OPEN Specialty Article: The University of Hong Kong Collection

The Application of Artificial Intelligence in Spine Surgery: A Scoping Review

Liangyu Shi, MS

Hongfei Wang, PhD

Graham Ka-Hon Shea, FRCSEd (O&T)

From the Department of Orthopaedics and Traumatology, Li Ka Shing University, The University of Hong Kong, Hong Kong SAR, China

Correspondence to Dr. Shea: gkshea@hku.hk

None of the following authors or any immediate family member has received anything of value from or has stock or stock options held in a commercial company or institution related directly or indirectly to the subject of this article: Dr. Shi, Dr. Wang, and Dr. Shea.

Authorship contribution: Dr. Shi: acquisition and analysis of data, and drafting of manuscript; Dr. Wang: acquisition and analysis of data, and drafting of manuscript; Dr. Shea: study conception, drafting, and review of manuscript.

Dr. Shi and Dr. Wang contributed equally to this work.

JAAOS Glob Res Rev 2025;9: e24.00405

DOI: 10.5435/JAAOSGlobal-D-24-00405

ABSTRACT

Background: A comprehensive review on the application of artificial intelligence (AI) within spine surgery as a specialty remains lacking.

Methods: This scoping review was conducted upon PubMed and EMBASE databases according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. Our analysis focused on publications from January 1, 2020, to March 31, 2024, with a specific focus on AI in the field of spine surgery. Review articles and articles predominantly concerning secondary validation of algorithms, medical physics, electronic devices, biomechanics, preclinical, and with a lack of clinical emphasis were excluded.

Results: One hundred five studies were included after our inclusion/exclusion criteria were applied. Most studies (n = 100) were conducted through supervised learning upon prelabeled data sets. Overall, 38 studies used conventional machine learning methods upon predefined features, whereas 67 used deep learning methods, predominantly for medical image analyses. Only 25.7% of studies (27/105) collected data from more than 1,000 patients for model development and validation. Data originated from only a single center in 72 studies. The most common application was prognostication (38/105), followed by diagnosis (35/105), medical image processing (29/105), and surgical assistance (3/105).

Conclusion: The application of AI within the domain of spine surgery has significant potential to advance patient-specific diagnosis, management, and surgical execution.

Introduction

Over the past decade, advances in artificial intelligence (AI) have evolved from being of academic interest to tools ready for use at the bedside. The capacity of AI to process a large amount of clinical data through sophisticated algorithms has surpassed the performance of seasoned physicians in selected contexts. Spine surgery is a medical discipline requiring not only astute clinical diagnosis but also meticulous surgical planning and execution. Given this breadth in the relevance of AI in spine surgery, it is invaluable that advances across the field—spanning both specific spine disorders and areas of application—may be summarized and critically appraised. This promises to provide insight into how research developments may be consolidated and refined toward benefitting patients. In this scoping review, we provide a comprehensive analysis of AI's role in spine surgery to date, with the objectives being to shed light on the current state of clinical integration and identify trajectories for future research and development.

Methods

Study Protocol

This scoping review was conducted in strict accordance with the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews.[1] The objective was to thoroughly chart the scope of literature concerning AI and spine surgery, with a particular focus on the integration of AI within clinical practice.

Inclusion and Exclusion Criteria

We employed a detailed search strategy, outlined in Supplementary Figure 1 (http://links.lww.com/JG9/A403), to identify articles listed upon PubMed and EMBASE databases and published between January 1, 2020, and March 31, 2024. Our search terms were selected to encompass AI-related technologies (including AI, machine learning [ML], deep learning [DL], and computer vision) alongside clinical and surgical aspects of spine surgery (medical image computing [MIC], assisted diagnosis, prognostication, surgical planning, and assistance). To ensure the incorporation of rigorously refereed articles of high quality, an additional inclusion criterion was the selection of articles published in journals ranked in the top 25% by Journal Citation Report (2023) impact factor rankings (JCR Q1) in the categories of "Clinical Medicine" and "Computer Science." We excluded abstracts, conference proceedings, reviews, commentaries, and editorial pieces. Studies without a clinically orientated focus were excluded. Article screening and selection was carried out independently by two experienced postgraduate researchers with a background in AI and clinical orthopaedics. Disagreements were resolved through reaching a consensus, ensuring that each selected article met our defined selection criterion.

