Review

SPINE RESEARCH

OPEN

Empowering the future of spinal surgery through digital and intelligent technologies

Chenfei Gao, Tianyu Yao, Tenghui Zhang, Wenyu Zhang, Jianxi Wang, Fazhi Zang, Huajiang Chen*

Abstract

Artificial intelligence (AI) has been widely applied in spinal surgery, contributing significantly to clinical disease diagnosis, surgical treatment decision-making, prognosis prediction, intraoperative intelligent navigation, surgical rehabilitation, and the advancement of surgical instruments. Nevertheless, current research predominantly focuses on evaluating model performance, often neglecting clear indicators of clinical utility. Moreover, several challenges persist, including low-quality datasets, heterogeneity in research reports, insufficient algorithm transparency, and limited clinical application scenarios. Looking ahead, by enhancing the reliability and clinical efficacy of algorithms from multiple perspectives, AI is expected to enable comprehensive management of spinal surgical diseases throughout the preoperative, intraoperative, and postoperative phases.

Keywords: artificial intelligence, diagnosis, prognosis, spine, surgical treatment

1. Overview of artificial intelligence research in spinal surgery

Artificial intelligence (AI) is a scientific discipline dedicated to enabling machines to simulate human intelligence.[1] Algorithmic models are the fundamental technical means for realizing AI. Machine learning (ML), a subfield of AI, enables machines to learn from experience, allowing them to improve their performance on specific tasks based on past experiences or provided data. The main objective of ML is prediction, which involves generating desired output values from features provided by model developers or automatically extracted from training data.[2] Deep learning (DL), an advanced form of ML, has convolutional neural networks at its core. DL automatically extracts structural features and weights from data through multiple layers of processing units, enabling more complex iterative computations that simulate the computational model of the human brain's neural network.[3]

With the rapid development and extensive application of AI technology in spinal surgery, existing research has established an application framework centered on the clinical diagnosis of spinal diseases and covering the entire process of spinal surgical treatment.[4–8] This framework includes decision-making for surgical treatment plans, prognosis prediction, intraoperative intelligent navigation assistance, and postoperative rehabilitation. The scale and quality of medical databases have a significant impact on the performance of AI algorithms, thereby determining the developmental potential and clinical application value of these models. Large-scale databases constructed through standardized data collection, preprocessing, and annotation provide reliable training data, which support the training and parameter optimization of AI models. Training and parameter optimization are critical steps in model development. During the training phase, the algorithm iteratively learns from medical data; in the parameter optimization phase, the trained model analyzes newly input medical data and makes further parameter adjustments based on reference standards. Finally, the output results of the algorithm model in the validation set are transformed into a confusion matrix, and the diagnostic level of the algorithm is quantified by calculating the corresponding evaluation metrics. Meanwhile, the diagnostic performance of the algorithm can be compared with that of spine surgeons or radiologists with different levels of seniority, demonstrating the effectiveness of AI in spinal surgery diagnosis.

2. Applications

2.1. Diagnosis

In the clinical diagnosis of spinal surgical diseases, the main conditions include degenerative diseases, trauma, deformities, inflammation, and tumors. The diagnosis of

---

CG, TY, and TZ contributed to this article equally.

Department of Orthopedics, Spine Center, Shanghai Changzheng Hospital, Naval Medical University, Shanghai, China.

* Correspondence: Huajiang Chen, Department of Orthopedics, Spine Center, Shanghai Changzheng Hospital, Naval Medical University, 415th Feng Yang Road, Shanghai 200003, China (e-mail: spine.chen@163.com).

Spine Research (2025) 11:1(e0003)

Received: 24 April 2025 / Received in final form: 24 April 2025 / Accepted: 28 April 2025

http://dx.doi.org/10.1097/br9.0000000000000003

these conditions mainly relies on imaging examinations. However, challenges such as low efficiency in imaging interpretation and its heavy dependence on physicians' expertise may lead to missed or misdiagnosed cases, especially for lesions that are difficult to detect or evaluate visually, such as subtle fractures, *in-situ* fractures, and incomplete fractures. This introduces a certain degree of uncertainty in clinical diagnoses. A reliable way to reduce these errors is to apply deterministic algorithms, which aim to minimize the influence of the viewer‘s subjectivity and unstable judgment.

