"GLOBAL HEALTH AND MODERN MEDICINE: INNOVATIONS AND PERSPECTIVES"
International Conference titled

Website:https://worldsiencepub.com/index.php/isngcs September edition,2026

USE OF AN ARTIFICIAL INTELLIGENCE ASSISTANT FOR DETECTING BONE FRACTURES ON RADIOGRAPHIC IMAGES: DIAGNOSTIC ACCURACY AND DECISION SUPPORT

Qiyamova Nasiba Bozor qizi
Hudoynazarov Fayyozbek Muhammadiyevich
Xasanov Azizbek Murod o'g'li
student of Samarkand state medical university
DOI:https://doi.org/10.5281/zenodo.22916263

Annotation

Artificial intelligence (AI) is increasingly being incorporated into medical imaging, particularly radiography, where large numbers of examinations must be interpreted accurately and efficiently. Fracture detection is an important application because subtle cortical breaks, nondisplaced fractures, and overlapping anatomical structures may be difficult to recognize on plain radiographs. This article reviews the role of AI assistants in radiographic fracture detection, the principles of image analysis, potential diagnostic benefits, limitations, and requirements for clinical implementation. Published evidence indicates that modern AI algorithms can achieve high sensitivity and specificity for fracture detection in selected datasets, although performance varies according to anatomical region, image quality, prevalence, algorithm architecture, and the characteristics of the patient population. AI should therefore be regarded as a clinical decision-support tool rather than an autonomous replacement for radiologists or other qualified clinicians. Proper validation, human oversight, data governance, and monitoring are essential for safe implementation.

Keywords

artificial intelligence; radiography; machine learning; deep learning; bone fracture; computer-aided diagnosis; diagnostic accuracy; radiologist; clinical decision support

Introduction

Plain radiography remains one of the most accessible and frequently used imaging methods for evaluating suspected skeletal injury. It is relatively inexpensive, rapid, widely available, and capable of demonstrating many clinically important fractures. Nevertheless, interpretation can be challenging when the fracture line is subtle, the fragments are minimally displaced, the anatomy is complex, or the image contains overlapping structures. Missed fractures may delay immobilization, surgical planning, rehabilitation, and other appropriate treatment. These clinical realities have encouraged the development of computer-aided methods that can assist healthcare professionals during image interpretation.

Artificial intelligence describes computational methods capable of learning patterns from data and producing predictions or classifications. In medical imaging, machine learning and especially deep learning have enabled algorithms to learn complex visual features directly from radiographic images. Convolutional neural networks and related architectures can be trained on large collections of labeled radiographs to distinguish examinations containing fractures from those without fractures, identify suspicious regions, or perform more detailed classification tasks.

© 2026lobal Health and Modern Medicine: Innovations and Perspectives. All Rights Reserved.

A 2024 systematic review and meta-analysis of 100 studies evaluating AI for fracture detection on plain radiographs reported pooled sensitivity of approximately 91.4% and specificity of approximately 92.1% in the studies with available confusion matrices. The reported summary area under the receiver operating characteristic curve was 0.968. These findings demonstrate substantial technical potential, but they do not mean that every AI system performs equally well in routine practice. Model performance is influenced by the data used for development and validation and by the clinical environment in which the system is deployed.

The purpose of this article is to describe how an AI assistant can support fracture detection on radiographs, examine the evidence for diagnostic performance, and discuss the methodological, clinical, and ethical requirements for responsible use.

Materials and Methods

This article was prepared as a narrative scientific review based on peer-reviewed literature concerning artificial intelligence, machine learning, deep learning, and fracture detection on plain radiographs. Particular attention was given to systematic reviews and meta-analyses evaluating diagnostic accuracy, as well as authoritative guidance concerning the clinical governance of AI in healthcare. The literature was considered with emphasis on sensitivity, specificity, receiver operating characteristics, dataset composition, external validation, human-AI interaction, and limitations relevant to clinical implementation.

The main analytical concepts considered were true-positive, false-positive, true-negative, and false-negative results. Sensitivity represents the proportion of actual fractures correctly identified by the system, whereas specificity represents the proportion of examinations without fracture that are correctly classified as negative. Because the clinical consequences of missed fractures and unnecessary additional assessment are different, both measures should be considered together. The article also considers whether AI is used as a second reader, triage system, abnormality detector, or integrated decision-support component.

No original patient dataset was collected or analyzed for this review. Therefore, the article does not claim an original diagnostic accuracy estimate. Published performance values are presented as evidence from previously conducted studies and should not be interpreted as performance of a particular local AI system.

