Review

SPINE RESEARCH

OPEN

Application of artificial intelligence in the diagnosis of spinal infectious diseases

A review of diagnostic and differential diagnostic performance

Haoming Sun[a], Chunwang Jia[a], Hongwei Wang[a], Xianghe Wang[a], Haocheng Xu[a], Xiaosheng Ma[a], Feizhou Lyu[a], Jianyuan Jiang[a], Shuo Yang[a,*], Hongli Wang[a,*]

Abstract

Spinal infectious diseases can lead to severe disability and adverse outcomes without timely and accurate diagnosis and treatment. Artificial intelligence (AI) and deep learning (DL) have brought new progress to imaging-based diagnosis of spinal infectious diseases. This review aimed to evaluate the diagnostic and differential diagnostic performance of AI models for spinal infectious diseases. A systematic literature search was performed in PubMed, Web of Science, and Embase for studies published between 2018 and 2026. Eligible studies focused on AI or DL models for detecting or differentiating spinal infectious diseases, and data including sensitivity, specificity, accuracy, and area under the receiver operating characteristic curve were extracted and analyzed. A total of eight studies were included. AI models showed generally favorable efficacy in the identification and differential diagnosis of spinal infections. For differentiating tuberculous spondylitis (TS) from brucellar spondylitis (BS), sensitivity ranged from 0.588 to 0.940 and specificity from 0.679 to 0.964. Most models were based on magnetic resonance imaging (MRI) or computed tomography (CT) and achieved high diagnostic accuracy, but the majority were retrospective, single-centered, and lacked large-scale external validation. AI serves as a reliable auxiliary tool for the rapid and accurate diagnosis of spinal infectious diseases, especially when combined with advanced imaging techniques. However, further multicenter, prospective validation is required before widespread clinical application. AI should complement rather than replace clinical expertise in routine diagnostic workflows.

Abbreviations: AI = artificial intelligence, AUC = area under the curve, CT = computed tomography, DL = deep learning, MRI = magnetic resonance imaging, ROC = receiver operating characteristic.

Keywords: artificial intelligence, deep learning, diagnose, infections, magnetic resonance imaging, spinal, spondylitis

1. Introduction

Spinal infections represent a group of rare conditions affecting vertebral bodies, intervertebral discs, paraspinal soft tissues, epidural space, meninges, and spinal cord. The causal factors, clinical presentations, and imaging features are a challenge because of the difficulty to differentiate them from other conditions, such as degenerative and inflammatory disorders and spinal neoplasm.[1] On the basis of the pathogen, bacterial spinal infections can be either pyogenic or granulomatous, such as tuberculosis or brucellosis.[2] According to the classification of pathogens, the types of spinal infection include pyogenic bacterial infection, specific infections such as tuberculosis, brucellosis, and fungal infections, as well as other rare pathogenic infections such as viral and parasitic infections.

Radiography, computed tomography (CT), magnetic resonance imaging (MRI), and nuclear imaging studies are complementary imaging modalities for the diagnosis and evaluation of spinal infections. Radiography is often the initial modality used for assessing back or neck pain due to its widespread availability and low cost.[3]

Artificial intelligence (AI) models can be classified from multiple perspectives. According to learning methods, they are mainly divided into supervised learning, unsupervised learning, semi-supervised learning, self-supervised learning, and reinforcement learning. According to

HS and CJ contributed equally to this work.

[a] Department of Orthopaedics, Huashan Hospital, Fudan University, Shanghai, China.

* Correspondence: Hongli Wang, Department of Orthopaedics, Huashan Hospital, Fudan University, No.12 Middle Urumqi Road, Shanghai 200040, China (e-mail: wanghongli0212@163.com); Shuo Yang, Department of Orthopaedics, Huashan Hospital, Fudan University, No.12 Middle Urumqi Road, Shanghai 200040, China (e-mail: yangshuo429@126.com).