2.1.1. Segmentation. Spine image segmentation is a fundamental part of precision medicine. Early AI research focused on identifying the specific coordinates of spinal anatomical structures. Based on this, subsequent studies have successfully segmented various anatomical structures, including vertebrae, spinal canals, intervertebral discs, and spinal cords. This enhanced segmentation is essential for in-depth analysis and clinical applications of spine images. Numerous studies have shown that trained AI algorithms can demonstrate high accuracy in spine image segmentation tasks, providing reliable technical support for clinical applications.[9–11]

2.1.2. Measurement. The measurement of spinal imaging parameters provides an essential basis for physicians to assess disease progression and select appropriate treatment options. Traditional manual measurement methods are inefficient and have significant inter- and intraobserver variability. The application of AI algorithms to achieve automated and precise measurement of spinal imaging is crucial for evaluating biomechanical parameters such as overall spinal curvature, intervertebral disc height, and relative vertebral positions. Analyzing the signal characteristics of vertebral bodies, intervertebral discs, and surrounding spinal ligaments using computed tomography (CT) and magnetic resonance imaging (MRI) has significant reference value in characterizing the severity of spinal diseases, providing a basis for determining the pathophysiological state of spinal disorders. With continuous technological advancements, the measurement of spinal morphological parameters is no longer limited to traditional medical imaging techniques. Meng et al.[12] used light-based depth sensing technology and DL algorithms to synthesize ordinary RGB-depth images of the backs of pediatric scoliosis patients into images comparable to X-rays. Through the analysis of these synthesized images, the model demonstrated high accuracy in predicting anatomical landmarks of the back and could also provide severity grading and curve types of adolescent idiopathic scoliosis based on the synthesized images. Currently, the model has been encapsulated into an open-platform application, and due to their convenience and wide availability, RGB images captured by smartphones have become a new source for obtaining spinal morphological information.[13] This innovative data collection method offers possibilities for patient self-monitoring and remote disease management by physicians, promising to drive transformative changes in the management model of spinal disorders.

2.1.3. Lesions detection. The evaluation of complex anatomical structures or subtle lesions often requires a great deal of time and effort. Therefore, AI-assisted diagnosis has become an urgent clinical need, especially for conditions that require prompt and accurate diagnosis, such as high cervical spine fractures. This urgency is particularly evident in complex cases with a high risk of misdiagnosis, including spinal tumors, tuberculosis, and chronic fractures. Gitto et al.[14] conducted a retrospective study involving diffusion-weighted and T2-weighted MRI images from 101 patients with histologically confirmed spinal bone tumors (22 benign, 38 primary malignant, and 41 metastatic). The study extracted 1702 radiomics features, and through stability testing and statistical screening, ultimately identified 10 key features. A tumor type classification model was constructed using a support vector machine, achieving a diagnostic accuracy of 76%, providing a quantitative basis for differentiating between benign and malignant tumors. However, the limited number of patients who underwent histological and pathological confirmation may introduce bias in both the training of the AI model and its subsequent clinical application. AIDOC, an AI platform approved by the US Food and Drug Administration for clinical use, initially reported a sensitivity of 91.7% and a specificity of 88.6% in detecting spinal fractures. However, in an external database test involving 1904 emergency cervical spine CT scans, it ultimately achieved a sensitivity of 54.9% and a specificity of 94.1%, revealing a decline in performance in detecting chronic fractures.[15] This case exemplifies the limited adaptability of AI, illustrating the challenges AI algorithms encounter in generalizing to clinical settings where disease prevalence and imaging protocols differ from those of the training dataset. Current research on AI in the diagnosis of spinal surgical diseases mainly focuses on binary classification tasks for individual disease types, which require AI to determine the presence or absence of a condition. However, there is a relative lack of research on multiclassification diagnostic tasks aimed at identifying specific clinical subtypes of positive lesions. Additionally, differential diagnostic tasks involving the multiclassification of diseases at the same anatomical site remain underexplored. Factors such as the difficulties in constructing standardized databases, the limited scale of existing databases, the complexity of annotation work, the large volume of tasks, and the high difficulty of recognizing multiclassification targets collectively contribute to the suboptimal performance of AI in complex multiclassification diagnostic tasks, hindering the progress of related research. In the future, more attention should be paid to complex multiclassification diagnostic tasks to effectively enhance the clinical utility of AI-assisted diagnosis.

2.1.4. Clinical data analysis. The diagnosis of diseases is not limited to imaging examinations, as clinical data also has significant diagnostic value. In addition to the patient's basic information, identical symptoms, neurological signs, diagnostic tests, patient-reported outcome measures, and laboratory results are all crucial characteristics. The 5 times sit-to-stand test is an objective assessment method commonly used in clinical settings to evaluate functional impairment. Staartjes et al. developed an AI model that incorporates a clustering algorithm. By inputting baseline data such as gender, age, height, weight, and smoking status, this model can determine the upper limit of normal values for different populations in the 5 times sit-to-stand test, enabling accurate assessment of patients' spinal dysfunction.[16] Integrating AI applications with textual records, laboratory data, and radiological imaging may enhance diagnostic and predictive performance, representing an interesting future direction for ML-assisted quality improvement. The extraction of these features and their integration with image data to form a multisource information fusion algorithm are current challenges that require further research.

