 40,191 images, achieving an AUROC of 0.998 and 95.5% localisation accuracy on true-positive images, a proof of concept for automated monitoring of fixation failure [49].

Intra-operative fluoroscopic analysis of reduction quality and implant position, including tip-apex distance in cephalomedullary nailing, is a logical extension of detection-domain CNNs but remains far less mature. Theatre translation additionally requires real-time inference, vendor-neutral integration with fluoroscopy workstations and certification as Software as a Medical Device.

GENERATIVE AI AND LARGE LANGUAGE MODELS IN ORTHOPAEDIC TRAUMA

The applications so far are discriminative: the model assigns a label to an image or a patient. Generative models produce text instead, and have entered orthopaedic practice through the clinic, the consent conversation and the electronic record rather than the radiology worklist. Two systematic reviews map this literature. Zhang and colleagues found wide heterogeneity in how performance was defined, with diagnostic accuracy of 55% to 93% and orthopaedic examination scores of 45% to 73.6%, and concluded that LLMs cannot replace orthopaedic professionals in the short term but may serve as a copilot [50]. Mo and colleagues, reviewing 60 studies, found that newer models outperformed their predecessors but that orthopaedic residents (74.2-75.3%) continued to outscore the models on in-training examination questions [51].

Clinical documentation

Documentation is the most rigorously tested application to date. Baker and colleagues conducted a blinded randomised trial comparing typing, dictation and ChatGPT for documenting the history of present illness from standardised patient histories. The LLM produced longer and, on the Physician Documentation Quality Instrument, higher-quality histories than either comparator, but introduced erroneous information into 36% of documents [52]. That combination of better prose and unreliable facts mandates clinician verification of every generated note before it enters the record. Related NLP applications, including automated cohort identification, fracture-registry population and severity coding from unstructured reports, make lighter demands on factual generation and are correspondingly closer to deployment [6].

Patient education

Patients now consult chatbots before and after fracture consultations, whether or not surgeons endorse it. Christy and colleagues found an online platform's answers to common distal radius fracture questions broadly appropriate but variably reliable [53], and the wider orthopaedic pattern is consistent: readability and tone are good, while accuracy, currency and source attribution are not [50]. For Indian practice, where patient material in regional languages is scarce, the pragmatic model is surgeon-curated rather than patient-facing: LLM-drafted counselling material reviewed and approved by the treating clinician before release.

Operative planning support

This is the weakest evidence domain and should be described as such. No validated LLM application for operative planning in orthopaedic trauma has been reported; published work is confined to examination and clinical-vignette questions rather than implant or approach recommendations for real patients, and Mo and colleagues identified image analysis as a functionality that remains largely unstudied [51]. Multimodal models combining radiographic input with clinical text are a plausible direction, but at present this is a research agenda rather than a clinical tool.

Limitations and hallucination risk

Four limitations constrain all current LLM use in trauma. First, hallucination: models generate fluent but fabricated content, most notoriously in citations. Bhattacharyya and colleagues found that most references generated by ChatGPT in medical content were fabricated or inaccurate [54]; the same failure mode applied to a clinical statement is more dangerous than a false radiographic positive, because it arrives with no image against which to check it. Second, non-determinism: the same prompt may yield different answers on different days or model versions, which is incompatible with the reproducibility expected of a medical device. Third, data governance: entering patient identifiers into public LLM endpoints raises confidentiality obligations that most Indian institutions have not addressed in policy. Fourth, regulatory status: general-purpose LLMs are not cleared as Software as a Medical Device for diagnostic use in any major jurisdiction, and responsibility for any resulting error rests entirely with the treating surgeon.

REHABILITATION AND WORKFLOW INTEGRATION

Pose-estimation and wearable-sensor systems allow objective monitoring of range-of-motion recovery and activity after fixation, and are amenable to smartphone-based tele-rehabilitation, a meaningful proposition for rural Indian follow-up where formal physiotherapy is scarce. Vogel and colleagues note, however, that much of the supporting evidence derives from other medical domains and has yet to be validated in orthopaedic trauma populations [55]. At departmental level, NLP-based triage of emergency radiology worklists could bring high-risk studies to attention earlier, although patient-level benefit is unproven [6].