© The Author(s) 2026. Published by Wolters Kluwer Health, LLC. on behalf of Higher Education Press. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Spine Research (2026) 2:3;147–153

Received: 14 October 2025 / Accepted: 31 March 2026

Published online 6 August 2026

http://dx.doi.org/10.1097/br9.0000000000000043

technical architectures, they can be categorized into traditional machine learning models and deep learning (DL) models centered on neural networks. DL further includes typical structures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformers. According to application tasks, they can be divided into natural language processing, computer vision, speech processing, multimodal, and recommendation system models. According to functional characteristics, they can be distinguished into discriminative models for judgment and prediction, and generative models for content generation. In addition, based on model scale and application scenarios, they can also be divided into dedicated small models for single tasks, large models with general capabilities, and domain-specific models for medical, financial, and other vertical fields, which together form the technical system of AI. DL-based image analysis not only matches the performance of traditional radiological approaches but, in specific clinical scenarios, has been shown to rival or even exceed expert-level diagnostic accuracy.[4]

This review aims to analyze the current literature on the use of AI and DL models for the detection of spinal infectious diseases. By examining the performance of these AI models, we seek to clarify their advantages and limitations, as well as explore their prospects and development directions, to better meet the needs of clinical practice and scientific research.

2. Material and methods

2.1. Literature search strategy

A systematic search was conducted in the PubMed, Web of Science, and Embase databases. It was conducted to identify studies on applying AI and DL models for detecting spinal infectious diseases. The search strategy included a combination of Medical Subject Headings terms and keywords, such as "artificial intelligence," "deep learning," "spondylitis," "infections," and "diagnostic imaging."

2.2. Study selection

The initial search yielded 11 records. After screening the titles and abstracts, one document was excluded. After analyzing the specific content of the remaining eight articles, two more documents were excluded. The following inclusion criteria were used: The study population comprised patients with suspected or confirmed spinal infections. The intervention involved the use of AI or DL models for the detection of spinal infections. The study compared the performance of AI models with that of radiologists or mature AI diagnostic models. The study outcome measures were diagnostic accuracy metrics. The study designs included diagnostic studies, prospective studies, retrospective studies, and randomized controlled trials evaluating the performance of AI models. The excluded documents were those that did not involve spinal infectious disease detection, discussed generative AI models (ChatGPT) unrelated to radiology, and focused solely on image dataset algorithms, data science, or mathematical models. The literature screening process is presented in Figure 1.

2.3. Data extraction

The data included study characteristics: author, year, country of origin, sample size, and population demographics; imaging modality: MRI, CT; AI model type: models such as CNN; performance metrics: area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy, recall, and F-score.

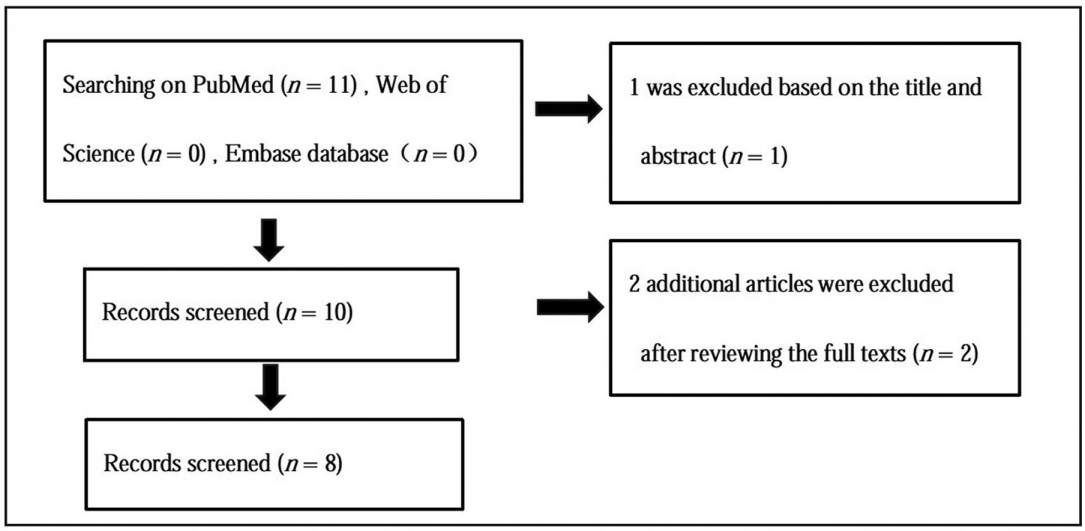


Figure 1. Literature screening process.