2.2. Decision support

The decision-making process for spinal disease treatment has always been a complex challenge in clinical practice. Key considerations affecting surgical decisions include factors such as the patient's clinical symptoms, imaging findings, overall physical condition, comorbidities, and surgical risks. Constructing predictive models to assess patients' surgical suitability can provide clinicians with better decision-making support. In recent years, numerous studies have focused on developing AI models to predict patients' surgical suitability and assist in clinical decision-making. Among them, some models specifically predict the likelihood of surgical treatment for patients with degenerative disc disease. These models incorporate multiple clinical parameters, including patients' baseline data, clinical symptoms, sagittal and coronal balance parameters, Oswestry Disability Index scores, and Scoliosis Research Society-22 quality of life scores.[17,18] Additionally, laboratory test data are integrated into the training systems of AI models. Specifically, preoperative routine blood tests, blood biochemical indicators, coagulation function, and postoperative drainage volume can all be used to predict patients’ surgical suitability and postoperative infection risks. AI models not only process these laboratory test data effectively but also use them to optimize surgical decision-making.

The development of artificial general intelligence (AGI) has brought new opportunities and transformations to various fields.[19] In the medical field, although AGI has powerful information processing capabilities, its application faces many challenges due to the specialized and complex nature of medicine. Especially in clinical issues related to spinal orthopedic surgery, the performance and reliability of AGI models still need to be thoroughly evaluated. In this context, some studies have assessed the potential applications of AGI models, such as ChatGPT, by querying them on topics including low back pain, degenerative spondylolisthesis, cervical radiculopathy, and spinal cord injury.[20,21] The research found that for most clinical questions that are well defined by clinical guidelines, ChatGPT's responses were consistent with the content of those guidelines.[22] However, AGI models have also revealed several problems in medical applications. Some responses are incomplete or difficult to understand, which affects the effective communication of information. When faced with questions for which clinical guidelines do not provide best practices, the answers generated by ChatGPT are often too general, inconsistent with existing literature, and may even include fabricated data or citations. These issues indicate that improvements and optimizations are still necessary before AGI models like ChatGPT can be reliably applied in clinical settings.[23] To enable AGI models to better serve the medical field, improvements are required in several aspects. First, the completeness and readability of the answers should be improved to ensure that patients and doctors can accurately understand the information. Second, the consistency of answers must be enhanced to be in line with medical literature and clinical practice. Additionally, the safety of the model should be improved to prevent errors such as data fabrication.[24]

2.3. Development of intelligent spinal surgery equipment

In spinal surgery, intelligent navigation mainly relies on surgical robot technology.[25] Traditional surgical robots perform 3-dimensional (3D) reconstructions of the patient's imaging data preoperatively to establish a biological coordinate system of the human body. This process enables precise intraoperative localization of lesions, real-time monitoring of surgical approaches, and increased operational convenience.

Intraoperative navigation technology in spinal surgery significantly improves the precision of surgical operations by integrating the real-time positional feedback of surgical instruments with the imaging of the patient's anatomical structures. Burström et al.[26] developed an automatic segmentation system for spinal surgery based on intraoperative cone-beam CT. This system uses ML methods to achieve automatic segmentation of the spine and pedicle recognition, providing screw trajectory suggestions with a clinically relevant accuracy of 86.1%. The 7D Surgical System[27,28] is a machine vision-based navigation system for spinal surgery. It uses optical topographic imaging technology, combined with preoperative CT scan data, to rapidly complete intraoperative navigation registration. In addition, it creates multiplanar reconstructions of the spine through machine vision structured light imaging, which assists in pedicle screw placement while reducing the reliance on traditional intraoperative fluoroscopy, thereby decreasing radiation exposure. Jecklin et al. developed the X23D system, which uses DL technology to convert multiview 2-dimensional (2D) fluoroscopic images into 3D images of the spine in real-time during surgery. The system achieved an average F1 score of 88% and a surface score of 71% on unseen data, with the latter being 22% higher than that of existing technologies.[29] The X23D system not only provides a crucial reference for the conversion of 2D to 3D images and surgical planning in intraoperative navigation but also opens up new possibilities for navigation and decision-making in spinal surgery, especially in cases of intraoperative condition changes or emergency surgeries. Edström et al.[30,31] investigated an augmented reality (AR) surgical navigation system. This system can rapidly acquire high-quality 3D imaging data during surgery through imaging devices such as cone-beam CT or intraoperative MRI, and perform the real-time 3D reconstruction. By using AR technology, the system integrates the reconstructed 3D images with the actual surgical scene, providing surgeons with intuitive navigation information. The accuracy of the 3D reconstruction reaches the submillimeter level, enabling clear visualization of the anatomical structures of the spine. This technology not only improves the precision and safety of surgeries but also significantly reduces surgical time and complexity, providing surgeons with stronger support for surgical navigation.