COMMERCIAL PRODUCTS AND REGULATORY LANDSCAPE

Commercial translation has run ahead of independent evaluation. Van Leeuwen and colleagues catalogued more than 100 commercially available AI radiology products in a single year and found the supporting scientific evidence limited for many [56]. Husarek and colleagues reported broadly favourable pooled accuracy across commercial fracture-detection systems but flagged industry-funded study design and sparse independent validation [20]. Several such systems, including Gleamer BoneView, are in routine European and North American emergency-department use [43,56].

Regulation is evolving unevenly. The European Union Medical Device Regulation and the EU AI Act impose post-deployment monitoring obligations, and the United States Food and Drug Administration increasingly emphasises post-market surveillance and transparent change control for adaptive algorithms. India regulates through the Central Drugs Standard Control Organisation under the Medical Device Rules, within which software-based devices are still being accommodated; deployment in Indian institutions therefore remains limited and largely at the discretion of individual centres.

CHALLENGES AND LIMITATIONS

Beyond study-level risk of bias, several cross-cutting challenges constrain translation. Training datasets are heavily skewed toward North American, European and East Asian populations; differences in bone density, body habitus, implant selection and radiographic protocol limit direct transferability [46]. Interpretability remains partial: saliency mapping offers visual plausibility but not causal explanation adequate for medico-legal scrutiny, and Zech and colleagues showed that models can achieve apparently high performance by latching onto confounders such as side markers and scanner artefacts [57].
Reference-standard variability confounds cross-study comparison, since report-based labels, expert consensus and operative confirmation are not interchangeable. Automation bias and algorithm aversion are both documented failure modes, so AI output should be framed as adjunctive rather than authoritative. Medico-legal accountability in India remains ambiguous, health-economic evaluation is sparse, and the capital and integration costs of commercial products are substantial for public-sector institutions.

Limitations of this review

This is a narrative review: selection was purposive, screening was not independently duplicated, no PRISMA flow diagram is presented and no quantitative pooling was undertaken, so selection bias cannot be excluded. Risk-of-bias appraisal was narrative and domain-based rather than a formal scored assessment of every included study. The commercial landscape in particular is likely to have changed between the final search and publication.

INDIAN AND LOW- AND MIDDLE-INCOME COUNTRY CONTEXT

Three features of Indian trauma practice shape AI priorities. First, road-traffic and occupational fracture volume is high and case-loads routinely exceed the specialist staffing of district hospitals, so decision support that raises the floor of resident and emergency-physician performance has greater marginal value here than tools that raise the specialist ceiling. Second, digital radiography is now widespread, but picture-archiving and communication systems are inconsistently deployed, which constrains real-time integration far more than algorithm accuracy does. Third, representative Indian datasets are scarce: no large-scale, consented Indian trauma imaging repository comparable to international resources such as FracAtlas [58] exists. Curating a multicentric Indian registry under appropriate ethical governance is the highest-yield investment available, and federated learning offers a privacy-preserving route suited to the Indian hospital landscape.

Future Directions

Priorities for the next five years follow from the deficiencies identified above: multicentre prospective validation with pre-registered protocols; uniform adherence to CLAIM, TRIPOD+AI and DECIDE-AI reporting standards; benchmark datasets inclusive of South Asian populations; pragmatic randomised trials testing whether AI reduces missed-fracture rates, shortens time to theatre or improves patient-reported outcomes; multimodal integration of imaging and clinical data; explainable architectures adequate for medico-legal accountability; and incorporation of AI appraisal into residency curricula. What the field does not need is another generation of internal-validation accuracy studies.