3. Results

3.1. Study selection

Initially, a literature search was conducted using the keywords "spondylitis," "infections," "diagnostic imaging," and "artificial intelligence," which retrieved 11 articles. After screening the titles and abstracts, two articles were excluded. Additionally, another article was excluded because it used a generative AI model unrelated to imaging. Eventually, a total of eight articles were included in this review.

3.2. Study characteristics

The included studies were published between 2018 and 2026, with sample sizes ranging from 94 to 1,027. Most of the studies (six out of eight) used MRI as the primary imaging method, while one study used CT as the main imaging method. These studies were mainly conducted in East Asia, including six in China, one in Republic of Korea, and one in Japan. CNNs were the dominant type of AI models. Table 1 presents the basic information of these studies, including nationality, year, sample size, imaging techniques used, names of AI models, and so on.

3.3. Performance of AI models

The AI models assessed in this review demonstrated a wide range of diagnostic performance with variability in AUC, sensitivity, specificity, accuracy, recall, and F-score.

3.3.1. Sensitivity. Sensitivity is an indicator used to evaluate the strength of a model's ability to avoid missed diagnoses, with a range from 0.588 to 0.940. ResNet101 exhibits the lowest sensitivity. After combining the radiomics model, the clinical model, and the DL model into the DLRN model (a model for differential diagnosis of tuberculous spondylitis [TS] and pyogenic spondylitis [PS]), its sensitivity reaches 0.588. On the contrary, in the model for differential diagnosis of PS and Modic change, CNN alone achieved a sensitivity of 0.940.

3.3.2. Specificity. Specificity is an indicator for evaluating the ability to avoid misdiagnosis, and its range is between 0.679 and 0.964. In the model for differential diagnosis of TS and PS, the deep convolutional neural network (DCNN) showed the lowest specificity, which was 0.625. Instead, the DLRN model (ResNet101), which had the lowest sensitivity, achieved the highest specificity of 0.964.

3.3.3. Accuracy. Overall, the accuracy ranges from 0.732 to 0.942. Among these models, the VGG19 model, specifically designed for distinguishing TS from brucellar spondylitis (BS), achieved the lowest accuracy of 0.732, while the CBAM-ResNeXt model, which is specialized in the differential diagnosis of spinal tuberculosis and BS, obtained the highest accuracy of 0.942.

3.3.4. Recall. The recall plays an important role in determining whether there is a missed diagnosis. Not all the studies mentioned provide data on the recall rate; however, among them, the CBAM-ResNeXt model, which does provide this data, has a recall rate of 0.937. This indicates that the probability of missed diagnosis with this model is extremely low.

3.3.5. F1-score. The F1-score is a comprehensive metric that combines precision and recall. It can more objectively measure a model's ability to balance between "reducing misjudgments" and "not missing target cases." In two studies on the differential diagnosis of spinal tuberculosis and spinal brucellosis, the CBAM-ResNeXt model achieved a score of 0.938, while the VGG19 combined model obtained a score of 0.805. This indicates that both models play an important role in reflecting the true positive rate and reducing the false positive rate.