AR technology is a technique that seamlessly integrates virtual information with the real world. This technology generates virtual objects, scenes, and other information through computers, precisely superimposing them onto the real environment, enabling real-time interaction between the virtual and the real. In spinal surgery, AR technology can display essential patient imaging data in real time and accurately fuse and present virtual information, such as the patient's 3D spinal model obtained preoperatively and intraoperatively, with the actual spinal site during surgery. This ability helps spinal surgeons improve surgical efficiency and safety. Furthermore, AR technology has been widely applied in surgical navigation assistance within spinal surgery. Through AR devices, such as head-mounted displays, surgeons can precisely integrate virtual surgical planning information with the patient's actual spinal region. For example, during pedicle screw implantation surgery, after wearing AR devices, surgeons can clearly visualize the virtual screw insertion path within their surgical field of view, similar to marking an accurate route on the real spine with special markers, providing intuitive and precise guidance for the surgical procedure. Auloge et al.[32] developed a navigation system that integrates AR and AI technologies for the treatment of vertebral compression fractures. The AR/AI navigation system uses AI software, AR technology, and 3D imaging guidance to autonomously identify anatomical landmarks, generate safe and precise needle trajectories, and superimpose virtual 3D anatomical data onto real-world 2D visual images in real time. This significantly improves instrument deployment accuracy. Studies indicate that the accuracy of percutaneous vertebroplasty guided by this technology is comparable to that of the standard fluoroscopy-guided group, yet it offers a markedly reduced radiation dose and shorter operative time. Moreover, AR technology provides real-time feedback during procedures, assisting surgeons in making timely adjustments to their maneuvers. For example, during screw implantation, the AR device displays the deviation between the actual position of the screw and the planned position in real time. If the actual implantation angle or depth deviates from the preset values, the AR system promptly issues an alert and visually indicates the necessary adjustment direction and magnitude through virtual images. Using this real-time feedback, the surgeon can quickly make adjustments to ensure the screw is accurately implanted in the ideal position, further enhancing the precision and safety of the surgery. AR technology offers a revolutionary approach to the education and training of spinal surgery. Liu et al. conducted a comparative study involving 12 medical students who used AR teaching methods and 13 who used traditional teaching techniques.[33] The findings showed that the AR-assisted teaching system significantly outperformed traditional methods in terms of teaching quality, student motivation, and mastery of anatomical structures. This research provides substantial experimental data supporting the implementation of AR technology in the clinical teaching of anatomy and surgery related to spinal tumors, highlighting its effectiveness in medical education. Hasan et al.[34] further pointed out that virtual reality and AR technologies are transforming the training paradigms for surgical trainees by providing immersive and interactive experiences. In the context of spinal surgery, these technologies allow trainees to practice surgical procedures in a simulated environment and assess their performance metrics through sophisticated analysis and algorithms, thereby enhancing their surgical competencies.

In addition to being used for assisting screw implantation, Li et al.[35] have developed the world's first spinal laminotomy surgical robot capable of autonomous lamina cutting operations. The team used the robot to autonomously complete 80 laminotomy planes. This robotic system can rapidly and accurately perform laminotomy, reducing the time required for manual adjustments and repeated confirmations during surgery. Additionally, its multi-mode safety system enables the robot to automatically identify biological tissue characteristics, avoiding damage to surrounding nerves and blood vessels. This not only improves surgical efficiency but also ensures the accuracy and safety of laminotomy procedures.

With the development of surgical robots, contemporary multiseries surgical robots at home and abroad have integrated many intelligent system functions during their research and development. These functions include automated instrument-to-anatomy registration, intelligent spatial coordinate tracking and positioning, rendering and enhancement of regions of interest in the surgical area, automatic grasping and recognition of anatomical structures, and intraoperative musculoskeletal feedback. These advancements have significantly increased the level of intelligence in intraoperative navigation. However, the author argues that the current stage of spinal surgical robots has not yet achieved true AI; instead, they function more as instant positioning systems with some intelligent features. It is expected that with the further integration of AI and navigation technology, more AI-oriented surgical robots will emerge in the future.