Principles of AI Analysis of Radiographic Images

An AI fracture-detection system generally begins with digital radiographic images stored in a standardized imaging environment. The images may be processed to normalize intensity, remove irrelevant information, and ensure a suitable input format. During model development, radiographs are paired with reference labels derived from radiologist reports, expert consensus, follow-up imaging, surgical findings, or other accepted reference standards. The algorithm then learns statistical relationships between image features and the target label.

Deep learning models are particularly useful because they can learn hierarchical visual representations. Early network layers may identify basic image structures such as edges and intensity transitions, while deeper layers can combine these features into patterns associated with cortical discontinuity, abnormal lucency, displacement, angulation, or other signs of fracture. Depending on the intended task, the output may be a probability of fracture, a bounding box around a suspicious region, a heat map, or a classification for a specific anatomical region.

In a clinical setting, an AI assistant can operate in several ways. It may analyze every radiograph and flag examinations considered suspicious, helping prioritize cases. It may act as a second reader by

© 2026lobal Health and Modern Medicine: Innovations and Perspectives. All Rights Reserved.

highlighting areas that deserve additional attention. It may also provide an explanation-oriented visualization, such as a heat map, to indicate where the model detected features associated with fracture. These functions can potentially improve workflow while preserving the radiologist's role in final interpretation.

Diagnostic Performance and Clinical Value

The strongest evidence currently available suggests that AI can detect many fractures on plain radiographs with high diagnostic performance under study conditions. The 2024 systematic review and meta-analysis by Nowroozi and colleagues included 100 studies and found pooled sensitivity of 91.43% and specificity of 92.12% among studies with sufficient data for pooled analysis. The performance was not uniform across all anatomical regions or study designs. The review also found associations between algorithm performance and factors such as anatomical site, publication period, and model architecture.

Wrist fracture detection is an especially active area of investigation because distal radius and scaphoid fractures can sometimes be subtle. A 2024 systematic review and meta-analysis identified 20 eligible studies evaluating AI for wrist fracture detection and compared AI performance with human readers. Such evidence is important because it moves the discussion beyond technical feasibility toward the clinically relevant question of how AI performs in relation to established interpretation workflows.

The practical value of AI is not limited to numerical accuracy. A useful assistant may reduce the probability that a subtle abnormality is overlooked, help prioritize examinations in busy emergency departments, and provide an additional layer of quality control. However, the benefit depends on integration into a workflow that allows clinicians to review the original images and consider the patient's symptoms, physical examination, mechanism of injury, and previous imaging.

AI may be particularly useful when the workload is high or when radiographic findings are difficult to recognize. At the same time, high sensitivity does not guarantee clinical safety. False-positive results can increase unnecessary additional review or imaging, while false-negative results can create false reassurance. Therefore, clinical effectiveness should be evaluated using patient-centered and workflow-related outcomes in addition to technical accuracy.

Human-AI Collaboration

The most appropriate role of an AI assistant is generally to support, rather than replace, qualified clinical judgment. Radiographic diagnosis is not based solely on pattern recognition. Clinicians integrate image findings with the patient's age, mechanism of trauma, symptoms, physical examination, prior history, and additional imaging when necessary. An AI model usually receives only a limited subset of this information.

Human-AI collaboration can be structured as a second-reader model in which the clinician first interprets the image and then compares the interpretation with AI output. Another approach is AI-first triage, in which examinations predicted to contain abnormalities are prioritized. Each workflow has potential advantages and risks. If clinicians become excessively dependent on an automated recommendation, automation bias may occur. Conversely, if AI findings are ignored because they are presented poorly or generate too many alerts, the system may provide little practical benefit.

For this reason, interface design is clinically important. The system should clearly communicate that its output is decision support, show the relevant anatomical location when possible, and make it easy for the clinician to review the original radiograph. AI performance should also be monitored after

© 2026lobal Health and Modern Medicine: Innovations and Perspectives. All Rights Reserved.

deployment because changes in equipment, imaging protocols, patient populations, and clinical prevalence can affect model behavior.

Limitations and Sources of Error

AI performance reported in research studies may not directly translate to a particular hospital. One major limitation is dataset shift. An algorithm trained on images from one institution may encounter different radiographic equipment, positioning practices, exposure parameters, image preprocessing, and patient demographics in another institution. These differences can reduce generalizability.

Another concern is the quality of reference labels. If training data contain incorrect or inconsistent labels, the model may learn patterns that do not accurately represent the intended clinical diagnosis. The same problem can affect evaluation datasets. Strong reference standards and independent external validation are therefore essential.