Results

Search Results

As shown in Figure 1, our initial search yielded 2,959 results, of which 890 studies were published in JCR Q1 journals. An additional 785 publications were excluded following the evaluation of their title, abstract, and full text because they did not exhibit a predominant emphasis on the intersection between AI and spine surgery. A comprehensive analysis of these articles was conducted to concerning the application of AI in the field of spine surgery. Country of origin for the corresponding author among each included publication was analyzed. As shown in Figure 2, more than half of studies originated from China and the United States, corresponding to 43 (40.9%) and 25 (23.8%) articles, respectively. Overall, eight studies (7.6%) were from Japan and seven studies (6.7%) from Korea. Canada, Singapore, and Switzerland contributed three studies (2.9%) each, whereas the remaining 13 articles were published in other countries.

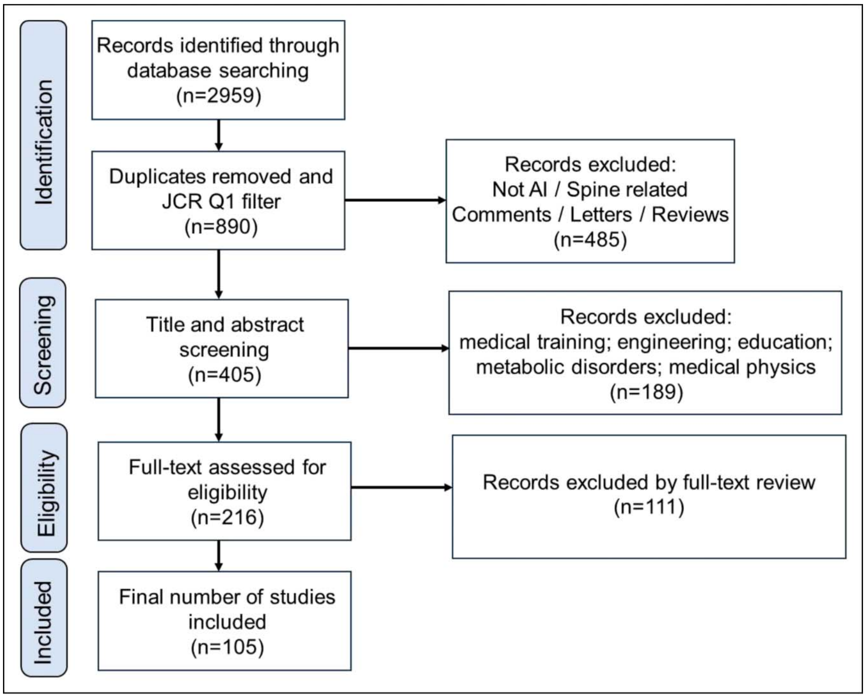
Clinical Data

Type of Data

Provision of clinical data is essential for DL. According to the type of input data, literature may be divided into being clinical parameters–based studies (n = 32), single-modality medical image–based studies (MRI [n = 25], CT [n = 13], radiograph [n = 22] and ultrasonography [n = 2]), sensor parameter–based studies (n = 4), and composite input-based studies (n = 7). Results are summarized in Figure 3.

As an example of clinical parameters, age, sex, disease duration, and preoperative Japanese Orthopedic Association score were integrated with ML methods to predict surgical outcomes in patients with degenerative cervical myelopathy (DCM).[2] Literature concerning MRI as input included two studies using quantified MRI features[3,4] and 23 studies using MRI images. The former approach relied on feature extraction software to quantify MRI features for ML input, such as for the classification of spinal tumors.[3] The latter directly approach employed DL to extract image features without prior feature engineering, such as for the automated detection of lumbar disk herniation (LDH).[5]

Figure 1



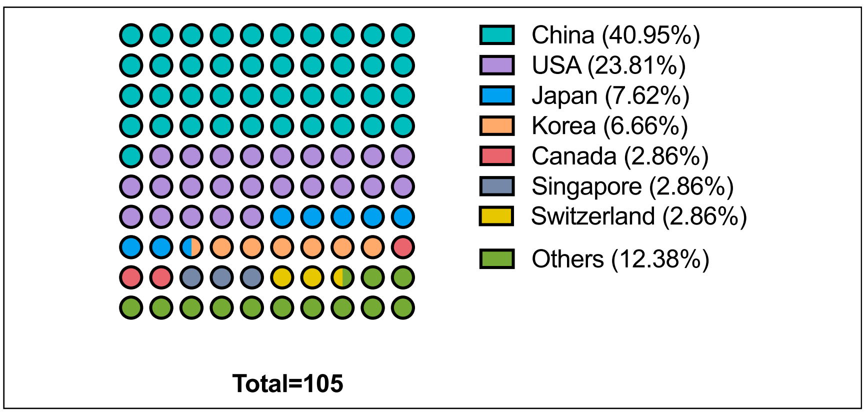
PRISMA flow diagram showing the scoping review process. PRISM = Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Overall, 21 studies used radiographic images directly as input, whereas one study used quantified radiograph morphological features.[6] Twelve studies considered CT images as input, and one study considered CT texture features.[7] Two studies on AI-based ultrasonography image segmentation considered scoliosis detection[8] and quantification of scoliosis-related measurements.[9]