3.3.6. Area under the ROC curve. The AUC value reflects the discriminative ability of a model. A value of 0.5 indicates random classification, which is meaningless. When the value is greater than 0.5, the closer it is to 1.0, the stronger the discriminative ability of the model. In the model for differential diagnosis of spinal tuberculosis and BS, the VGG19 combined model achieves the highest AUC value of 0.973, the CBAM-ResNeXt model has an AUC value of 0.953, and the Rad + ResNet34-SR model has an AUC value of 0.952. Additionally, the CNN model for differential diagnosis of spinal Modic changes and spinal pyogenic infection obtains an AUC value of 0.950. The VGG19 model achieves the lowest AUC value

Table 1

Basic characteristics of the included studies

Study	Year	Nationality	Sample size	Population	Imaging	AI model
Wang et al.[5]	2025	China	94	Patients with TS or BS	CT	ResNet34, ResNet18
Chen et al.[6]	2024	China	383	Patients with TS or BS	MRI	VGG19, VGG16, ResNet18, DenseNet121
Wang et al.[4]	2025	China	310	Patients with TS or BS	MRI	CBAM-ResNeXt
Yang et al.[7]	2025	China	147	Patients with TS or PS	MRI	ResNet101
Kim et al.[8]	2018	Korea	161	Patients with TS or PS	MRI	DCNN
Mukaihata et al.[9]	2022	Japan	100	Patients with PS or Modic change	MRI	CNN
Song et al.[10]	2026	China	235	Patients with TS or BS	MRI	VGG19
Sen et al.[11]	2026	China	1,027	Patients with TS	CT	ResNet50

AI = artificial intelligence, BS = brucellar spondylitis, CNN = convolutional neural network, CT = computed tomography, DCNN = deep convolutional neural network, MRI = magnetic resonance imaging, PS = pyogenic spondylitis, TS = tuberculous spondylitis.

Table 2

The performance of AI models

Study	AI model	Function	AUC	Sensitivity	Specificity	Accuracy	Recall	F1-score	Limitations
Wang et al.[5]	ResNet34, ResNet18	Differentiating TS and BS with CT	0.952	0.909	0.941	0.929			Lack of prospective, multi-center, and confirmatory studies
Chen et al.[6]	VGG19, VGG16, ResNet18, DenseNet121	Differentiating TS and BS with MRI	0.973	0.900	0.944	0.921		0.923	Contrast agent allergy and lack of external validation
Wang et al.[4]	CBAM-ResNeXt	Differentiating TS and BS with MRI	0.953		0.948	0.942	0.937	0.938	Limited sample size and restriction to only two types of infections
Yang et al.[7]	ResNet101	Differentiating TS and PS with MRI	0.859	0.588	0.964	0.833			Limited sample size
Kim et al.[8]	DCNN	Differentiating TS and PS with MRI	0.802	0.850	0.679	0.764			Lack of external validation and prospective studies
Mukaihata et al.[9]	CNN	Differentiating PS and Modic with MRI	0.950	0.940	0.900	0.920			Not all patients with Modic changes were diagnosed using biopsy cultures
Song et al.[10]	VGG19	Differentiating TS and BS with MRI	0.765			0.732		0.805	Limited types of infections
Sen et al.[11]	ResNet50	Differentiating TS with CT	0.941	0.843		0.790			Selection bias

AUC = area under the ROC curve, BS = brucellar spondylitis, CNN = convolutional neural network, CT = computed tomography, DCNN = deep convolutional neural network, MRI = magnetic resonance imaging, PS = Pyogenic Spondylitis, TS = Tuberculous Spondylitis.

of 0.765. This data indicates that these models possess strong capabilities and significant clinical potential in differential diagnosis.

The performance of AI models is listed in Table 2.