CONCLUSION

Artificial intelligence, and particularly CNN-based deep learning, has achieved specialist-level diagnostic accuracy for fracture detection on conventional radiographs across the major sites of orthopaedic trauma, and demonstrably improves the performance of less-experienced clinicians when used as decision support. Its record elsewhere is weaker: classification lags behind detection, outcome prediction offers little over conventional regression, and generative language models remain unreliable in factual output. Roughly half of the diagnostic literature carries a high risk of bias and only about one study in nine reports external validation, so published accuracy should be read as an upper bound rather than an expectation. Translation into routine Indian practice requires external validation on representative datasets, prospective clinical-impact evidence, clear medico-legal frameworks and investment in national trauma imaging registries. Until that evidence exists, the orthopaedic surgeon should treat artificial intelligence as a clinical adjunct and retain full diagnostic and therapeutic accountability.

Declaration by Authors

Ethical Approval: Not applicable. This narrative review synthesised published studies and involved no human participants, identifiable patient data or animal experiments; ethics committee approval was therefore not required.

Acknowledgement: None.

Source of Funding: None. This review received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of Interest: The authors declare no conflict of interest.

Authors' Contributions:

AY: conceptualisation, literature search, synthesis and drafting. P: screening, data verification and critical revision for intellectual content. SK: literature search, tables and figure, and critical revision. All authors approved the final manuscript and agree to be accountable for all aspects of the work.

REFERENCES

1. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521(7553): 436-44. doi: https://doi.org/10.1038/nature14539.
2. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med. 2022; 28(1): 31-8. doi: 10.1038/s41591-021-01614-0.
3. Chen K, Stotter C, Klestil T, Nehrer S. Artificial intelligence in orthopedic radiography analysis: a narrative review. Diagnostics (Basel). 2022; 12(9): 2235. doi: 10.3390/diagnostics12092235.
4. Zech JR, Santomartino SM, Yi PH. Artificial Intelligence (AI) for Fracture Diagnosis: An Overview of Current Products and Considerations for Clinical Adoption, From the AJR Special Series on AI Applications. AJR Am J Roentgenol. 2022;219(6):869-878. doi: 10.2214/AJR.22.27873.
5. Guly HR. Diagnostic errors in an accident and emergency department. Emerg Med J. 2001; 18(4): 263–269. doi: 10.1136/emj.18.4.263
6. Misir A. Artificial intelligence in orthopedic trauma: a comprehensive review. Injury. 2025; 56(8):112570. doi: 10.1016/j.injury.2025.112570.
7. Kuo RYL, Harrison C, Curran TA, Jones B, Freethy A, Cussons D, et al. Artificial intelligence in fracture detection: a systematic review and meta-analysis. Radiology. 2022; 304(1):50-62. doi: 10.1148/radiol.211785.
8. Yang S, Yin B, Cao W, Feng C, Fan G, He S. Diagnostic accuracy of deep learning in orthopaedic fractures: a systematic review and meta-analysis. Clin Radiol. 2020; 75(9): 713.e17-713.e28. doi: 10.1016/j.crad.2020.05.021.
9. Jung J, Dai J, Liu B, Wu Q. Artificial intelligence in fracture detection with different image modalities and data types: a systematic review and meta-analysis. PLOS Digit Health. 2024; 3(1): e0000438. doi: 10.1371/journal.pdig.0000438.
10. Lex JR, Di Michele J, Koucheki R, Pincus D, Whyne C, Ravi B. Artificial intelligence for hip fracture detection and outcome prediction: a systematic review and meta-analysis. JAMA Netw Open. 2023; 6(3): e233391. doi: 10.1001/jamanetworkopen.2023.3391.
11. Oppenheimer J, Lüken S, Hamm B, Niehues SM. A prospective approach to integration of AI fracture detection software in radiographs into clinical workflow. Life (Basel). 2023; 13(1):223. doi: 10.3390/life13010223.
12. Canoni-Meynet L, Verdot P, Danner A, Calame P, Aubry S. Added value of an artificial intelligence solution for fracture detection in the radiologist's daily trauma emergencies workflow. Diagn Interv Imaging. 2022; 103(12):594-600. doi: 10.1016/j.diii.2022.06.004.
13. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011; 155(8):529-36. doi: 10.7326/0003-4819-155-8-201110180-00009.
14. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019; 170(1):51-58. doi: 10.7326/M18-1376.
15. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021; 372: n71. doi: 10.1136/bmj. n71.
16. Mongan J, Moy L, Kahn CE Jr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. Radiol Artif Intell. 2020; 2(2): e200029. doi: 10.1148/ryai.2020200029.
17. 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. doi: 10.1136/bmj-2023-078378.
18. Langerhuizen DWG, Janssen SJ, Mallee WH, van den Bekerom MPJ, Ring D,