Fractures may also be difficult to classify when there is severe overlap, poor positioning, motion artifact, incomplete anatomy, or unusual pathology. Pediatric patients, postoperative changes, pathological fractures, and rare fracture patterns may require separate validation because their radiographic appearance differs from common adult traumatic fractures.

Ethical and governance issues must also be considered. The World Health Organization emphasizes that AI in healthcare should be developed and used with attention to human rights, accountability, transparency, equity, and appropriate human oversight. In addition, patient privacy and secure handling of medical images are essential. AI systems should be continuously evaluated after implementation rather than treated as permanently reliable once initial validation has been completed.

Discussion

The development of AI assistants for radiographic fracture detection represents a significant intersection between medical imaging and computational medicine. The available literature demonstrates that AI can recognize fracture-related patterns at a level that may be clinically useful in selected settings. However, diagnostic accuracy alone does not establish that an AI system improves patient outcomes. A clinically valuable system must fit the workflow, provide understandable outputs, operate safely, and maintain performance across relevant populations.

An important distinction should be made between an algorithm's performance in a controlled research dataset and its performance in routine practice. Retrospective datasets may have a higher prevalence of fractures than a general emergency population, which can influence predictive values. External prospective evaluation is therefore important. Developers and healthcare institutions should assess sensitivity, specificity, calibration, false-negative cases, false-positive cases, turnaround time, and the effect of AI on clinical decisions.

Regulatory oversight is another important component. The United States Food and Drug Administration maintains information on AI-enabled medical devices and evaluates the safety and effectiveness of devices according to their intended uses. The regulatory framework recognizes that different AI applications, such as triage, diagnosis, risk assessment, and image processing, can require different evaluation approaches. This reinforces the principle that an AI assistant should be validated for its specific clinical purpose rather than judged solely on general claims about AI.

For healthcare institutions considering implementation, a structured pathway is appropriate. First, the intended clinical problem should be defined precisely. Second, the algorithm should undergo local technical and clinical validation. Third, users should receive training on interpretation of AI output and recognition of system limitations. Fourth, performance should be monitored continuously, with a

© 2026lobal Health and Modern Medicine: Innovations and Perspectives. All Rights Reserved.

mechanism for reporting errors and adverse events. Finally, patient privacy, cybersecurity, accountability, and equitable access should be incorporated into governance.

AI should therefore be understood as an additional information source within a diagnostic system. Its greatest value may arise when it complements professional expertise, reduces repetitive workload, and draws attention to findings that might otherwise be overlooked. The final clinical decision should remain based on comprehensive patient assessment and appropriate professional responsibility.

Conclusion

Artificial intelligence assistants have substantial potential to support fracture detection on plain radiographs. Current evidence demonstrates high diagnostic performance in many research settings, with systematic-review data showing pooled sensitivity and specificity above 90% in a large body of published studies. Nevertheless, performance varies by anatomical site, dataset, algorithm, image quality, and clinical environment. AI should not be treated as an autonomous substitute for radiologists or other qualified clinicians. Safe implementation requires external validation, human oversight, transparent communication of limitations, protection of patient data, and continuous monitoring. Future research should focus not only on improving algorithmic accuracy but also on prospective clinical evaluation, workflow integration, equity, interpretability, and measurable effects on patient care.

References

1. Nowroozi A, Salehi MA, Shobeiri P, et al. Artificial intelligence diagnostic accuracy in fracture detection from plain radiographs and comparing it with clinicians: a systematic review and meta-analysis. Clinical Radiology. 2024;79(8):579–588.
2. Suen K, et al. Accuracy of wrist fracture detection on radiographs by artificial intelligence compared to human clinicians: a systematic review and meta-analysis. European Journal of Radiology. 2024;174:111593.
3. World Health Organization. Ethics and Governance of Artificial Intelligence for Health. Geneva: World Health Organization; 2021.
4. World Health Organization. Ethics and Governance of Artificial Intelligence for Health: WHO Guidance Executive Summary. Geneva: World Health Organization; 2021.
5. U.S. Food and Drug Administration. Artificial Intelligence-Enabled Medical Devices. Center for Devices and Radiological Health. 2026.
6. U.S. Food and Drug Administration. Artificial Intelligence Program: Research on AI/ML-Based Medical Devices. Center for Devices and Radiological Health.
7. U.S. Food and Drug Administration. Regulatory Evaluation of New Artificial Intelligence Uses for Improving and Automating Medical Practices. Center for Devices and Radiological Health.
8. U.S. Food and Drug Administration. Artificial Intelligence and Machine Learning in Software as a Medical Device. Regulatory science and medical-device guidance resources.

© 2026lobal Health and Modern Medicine: Innovations and Perspectives. All Rights Reserved.