4. Discussion

This systematic review assesses the diagnostic performance of AI models in detecting spinal infection diseases based on eight studies published between 2018 and 2026. It includes three studies on differentiating spinal tuberculosis from BS, two studies on differentiating spinal tuberculosis from pyogenic spondylitis, and one study on differentiating pyogenic spondylitis from spinal Modic changes. Diagnostic imaging mainly uses MRI, with two studies employing CT scanning. MRI is considered the gold standard due to its superior ability to visualize soft tissue structures and assess the extent of infection.[12] Among MRI applications, the T2 sequence is most commonly used, while some studies also adopt multisequence imaging. The results show that the diagnostic accuracy of multisequence models is higher than that of single-sequence models. The study using CT applied SR-CT (super-resolution CT) to perform fine processing on patients' images, and the results also indicated that the higher the quality of the images, the more accurate the diagnosis of the AI model. Meanwhile, some AI models have incorporated clinical data for training, and the outcomes of integrating clinical data vary across different studies. In Chen et al.[6] after the VGG19 combined model[6] was trained with the inclusion of age, gender, site of spinal infection, and number of involved vertebrae, the model's AUC value decreased from the original 0.973 to 0.912, its sensitivity dropped from 0.900 to 0.850, and its specificity fell from 0.944 to 0.861. This study holds that this may be because the DL models we constructed had a strong ability to diagnose TS and BS, and relatively simple clinical information did not benefit the DL models[6]. However, in Yang et al.[7] when DLR[7] model was combined with clinical data, including details such as routine blood test results and a variety of imaging features from the DLRN[8] model, its diagnostic performance achieved a significant improvement. These two distinctly different results may indicate that whether the inclusion of clinical data in an AI model improves its diagnostic performance could be influenced by a variety of factors, such as the performance of the model itself and the correlation between the clinical data and the corresponding disease. This requires further research and verification.

AI integration into clinical workflows offers several potential advantages. In particular, the application of AI in radiology can augment diagnostic processes by providing automated decision support, improving diagnostic speed, and reducing human error.[13] AI is a vast field that includes a variety of methodologies, such as computer vision, natural language processing, and machine learning.[14] AI can increase the value that musculoskeletal imagers provide to their patients and to referring clinicians by improving image quality, patient centricity, imaging efficiency, and diagnostic accuracy.[15] Although the

number of AI models specifically developed for the diagnosis of spinal infectious diseases is currently relatively small, AI has already played a significant role in the diagnosis of other orthopedic conditions, such as automatically detecting and classifying lumbar spinal canal, lateral recess, and neural foraminal stenoses on lumbar spine MRI scans,[16,17] classification of vertebral fractures,[18,19] evaluation of intervertebral disc degeneration,[20] and differential diagnosis of spinal tumors and inflammatory lesions.[9,21] It can thus be seen that the application of AI in the future medical field is absolutely indispensable and will be involved in the diagnosis and treatment of more diseases.

AI and machine learning technology play a crucial role in drug discovery and development.[22] The most common applications of AI in drug treatment have to do with matching patients to their optimal drug or combination of drugs, predicting drug-target or drug–drug interactions, and optimizing treatment protocols.[23] AI algorithms can analyze medical imaging data, such as X-rays, CT scans, and MRI, to detect abnormalities and assist in early disease detection.[24] This capability translates to a significantly reduced chance of misdiagnosis and leads to better patient outcomes.[25] Also, AI-powered virtual applications can analyze an individual's health data,[25] including their genetics,[26] medical history,[27] lifestyle factors,[28] and current health status, to create personalized treatment plans.[25] AI also holds significant importance in the selection of drug targets. AI has been used to predict drug-target interactions,[29] measure the binding affinity of a drug,[30] and select and validate targets.[31] By exploring interactions between drugs and their targets, AI can predict combinations that exhibit enhanced therapeutic effects while minimizing adverse reactions.[32]

AI can also play an essential role in training and education. Radiology trainees and early-career clinicians could benefit from AI-assisted tools that provide instant feedback on image interpretation.[13] For clinicians and medical students, the emergence of AI can help them learn the imaging features of more diseases even when they have not been exposed to many real cases. Moreover, when junior doctors are not sensitive to the early imaging changes of some spinal infections, AI can also assist in the early identification of diseases to avoid serious consequences.