Medical images may be combined with clinical parameters for AI modeling, with examples being radiographs,[10,11] MRI scans,[12,13] and CT scans in combination with clinical parameters.[7] Seven studies involved more than one data modality as input. Inputs of paired images were also reported from two studies for medical image synthesis, including the generation of lumbar spine CT scans from MRIs[14] and scoliosis radiographs from back photographs.[15]

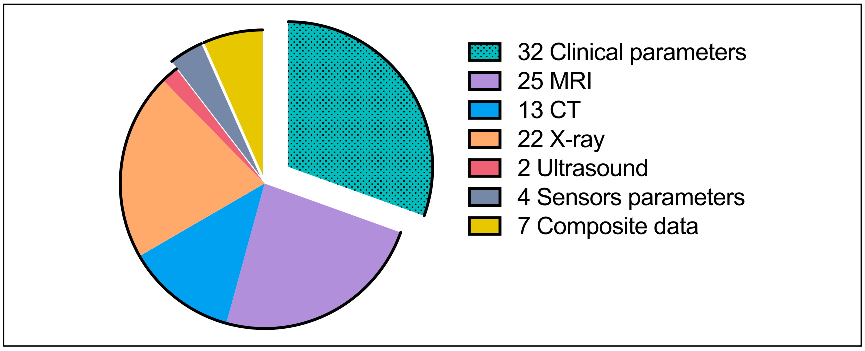
Four studies were conducted upon sensor-based parameters, with examples being radio frequency sensing for early diagnosis of spinal disorders,[16] gait acquisition with wearable inertial sensors for spinal cord injury

Figure 2



Visual representation showing country of origin of study corresponding author.

Figure 3



Pie chart showingsummary on the type of data used.

evaluation and prognostication,[17] hand movement video capture,[18] and smartphone-based grip-and-release parameters[19] for cervical myelopathy screening and severity evaluation.

Data Size

We investigated the cumulative size of assembled data sets and did not distinguish between training, validation, and testing sets because of variability in data division ratios. MRI and CT scans obtained at the same time point were considered as a single data belonging to one subject, rather than by dividing according to each two-dimensional slice. Data size among the literature was highly variable. As shown in Figure 4, 14 studies involved less than 100 subjects, 64 studies consisted of subject counts between 101 and 1,000, whereas 22 studies contained more than 1,000 but less than 10,000. Five studies reported more than 10,000 subjects, and among these, four studies used clinical parameters as model inputs and one study used radiograph images.

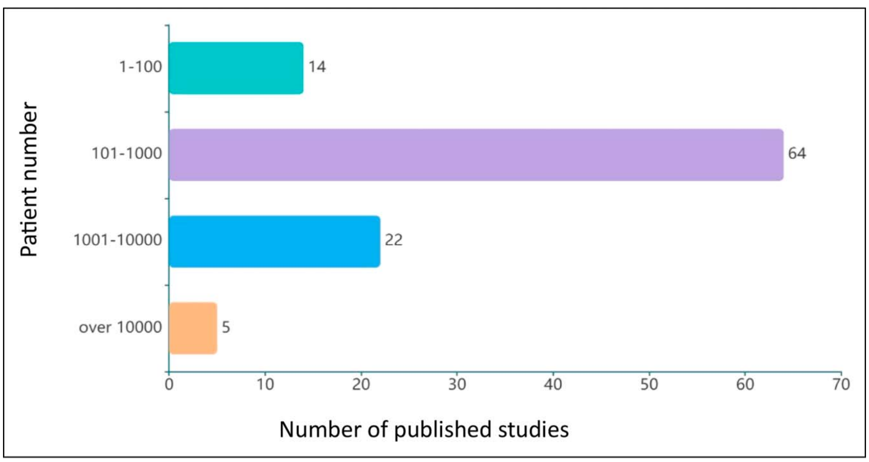
Independent testing with a separate data set is essential for AI model performance evaluation, and multicenter validation may be considered an ideal benchmark for real-world performance and generalizability. Overall, 72 studies collected data from a single center, whereas 23 studies used multicenter data for model development and validation. Ten studies used data from public databases.