2.4. Postoperative management

With the escalating number of spinal surgeries, postoperative management and rehabilitation have emerged as pivotal determinants of patient outcomes and quality of life. In this scenario, AI technology, capitalizing on its distinct advantages in data analysis, pattern recognition, and real-time feedback, is increasingly integrated into various facets of postoperative management in spinal surgery. This integration has catalyzed a paradigm shift from traditional experience-based approaches to data-driven models. The application of AI in postoperative management is predominantly manifested in several domains: prediction of postoperative complications and risks, management of postoperative recovery and quality of life assessment, intelligent early-warning systems in remote monitoring, data collection and functional evaluation via wearable devices, personalized rehabilitation training guidance, and intelligent prediction of long-term outcomes. These applications not only substantially enhance the precision and efficiency of postoperative rehabilitation but also facilitate the establishment of a more intimate and real-time treatment feedback loop between healthcare providers and patients. Consequently, AI-enabled postoperative management is evolving into an indispensable component of the intelligent development of the entire spinal surgery process. The prediction of postoperative complications represents a significant application of AI in spinal surgery, encompassing factors such as infection rates, extended hospital stays, and the requirement for reoperation. A retrospective study utilizing multicenter data from China developed a predictive model for deep surgical site infection (SSI) following posterior spinal fixation surgery.[36] This study juxtaposed knowledge-driven models, which were grounded in clinical expertise, with data-driven models. Through comprehensive analysis, 12 key predictive factors were identified: age, body mass index, diabetes, steroid use, hypoalbuminemia, surgery duration, blood loss, and the number of instrumented segments. Based on these factors, the A-DOUBLE-SSI risk scoring system was established. This system enables stratified patient assessment; for example, patients with a score ≤ 8 have an infection rate of approximately 1.06%, whereas those with a score > 15 exhibit a significantly elevated infection rate of up to 40.6%. This tool facilitates the preoperative identification of high-risk patients and the implementation of enhanced preventive measures. The length of postoperative hospital stay is a critical indicator of patient recovery and the occurrence of complications. AI models have demonstrated effectiveness in predicting patients who may experience prolonged hospital stays. A study focused on lumbar spinal stenosis surgery developed a web-based ML tool for predicting extended postoperative hospital stays.[37] By collecting data from 540 patients and employing a random forest model in conjunction with SHAP and LIME algorithms for model result interpretation, the study achieved a high prediction accuracy, with an area under the curve (AUC) of 0.83 indicating a good discriminative ability to correctly identify patients at risk of prolonged postoperative hospital stay. Model interpretation revealed that intraoperative blood loss was the most influential factor contributing to prolonged hospital stays. Additionally, the model was deployed as a web application for clinical use, enabling the early identification of high-risk patients with extended hospitalization needs.

AI technology has been extensively employed to predict adverse outcomes following spinal surgery, including reoperation and readmission. Rodrigues et al.[38] conducted a comparative analysis of DL and traditional ML algorithms in predicting these adverse outcomes after anterior cervical discectomy and fusion. Their findings indicated that although DL models showed marginal performance improvements, their overall effectiveness was comparable to traditional methods. This study, which analyzed data from nearly 1800 patients, aimed to predict the risks of complications within 90 days, readmission within 90 days, and reoperation within 2 years. The results underscored the limited information content of existing clinical variables, highlighting the need to introduce new predictive factors to significantly enhance the accuracy of long-term risk prediction. Merali et al.[39], leveraging ML techniques, utilized the AOSpine multicenter prospective database to forecast postoperative outcomes in patients with degenerative cervical myelopathy (DCM). They developed a classification model based on random forests, which accurately predicted postoperative functional recovery (modified Japanese Orthopaedic Association score) and quality of life improvement (Short-Form 6D) by thoroughly analyzing patients' preoperative clinical characteristics, symptom severity, and quality of life indicators. The model achieved an accuracy rate of 77% on an independent test set, with a peak AUC value of 0.73.

Postoperative functional recovery and quality-of-life improvement are critical metrics for evaluating surgical success. The application of AI technology in pain management and functional recovery prediction aims to facilitate the formulation of personalized rehabilitation programs. Gabriel et al.[40] reported that ML models have been utilized to predict the risk of prolonged opioid use following various surgical procedures, including spinal surgery. Other researchers have developed models to identify patients at risk of developing persistent postoperative opioid dependence or opioid use disorder. For example, Karhade et al.[41] analyzed data from approximately 5413 spinal surgery patients and, using a variety of algorithms, accurately predicted that 7.7% of patients continued opioid use 90 to 180 days postoperatively. The Elastic Net

Logistic Regression model, which demonstrated the highest effectiveness, achieved a C-statistic of 0.81, indicating excellent discriminative power. This model also identified several significant risk factors, such as the use of implants during surgery, preoperative opioid use duration, and comorbid depression, which emerged as the strongest predictors of persistent postoperative medication use. These AI-based tools hold promise for identifying high-risk populations for pain management before patient discharge, thereby enabling enhanced follow-up interventions, such as the development of safer analgesic protocols and close medication monitoring, to reduce the incidence of chronic postoperative pain and opioid addiction.