AI is significantly transforming surgery by enhancing precision, decision-making, and patient outcomes.[33] Before surgery, AI analyzes medical records and imaging, offering personalized treatment plans and risk assessments to optimize planning.[34,35] During surgery, AI monitors in real-time, identifies key structures, and helps surgeons perform precise operations, reducing errors.[36–38] After surgery, AI monitors recovery, predicts and prevents complications, and improves care quality.[39,40] Surgeons, working with data scientists, use AI to capture and interpret new clinical data. This improves surgical techniques and enhances patient care.[41–43]

Although many AI models currently match or even surpass professional doctors in diagnostic performance, especially in the field of medical imaging, doctors must be proficient in utilizing AI rather than allowing it to replace their own thinking or remaining unaware of the underlying logic behind AI's judgments. One concern with AI is the "black box" problem, where the underlying reasoning behind AI outputs is unclear.[44] This requires both doctors and AI models to make progress: AI models should also train and update their standards based on the latest and most accurate diagnostic evidence, while doctors need to provide more accurate data and samples for AI based on actual clinical situations. Only in this way can the healthy development of AI in the medical field be promoted.

This article reviews the research status and clinical application progress of AI models in the diagnosis of spinal infections, but certain limitations still exist:

First, most of the studies included in this review are single-center, retrospective observational studies.

Second, the standards for the construction and validation of AI models in the existing literature are not yet unified.

Third, studies on clinical applicability remain insufficiently in-depth. Most studies focus on indicators such as diagnostic accuracy, while explorations of model interpretability, impact on clinical decision-making, health economic benefits, and integration into real clinical workflows are limited.

In addition, this review did not conduct detailed stratified analyses of subgroup differences across different etiologies, disease stages, and imaging equipment conditions.

5. Conclusion

This article systematically reviews the imaging-based models for diagnosing spinal infectious diseases constructed in eight studies published from 2018 to 2026, among which CNNs are the most common. These models serve different functions, and their diagnostic performance is generally on par with or even superior to that of professional doctors, demonstrating great clinical potential. However, most of these studies have small sample sizes and lack external validation and multicenter validation, which requires further research. For models mainly used for imaging diagnosis, their performance varies significantly after incorporating different types of clinical data. Clinicians should further identify clinical data that have a higher correlation with disease diagnosis and improve the quality of medical images; only then can the performance of AI models be enhanced. AI should not replace doctors but rather assist them in diagnosis and treatment.

Acknowledgements

The authors thank all researchers who provided the original studies included in this review. We appreciate the support from the Department of Orthopaedics, Huashan Hospital, Fudan University, Shanghai, China.

Ethical statement

Not applicable.

Conflicts of interest

Hongli Wang serves as a member of the Editorial Board of this journal. He was not involved in the editorial review or decision-making process for this manuscript. All editorial decisions were made independently by other members of the Editorial Board who have no conflicts of interest.

Funding source

This work was supported by 2024 Shanghai Oriental Talent Program Youth Project (DFYCQN04) and 2025 The "Clinical Famous Doctor Cultivation Project" of Shanghai Medical College, Fudan University (DGF828030-1/005).

Data availability statement

Not applicable.

Author Contributions

HS and CJ—contributed equally to this work, including literature search, data extraction, statistical analysis, and drafting of the manuscript.
HW, XW, HX, XM, FL, and JJ—participated in data collection, literature evaluation, and critical revision of the intellectual content.
SY and HW—contributed to the study conception and design, supervised the entire work, revised the manuscript critically for important intellectual content, and approved the final version to be published.
All authors have read and agreed to the published version of the manuscript.