Kerkhoffs GMMJ, et al. What are the applications and limitations of artificial intelligence for fracture detection and classification in orthopaedic trauma imaging? A systematic review. Clin Orthop Relat Res. 2019; 477(11):2482-2491. doi: 10.1097/CORR.0000000000000848.
19. 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(1): 303. doi: 10.1007/s44196-024-00718-y.
20. 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(1): 23053. doi: 10.1038/s41598-024-73058-8.
21. Oude Nijhuis KD, Dankelman LHM, Wiersma JP, Barvelink B, IJpma FFA, Verhofstad MHJ, et al. AI for detection, classification and prediction of loss of alignment of distal radius fractures: a systematic review. Eur J Trauma Emerg Surg. 2024; 50(6): 2819-2831. doi: 10.1007/s00068-024-02557-0.
22. Yu AC, Mohajer B, Eng J. External validation of deep learning algorithms for radiologic diagnosis: a systematic review. Radiol Artif Intell. 2022; 4(3): e210064. doi: 10.1148/ryai.210064.
23. Urakawa T, Tanaka Y, Goto S, Matsuzawa H, Watanabe K, Endo N. Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. Skeletal Radiol. 2019; 48(2): 239-244. doi: 10.1007/s00256-018-3016-3.
24. Cheng CT, Ho TY, Lee TY, Chang CC, Chou CC, Chen CC, et al. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. Eur Radiol. 2019; 29(10): 5469-5477. doi: 10.1007/s00330-019-06167-y.
25. Krogue JD, Cheng KV, Hwang KM, Toogood P, Meinberg EG, Geiger EJ, et al. Automatic hip fracture identification and functional subclassification with deep learning. Radiol Artif Intell. 2020; 2(2): e190023. doi: 10.1148/ryai.2020190023.
26. Mutasa S, Varada S, Goel A, Wong TT, Rasiej MJ. Advanced deep learning techniques applied to automated femoral neck fracture detection and classification. J Digit Imaging. 2020; 33(5): 1209-1217. doi: 10.1007/s10278-020-00364-8.
27. Yamada Y, Maki S, Kishida S, Nagai H, Arima J, Yamakawa N, et al. Automated classification of hip fractures using deep convolutional neural networks with orthopedic surgeon-level accuracy: ensemble decision-making with antero-posterior and lateral radiographs. Acta Orthop. 2020; 91(6):699-704. doi: 10.1080/17453674.2020.1803664.
28. Suzuki T, Maki S, Yamazaki T, Wakita H, Toguchi Y, Horii M, et al. Detecting distal radial fractures from wrist radiographs using a deep convolutional neural network with an accuracy comparable to hand orthopedic surgeons. J Digit Imaging. 2022; 35(1):39-46. doi: 10.1007/s10278-021-00519-1.
29. Thian YL, Li Y, Jagmohan P, Sia D, Chan VEY, Tan RT. Convolutional neural networks for automated fracture detection and localization on wrist radiographs. Radiol Artif Intell. 2019; 1(1): e180001. doi: 10.1148/ryai.2019180001.
30. Gan K, Xu D, Lin Y, Shen Y, Zhang T, Hu K, et al. Artificial intelligence detection of distal radius fractures: a comparison between the convolutional neural network and professional assessments. Acta Orthop. 2019; 90(4):394–400. doi: 10.1080/17453674.2019.1600125
31. Oude Nijhuis KD, Barvelink B, Prijs J, Zhao Y, Liao Z, Jaarsma RL, et al. An open source convolutional neural network to detect and localize distal radius fractures on plain radiographs. Eur J Trauma Emerg Surg. 2025; 51(1):26. doi: 10.1007/s00068-024-02731-4.
32. Lindsey R, Daluiski A, Chopra S, Lachapelle A, Mozer M, Sicular S, et al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad Sci U S A. 2018; 115(45):11591-11596. doi: 10.1073/pnas.1806905115.
33. Hendrix N, Scholten E, Vernhout B, Bruijnen S, Maresch B, de Jong M, et al. Development and validation of a convolutional neural network for automated detection of scaphoid fractures on conventional radiographs. Radiol Artif Intell. 2021; 3(4): e200260. doi: 10.1148/ryai.2021200260.