Yagi et al.[42] developed and validated a model for predicting health-related quality of life (HRQoL) at a 2-year follow-up for patients undergoing decompression surgery for lumbar spinal stenosis. Trained on multicenter data from 3 different medical institutions, this model incorporated patients' sociodemographic, clinical, and surgical factors and accurately predicted postoperative changes in HRQoL. Such predictive capabilities empower physicians to conduct more accurate preoperative prognosis assessments, especially in identifying patients who may not achieve the minimal clinically important difference in postoperative improvement. This, in turn, enables more comprehensive patient education and the implementation of adjunctive therapeutic measures before surgery. Ames et al.[43] employed 6 ML algorithms to develop a prediction model based on 150 variables for 561 patients who underwent corrective surgery for adult spinal deformity. The model systematically predicted response trends for each item in the Scoliosis Research Society-22 Revised questionnaire at 1-year and 2-year postoperative intervals. The results showed that the model excelled in predicting issues closely related to pain, mobility, and social function, achieving a maximum AUC of 86.9%. This study introduced AI methods at a detailed level of patient outcomes, paving the way for data-driven individualized surgical consultations. Khan et al.[44] systematically evaluated the performance of 8 ML algorithms in predicting postoperative HRQoL improvement, using data from 193 patients with mild DCM from the AOSpine prospective database. The results indicated that the Gradient Boosting Machine model outperformed others in predicting mental health improvement, as measured by the Mental Component Summary, while the Earth model demonstrated optimal performance in predicting physical health improvement, as indicated by the Physical Component Summary. The test set AUC values for the Gradient Boosting Machine and Earth models were 0.77 and 0.78, respectively. This research provides data-driven support for making precise surgical decisions in the mild DCM population, further validating the practical application of ML in prognostic prediction within spinal surgery.

Beyond predicting long-term outcomes, AI is increasingly being utilized for real-time monitoring during the postoperative recovery phase. Various AI algorithms have been integrated into mobile applications, body tracking systems, and wearable sensors to capture objective indicators of functional recovery.[45] Through smart devices worn by patients after surgery, AI can continuously assess motor abilities, walking balance, daily activity levels, and other parameters. This enables the prompt identification of functional recovery delays and triggers timely alerts to the medical team for intervention. For example, AI analysis of gait and posture data offers an objective assessment of recovery progress, providing more precise insights compared to traditional subjective outpatient follow-up reports.[46,47] In summary, AI plays a dual role in postoperative recovery management: it not only predicts long-term functional outcomes and formulates individualized follow-up plans but also monitors key recovery indicators in real time, enabling early warnings and interventions that ultimately enhance patients' quality of life.

3. Summary and prespectives

Current research achievements in the application of AI in spinal surgery have been concentrated on medical image recognition and diagnostic support. Nevertheless, the substantial potential exists for further development in areas such as surgical treatment decision-making, prognosis prediction, intraoperative intelligent navigation assistance, and postoperative rehabilitation therapy. Existing AI research is confronted with several limitations. First, the research and design frameworks of current models lack unified standards, encompassing aspects such as the construction of clinical and imaging databases, the design of annotation schemes, and the validation of model performance. The Checklist for AI in Medical Imaging (CLAIM) published in *Radiology Artificial Intelligence* offers researchers a reference for standardized research processes, thereby promoting clear, transparent, and reproducible scientific communication in the field of medical imaging research.[48,49] Simultaneously, there is an urgent need to develop domestic guidelines for the implementation and application of AI research in spinal surgery. Second, while evaluating model efficacy is vital during AI algorithm development, the focus should shift from statistical metrics to real-world clinical impacts. We need to prioritize assessing how well these models solve specific clinical problems, ensuring their practical value in healthcare. Therefore, there is an urgent need for scientific and standardized interpretation of model output results. Before AI medical devices are approved for clinical use, researchers must disclose the models' expected performance, intended applications, limitations, and potential impacts.

In the future, concerted efforts will be made to promote the establishment of high-quality, multicenter databases and the formulation of unified data collection and annotation standards, which are fundamental for constructing AI imaging models. Through the integration of imaging, genomics, and clinical data for multimodal fusion, more comprehensive and precise AI models will be developed. In the diagnosis and treatment of spinal diseases, greater emphasis should be placed on integrating AI technology with existing medical diagnosis and treatment processes. Leveraging language models to achieve accurate diagnoses and formulate personalized treatment plans for spinal diseases will be a key direction. At the same time, establishing ethical guidelines and regulatory frameworks for AI medical technology to safeguard patient privacy and data security remains an essential and critical task.

Acknowledgments

Not applicable.

Ethical statement

Not applicable.

Conflicts of interest

The authors have no conflicts of interest to disclose.

Funding source

This work is sponsored by the Science Research Foundation (BHJ23C012).

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Author contributions

HC participated in research design and critical revision of the article. CG, TY, TZ, and WZ participated in the writing of the article. JW and FZ participated in the performance of the research.

References

[1] Yu K, Healey E, Leong T, Kohane IS, Manrai AK. Medical artificial intelligence and human values. N Engl J Med. 2024;390:1895–904.

[2] Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349:255–60.

[3] Toma A, Diller G, Lawler PR. Deep learning in medicine. JACC Adv. 2022;1:100017.