References

[1] Arbelaez A, Restrepo F, Castillo M. Spinal infections: clinical and imaging features. *Top Magn Reson Imaging*. 2014;23:303–14.
[2] Babic, M, Simpfendorfer, C. Infections of the spine. *Infect Dis Clin N. Am*. 2017;31(2):279–97.
[3] Wangaryattawanich P, Condos AM, Rath TJ. Bacterial and viral infectious disease of the spine. *Magn Reson Imaging Clin N Am*. 2024;32:313–33.
[4] Wang B, Wei J, Wang Z. Development of a deep learning-based MRI diagnostic model for human brucella spondylitis. *Biomed Eng Online*. 2025;24(1):87.
[5] Wang K, Qi L, Li J, Zhang M, Du H. CT-based 3D super-resolution radiomics for the differential diagnosis of brucella vs. Tuberculous spondylitis using deep learning. *Current Med Imag*. 2025;21:e15734056380084.
[6] Chen J, Guo X, Liu X, et al. Differentiation of tuberculous and brucellar spondylitis using conventional MRI-based deep learning algorithms. *Eur J Radiol*. 2024;178:111655.
[7] Yang X, Tian N, Zhang Y, et al. Deep learning radiomics model based on contrast-enhanced MRI for distinguishing between tuberculous spondylitis and pyogenic spondylitis. *Eur Spine J*. 2025;34(9):3702–13.
[8] Kim K, Kim S, Lee YH, et al. Performance of the deep convolutional neural network based magnetic resonance image scoring algorithm for differentiating between tuberculous and pyogenic spondylitis. *Sci Rep*. 2018;8(1):13124.
[9] Mukaihata T, Maki S, Eguchi Y, et al. Differentiating magnetic resonance images of pyogenic spondylitis and spinal modic change using a convolutional neural network. *Spine*. 2023;48:288–294.
[10] Yasin P, Tuersun A, Ashir A, et al. Comprehensive comparative analysis of explainable deep learning model for differentiation of brucellar spondylitis and tuberculous spondylitis through MRI sequences. *Eur J Med Res*. 2026;31(1):146.
[11] Mo S, Liu C, Xue J, et al. Deep learning diagnosis model of spinal tuberculosis based on CT bone window gradient attention mechanism: multi-center study. *Comput Assist Surg*. 2026;31:2599329.
[12] Xu Z, Zhu W, Zhou S, et al. Diagnostic imaging confusion in infectious spondylitis. *J Orthop Surg Res*. 2025;20:496.
[13] Liawrungrueang W, Cholamjiak W, Promsri A, et al. Artificial intelligence for cervical spine fracture detection: a systematic review of diagnostic performance and clinical potential. *Global Spine J*. 2025;15:2547–58.
[14] Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput Sci*. 2021;2:420.
[15] Gyftopoulos S, Lin D, Knoll F, Doshi AM, Rodrigues TC, Recht MP. Artificial intelligence in musculoskeletal imaging: current status and future directions. *AJR Am J Roentgenol*. 2019;213:506–13.
[16] Lim DSW, Makmur A, Zhu L, et al. Improved productivity using deep learning–assisted reporting for lumbar spine MRI. *Radiology*. 2022;305:160–6.
[17] Hallinan JTPD, Zhu L, Yang K, et al. Deep learning model for automated detection and classification of central canal, lateral recess, and neural foraminal stenosis at lumbar spine MRI. *Radiology*. 2021;300:130–8.
[18] Liu B, Jin Y, Feng S, et al. Benign vs malignant vertebral compression fractures with MRI: a comparison between automatic deep learning network and radiologist's assessment. *Eur Radiol*. 2023;33(7):5060–8.
[19] Zhang H, Yuan G, Wang C, et al. Differentiation of benign versus malignant indistinguishable vertebral compression fractures by different machine learning with MRI-based radiomic features. *Eur Radiol*. 2023;33(7):5069–76.
[20] Zheng HD, Sun YL, Kong DW, et al. Deep learning-based high-accuracy quantitation for lumbar intervertebral disc degeneration from MRI. *Nat Commun*. 2022;13:841.
[21] Liu H, Jiao M, Yuan Y. Benign and malignant diagnosis of spinal tumors based on deep learning and weighted fusion framework on MRI. *Insights Imag*. 2022;13(1):87.
[22] Gupta R, Srivastava D, Sahu M, et al. Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers*. 2021;25:1315–60.
[23] Romm EL, Tsigelny IF. Artificial intelligence in drug treatment. *Annu Rev Pharmacol Toxicol*. 2020;60(1):353–69.
[24] Yoon HJ. Medical image analysis using artificial intelligence. *Prog Med Phys*. 2019;30(2):49–58.
[25] Visan AI, Negut I. Integrating artificial intelligence for drug discovery in the context of revolutionizing drug delivery. *Life (Basel)*. 2024;14:233.
[26] Sohail A. Genetic algorithms in the fields of artificial intelligence and data sciences. *Ann Data Sci*. 2021;10:1007–18.
[27] LeeKim SHS. Prospect of artificial intelligence based on electronic medical record. *J Lipid Atheroscler*. 2021;10:282–90.
[28] ContrerasVehi IJ. Artificial intelligence for diabetes management and decision support: literature review. *J Med Internet Res*. 2018;20:e10775.
[29] Yaseen BT. Drug–target interaction prediction using artificial intelligence. *Appl Nanosci*. 2023;13:3335–45.