34. Hendrix N, Hendrix W, van Dijke K, Maresch B, Maas M, Bollen S, et al. Musculoskeletal radiologist-level performance by using deep learning for detection of scaphoid fractures on conventional multi-view radiographs of hand and wrist. Eur Radiol. 2023; 33(3):1575-1588. doi: 10.1007/s00330-022-09205-4.
35. Ozkaya E, Topal FE, Bulut T, Gursoy M, Ozuysal M, Karakaya Z. Evaluation of an artificial intelligence system for diagnosing scaphoid fracture on direct radiography. Eur J Trauma Emerg Surg. 2022; 48(1):585-592. doi: 10.1007/s00068-020-01468-0.
36. Chung SW, Han SS, Lee JW, Oh KS, Kim NR, Yoon JP, et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. Acta Orthop. 2018; 89(4): 468-473. doi: 10.1080/17453674.2018.1453714.
37. Prijs J, Liao Z, To MS, Verjans J, Jutte PC, Stirler V, et al. Development and external validation of automated detection, classification, and localization of ankle fractures: inside the black box of a convolutional neural network (CNN). Eur J Trauma Emerg Surg. 2023;49(2):1057-1069. doi: 10.1007/s00068-022-02136-1.
38. Olczak J, Emilson F, Razavian A, Antonsson T, Stark A, Gordon M. Ankle fracture classification using deep learning: automating detailed AO Foundation/Orthopedic Trauma Association (AO/OTA) 2018 malleolar fracture identification reaches a high degree of correct classification. Acta Orthop. 2021; 92(1):102-108. doi: 10.1080/17453674.2020.1837420.
39. Weikert T, Noordtzij LA, Bremerich J, Stieltjes B, Parmar V, Cyriac J, et al. Assessment of a deep learning algorithm for the detection of rib fractures on whole-body trauma computed tomography. Korean J Radiol. 2020; 21(7):891–899. doi: 10.3348/kjr.2019.0653
40. Husarek J, Fuchss AMC, Ruder TD, Mougiakakou S, Exadaktylos A, Wahedi K, et al. Artificial intelligence for fracture detection on computed tomography: a comprehensive systematic review and meta-analysis of diagnostic test accuracy in non-commercial and commercial solutions. Emerg Radiol. 2026; 33(2):351-375. doi: 10.1007/s10140-026-02437-7.
41. Hong N, Cho SW, Shin S, Lee S, Jang SA, Roh S, et al. Deep-learning-based detection of vertebral fracture and osteoporosis using lateral spine X-ray radiography. J Bone Miner Res. 2023; 38(6):887-895. doi: 10.1002/jbmr.4814.
42. Dupuis M, Delbos L, Veil R, Adamsbaum C. External validation of a commercially available deep learning algorithm for fracture detection in children. Diagn Interv Imaging. 2022; 103(3):151-159. doi: 10.1016/j.diii.2021.10.007.
43. Regnard NE, Lanseur B, Ventre J, Ducarouge A, Clovis L, Lassalle L, et al. Assessment of performances of a deep learning algorithm for the detection of limbs and pelvic fractures, dislocations, focal bone lesions, and elbow effusions on trauma X-rays. Eur J Radiol. 2022; 154:110447. doi: 10.1016/j.ejrad.2022.110447.
44. Anderson PG, Baum GL, Keathley N, Sicular S, Venkatesh S, Sharma A, et al. Deep learning assistance closes the accuracy gap in fracture detection across clinician types. Clin Orthop Relat Res. 2023; 481(3):580-588. doi: 10.1097/CORR.0000000000002385.
45. Cohen M, Puntonet J, Sanchez J, Kierszbaum E, Crema M, Soyer P, et al. Artificial intelligence vs. radiologist: accuracy of wrist fracture detection on radiographs. Eur Radiol. 2023; 33(6):3974-3983. doi: 10.1007/s00330-022-09349-3.
46. 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: are these externally validated and ready for clinical application? Bone Jt Open. 2021; 2(10):879–885. doi: 10.1302/2633-1462.210.BJO-2021-0133
47. Jeon YD, Jung KH, Kim MS, Kim H, Yoon DK, Park KB. Clinical validation of artificial intelligence-based preoperative virtual reduction for Neer 3- or 4-part proximal humerus fractures. BMC Musculoskelet Disord. 2024; 25(1):669. doi: 10.1186/s12891-024-07798-z.
48. Farber SH, Pacult MA, Godzik J, Walker CT, Turner JD, Porter RW, et al. Robotics in spine surgery: a technical overview and review of key concepts. Front Surg. 2021; 8:578674. doi: 10.3389/fsurg.2021.578674.