[4] Ren G, Yu K, Xie Z, et al. Current applications of machine learning in spine: from clinical view. Global Spine J. 2022;12:1827–40.

[5] Lopez CD, Boddapati V, Lombardi JM, et al. Artificial learning and machine learning applications in spine surgery: a systematic review. Global Spine J. 2022;12:1561–72.

[6] Hornung AL, Hornung CM, Mallow GM, et al. Artificial intelligence in spine care: current applications and future utility. Eur Spine J. 2022;31:2057–81.

[7] Hornung AL, Hornung CM, Mallow GM, et al. Artificial intelligence and spine imaging: limitations, regulatory issues and future direction. Eur Spine J. 2022;31:2007–21.

[8] Galbusera F, Casaroli G, Bassani T. Artificial intelligence and machine learning in spine research. JOR Spine. 2019;2:e1044.

[9] Da Mutten R, Zanier O, Theiler S, et al. Whole spine segmentation using object detection and semantic segmentation. Neurospine. 2024;21:57–67.

[10] Zou L, Guo L, Zhang R, et al. VLTENet: a deep-learning-based vertebra localization and tilt estimation network for automatic Cobb angle estimation. IEEE J Biomed Health Inf. 2023;27:3002–13.

[11] Mohanty R, Allabun S, Solanki SS, et al. NAMSTCD: a novel augmented model for spinal cord segmentation and tumor classification using deep nets. Diagnostics (Basel, Switzerland). 2023;13:1417.

[12] Meng N, Cheung JPY, Wong KK, et al. An artificial intelligence powered platform for auto-analyses of spine alignment irrespective of image quality with prospective validation. eClinicalMedicine. 2022;43:101252.

[13] Zhang T, Zhu C, Zhao Y, et al. Deep learning model to classify and monitor idiopathic scoliosis in adolescents using a single smartphone photograph. JAMA Netw Open. 2023;6:e2330617–e2330617.

[14] Gitto S, Bologna M, Corino VDA, et al. Diffusion-weighted MRI radiomics of spine bone tumors: feature stability and machine learning-based classification performance. Radiol Med. 2022;127:518–25.

[15] Voter AF, Larson ME, Garrett JW, Yu JPJ. Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures. Am J Neuroradiol. 2021;42:1550–6.

[16] Staartjes VE, Klukowska AM, Vieli M, et al. Machine learning–augmented objective functional testing in the degenerative spine: quantifying impairment using patient-specific five-repetition sit-to-stand assessment. Neurosurg Focus. 2021;51:E8.

[17] Broida SE, Schrum ML, Yoon E, et al. Improving surgical triage in spine clinic: predicting likelihood of surgery using machine learning. World Neurosurg. 2022;163:e192–8.

[18] Baroncini A, Campagner A, Cabitza F, et al. The use of machine learning for the prediction of response to follow-up in spine registries. Int J Med Inform. 2025;195:105752.

[19] Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29:1930–40.

[20] Karhade AV, Bongers MER, Groot OQ, et al. Development of machine learning and natural language processing algorithms for preoperative prediction and automated identification of intraoperative vascular injury in anterior lumbar spine surgery. Spine J. 2021;21:1635–42.

[21] Karhade AV, Lavoie-Gagne O, Agaronnik N, et al. Natural language processing for prediction of readmission in posterior lumbar fusion patients: which free-text notes have the most utility? Spine J. 2022;22:272–7.

[22] Alanazi AH, Cradock A, Rainford L. Development of lumbar spine MRI referrals vetting models using machine learning and deep learning algorithms: comparison models vs healthcare professionals. Radiography (Lond). 2022;28:674–83.

[23] Pedro KM, Alvi MA, Hejrati N, Quddusi AI, Singh A, Fehlings MG. Machine learning-based cluster analysis identifies four unique phenotypes of patients with degenerative cervical myelopathy with distinct clinical profiles and long-term functional and neurological outcomes. Ebiomedicine. 2024;106:105226.

[24] Sarikonda A, Isch E, Self M, et al. Evaluating the adherence of large language models to surgical guidelines: a comparative analysis of Chatbot recommendations and North American Spine Society (NASS) coverage criteria. Cureus. 2024;16:e68521.

[25] Zhou S, Zhou F, Sun Y, et al. The application of artificial intelligence in spine surgery. Front Surg. 2022;9:885599.

[26] Burström G, Buerger C, Hoppenbrouwers J, et al. Machine learning for automated 3-dimensional segmentation of the spine and suggested placement of pedicle screws based on intraoperative cone-beam computer tomography. J Neurosurg Spine. 2019;31:147–54.

[27] Kalfas IH. Machine vision navigation in spine surgery. Front Surg. 2021;8:640554.

[28] Jakubovic R, Guha D, Gupta S, et al. High speed, high density intraoperative 3D optical topographical imaging with efficient registration to MRI and CT for craniospinal surgical navigation. Sci Rep. 2018;8:14894.