[30] Kumar R, Sharma A, Siddiqui MH, Tiwari RK. Prediction of drug-plasma protein binding using artificial intelligence based algorithms. *Comb Chem High Throughput Screen*. 2018;21:57–64.
[31] Romeo-Guitart D, Forés J, Herrando-Grabulosa M, et al. Neuroprotective drug for nerve trauma revealed using artificial intelligence. *Sci Rep*. 2018;8:1879.
[32] Rani P, Dutta K, Kumar V, et al. Artificial intelligence techniques for prediction of drug synergy in malignant diseases: past, present, and future. *Comput Biol Med*. 2022;144:105334.
[33] Li H, Han Z, Wu H, et al. Artificial intelligence in surgery: evolution, trends, and future directions. *Int J Surg*. 2025;111:2101–11.
[34] Li J, Tang T, Wu E. RARPKB: a knowledge-guide decision support platform for personalized robot-assisted surgery in prostate cancer. *Int J Surg*. 2024;110(6):3412–24.
[35] Wang XC, Zhang L, Lin JB, et al. Imaging diagnosis and differential diagnosis of extraskeletal osteosarcoma. *BMC Cancer*. 2024;24(1):11.
[36] Bellos T, Manolitsis I, Katsimperis S, et al. Artificial intelligence in urologic robotic oncologic surgery: a narrative review. *Cancers (Basel)*. 2024;16:1775.
[37] Lee KS, Jung SH, Kim DH, et al. Artificial intelligence- and computer-assisted navigation for shoulder surgery. *J Orthop Surg (Hong Kong)*. 2024;32:793515038.
[38] Barone M, De Bernardis R, Persichetti P. Artificial intelligence in plastic surgery: analysis of applications, perspectives, and psychological impact. *Aesth Plast Surg*. 2024;49:1637–9.
[39] Wilson HH, Ma C, Ku D, et al. Deep learning model utilizing clinical data alone outperforms image-based model for hernia recurrence following abdominal wall reconstruction with long-term follow up. *Surg Endosc*. 2024;38:3984–91.
[40] Zhang B, Zhu J, Chen T, et al. Prediction of additional hospital days in patients undergoing cervical spine surgery with machine learning methods. *Comput Assist Surg (Abingdon)*. 2024;29:2345066.
[41] Zhang J, Luo Z, Zhang R, et al. The transition of surgical simulation training and its learning curve: a bibliometric analysis from 2000 to 2023. *Int J Surg*. 2024;110:3326–37.
[42] Sengun B, Iscan Y, Yazici ZA, et al. Utilization of artificial intelligence in minimally invasive right adrenalectomy: recognition of anatomical landmarks with deep learning. *Acta Chir Belg*. 2024;124:492–8.
[43] Balch JA, Abbott KL, Loftus TJ. Critical appraisal of artificial intelligence in robotic surgery. *Surgery*. 2024;176:558–9.
[44] Voter AF, Larson ME, Garrett JW, Yu J-PJ. Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures. *AJNR Am J Neuroradiol*. 2021;42:1550–6.

How to cite this article: Sun H, Jia C, Wang H, Wang X, Xu H, Ma X, Lyu F, Jiang J, Yang S, Wang H. Application of artificial intelligence in the diagnosis of spinal infectious diseases: A review of diagnostic and differential diagnostic performance. *Spine Res*. 2026;2(3):147–153. doi: 10.1097/br9.0000000000000043