49. Tan JR, Gao Y, Raghuraman R, Ting D, Wong KM, Cheng LT, et al. Application of deep learning algorithms in classification and localization of implant cutout for the postoperative hip. Skeletal Radiol. 2025; 54(1):67-75. doi: 10.1007/s00256-024-04692-6.
50. 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. doi: 10.2196/59607
51. 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(16):5876. doi: 10.3390/jcm14165876.
52. Baker HP, Dwyer E, Kalidoss S, Hynes K, Wolf J, Strelzow JA. ChatGPT's ability to assist with clinical documentation: a randomized controlled trial. J Am Acad Orthop Surg. 2024; 32(3):123-129. doi: 10.5435/JAAOS-D-23-00474.
53. Christy M, Morris MT, Goldfarb CA, Dy CJ. Appropriateness and reliability of an online artificial intelligence platform's responses to common questions regarding distal radius fractures. J Hand Surg Am. 2024; 49(2):91-98. doi: 10.1016/j.jhsa.2023.10.019.
54. Bhattacharyya M, Miller VM, Bhattacharyya D, Miller LE. High rates of fabricated and inaccurate references in ChatGPT-generated medical content. Cureus. 2023; 15(5): e39238. doi: 10.7759/cureus.39238
55. Vogel C, Grimm B, Marmor MT, Sivananthan S, Richter PH, Yarboro S, et al. Wearable sensors in other medical domains with application potential for orthopedic trauma surgery: a narrative review. J Clin Med. 2024; 13(11):3134. doi: 10.3390/jcm13113134.
56. van Leeuwen KG, Schalekamp S, Rutten MJCM, van Ginneken B, de Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol. 2021; 31(6):3797-3804. doi: 10.1007/s00330-021-07892-z.
57. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLoS Med. 2018; 15(11): e1002683. doi: 10.1371/journal.pmed.1002683.
58. Abedeen I, Rahman MA, Prottyasha FZ, Ahmed T, Chowdhury TM, Shatabda S. FracAtlas: a dataset for fracture classification, localization and segmentation of musculoskeletal radiographs. Sci Data. 2023; 10(1):521. doi: 10.1038/s41597-023-02432-4.

How to cite this article: Yadav A, Pushpa, Kumar S. Current applications of artificial intelligence in orthopaedic trauma: a narrative review. *Int. J. Sci. Healthc. Res.* 2026; 11(3): 259-273. DOI: *https://doi.org/10.52403/ijshr.20260330*

******