[29] Jecklin S, Jancik C, Farshad M, Fürnstahl P, Esfandiari H. X23D—intraoperative 3D lumbar spine shape reconstruction based on sparse multi-view X-ray data. J Imaging. 2022;8:271.

[30] Edström E, Burström G, Nachabe R, Gerdhem P, Elmi Terander A. A novel augmented-reality-based surgical navigation system for spine surgery in a hybrid operating room: design, workflow, and clinical applications. Oper Neurosurg. 2020;18:496–502.

[31] Edström E, Burström G, Omar A, et al. Augmented reality surgical navigation in spine surgery to minimize staff radiation exposure. Spine. 2020;45:E45–53.

[32] Auloge P, Cazzato RL, Ramamurthy N, et al. Augmented reality and artificial intelligence-based navigation during percutaneous vertebroplasty: a pilot randomised clinical trial. Eur Spine J. 2020;29:1580–9.

[33] Liu S, Yang J, Jin H, et al. Exploration of the application of augmented reality technology for teaching spinal tumor's anatomy and surgical techniques. Front Med. 2024;11:1403423.

[34] Hasan S, Miller A, Higginbotham D, Saleh ES, McCarty S. Virtual and augmented reality in spine surgery: an era of immersive healthcare. Cureus. 2023;15:e43964.

[35] Li Z, Wang C, Song X, et al. Accuracy evaluation of a novel spinal robotic system for autonomous laminectomy in thoracic and lumbar vertebrae: a cadaveric study. J Bone Joint Surg Am. 2023;105:943–50.

[36] Cheng L, Liu J, Lian L, et al. Predicting deep surgical site infection in patients receiving open posterior instrumented thoracolumbar surgery: A-DOUBLE-SSI risk score – a large retrospective multicenter cohort study in China. Int J Surg. 2023;109:2276–85.

[37] Yasheng P, Yusufu A, Yimiti Y, Luan H, Peng C, Song X. Web-based machine learning application for interpretable prediction of prolonged length of stay after lumbar spinal stenosis surgery: a retrospective cohort study with explainable AI. Front Physiol. 2025;16:1542240.

[38] Rodrigues AJ, Schonfeld E, Varshneya K, et al. Comparison of deep learning and classical machine learning algorithms to predict postoperative outcomes for anterior cervical discectomy and fusion procedures with state-of-the-art performance. Spine. 2022;47:1637–44.

[39] Merali ZG, Witiw CD, Badhiwala JH, Wilson JR, Fehlings MG. Using a machine learning approach to predict outcome after surgery for degenerative cervical myelopathy. PLoS One. 2019;14:e0215133.

[40] Gabriel RA, Park BH, Hsu C, Macias AA. A review of leveraging artificial intelligence to predict persistent postoperative opioid use and opioid use disorder and its ethical considerations. Curr Pain Headache Rep. 2025;29:30.

[41] Karhade AV, Ogink PT, Thio QCBS, et al. Development of machine learning algorithms for prediction of prolonged opioid prescription after surgery for lumbar disc herniation. Spine J. 2019;19:1764–71.

[42] Yagi M, Michikawa T, Yamamoto T, et al; Keio Spine Research Group. Development and validation of machine learning-based predictive model for clinical outcome of decompression surgery for lumbar spinal canal stenosis. Spine J. 2022;22:1768–77.

[43] Ames CP, Smith JS, Pellisé F, et al; European Spine Study Group. Development of predictive models for all individual questions of SRS-22R after adult spinal deformity surgery: a step toward individualized medicine. Eur Spine J. 2019;28:1998–2011.

[44] Khan O, Badhiwala JH, Witiw CD, Wilson JR, Fehlings MG. Machine learning algorithms for prediction of health-related quality-of-life after surgery for mild degenerative cervical myelopathy. Spine J. 2021;21:1659–69.

[45] Chakravorty A, Mobbs RJ, Anderson DB, et al. The role of wearable devices and objective gait analysis for the assessment and monitoring of patients with lumbar spinal stenosis: systematic review. BMC Musculoskelet Disord. 2019;20:288.

[46] Lee SI, Campion A, Huang A, et al. Identifying predictors for postoperative clinical outcome in lumbar spinal stenosis patients using smart-shoe technology. J Neuroeng Rehabil. 2017;14:77.

[47] Sun J, Liu Y, Yan S, et al. Clinical gait evaluation of patients with lumbar spine stenosis. Orthop Surg. 2018;10:32–9.

[48] 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:e200029.

[49] Tejani AS, Klontzas ME, Gatti AA, et al; CLAIM 2024 Update Panel. Checklist for artificial intelligence in medical imaging (CLAIM): 2024 Update. Radiol Artif Intell. 2024;6:e240300.