Application of Artificial Intelligence in Spine Surgery

Medical Image Computing

MIC is an interdisciplinary field concerning computer science, information technology, and medical imaging. MIC uses computer-based techniques to analyze and interpret medical images, such as radiographs, MRIs,

Figure 4



Horizontal bar chart showingpatient number recruited for study data set.

CT scans, and ultrasounds. We outline several common modes of application below -

- SegmentationThe objective of image segmentation is to divide an image into segments that correspond to distinct anatomical structures or regions of interest, often in preparation for further focused analysis. Conventional manual segmentation techniques, in which radiologists or clinicians outline regions of interest, are labor intensive and subjective. By contrast, AI-powered segmentation offers an automated and standardized approach. Twenty-nine out of the 105 publications in this review specifically concerned MIC.
- Segmentation based on radiographsThe segmentation of radiographic images in spine surgery most commonly concerned the detection of vertebrae.[20] Precise identification and division of vertebrae can enhance the efficiency of diagnosing spinal conditions such as fractures and degenerative disk disease. This capability may be especially well-suited in the emergency ward and primary care environments where prompt assessments in the presence of nonspine specialists is crucial, for example, in the recognition of osteoporotic vertebral collapse in the elderly.[21]
- Segmentation based on CT images.CTs similarly enable for automated segmentation and labeling of vertebrae along the entire spine.[22,23] An example of application following segmentation is planning for stereotactic radiation therapy. More than 90% of automated treatment plans for bone metastases to the spine following model training was demonstrated to be viable upon evaluation by a radiation oncologist.[24] Another study achieved automated localization and segmentation of lumbar cancellous bone based on CT scans,[25] which may be the basis for the diagnosing osteoporosis or osteopetrosis based on Hounsfield units, and for the assessment of pedicle screw pullout strength.
- Segmentation based on MRI scansMRI scans are especially useful for segmentation of soft tissue such as intervertebral disks,[26-29] dural sac, nerve roots,[30] paravertebral musculature,[31] and bone marrow fat.[32] Pfirrmann grading for lumbar disk degeneration grade is determined by intervertebral height and signal intensity within the nucleus pulposus, whereas Modic changes dependent on vertebral bone marrow signal intensity.[33] AI-assisted disk and vertebra segmentation is a prerequisite to obtaining such measurements.[28]

Assisted measurement

Measurement of distinct parameters from spinal imaging is crucial for diagnosing various spinal conditions. AI-based algorithms can negate errors and expedient data collection in comparison to manual measurement.[34] Literature concerning assisted spinal parameter measurements may be categorized into four main areas.

- Vertebrae landmark detection and localizationAlgorithms have enabled for the localization and identification of cervical, thoracic, and lumbar vertebrae from CT scans,[35] which provide the foundation for diagnoses and therapy. An example is in the diagnosis of cervical instability through a U-Net-based algorithm identifying vertebral landmarks on dynamic cervical radiographs and thereafter calculating the relative angles of motion between adjacent vertebrae.[36]
- Cobb angle and spinal curvature measurementsUpon manual measurements, 95% confidence intervals for Cobb angles vary from 2.8° to 8°.[37] Conversely, upon AI-assisted measurements, the mean absolute error (MAE) was more consistent at $<3°$.[38-40] AI also facilitates measurement of spinal curvatures upon off-center, angulated, and blemished images, as well as from photographs captured from smartphones or screenshots.[41] Ultrasonic images have demonstrated utility for scoliosis screening by extracting landmarks (i.e. spinous processes) to build a three-dimensional spinal profile while avoiding radiation.[9] Another clinically relevant application is in the quantitative measurement of thoracolumbar compression fractures to facilitate management.[42]
- Spinopelvic parameter measurementAn accurate assessment of spinopelvic parameters is especially important for surgical planning in deformity correction toward improving pain and functional outcomes. In addition to Cobb angles, AI may be harnessed upon standing whole-spine radiographs to automatically measure spinopelvic parameters, including coronal vertical axis, sagittal vertical axis, thoracic kyphosis, lumbar lordosis, T9 spinopelvic inclination, frontal pelvic asymmetry, sacral slope, pelvic tilt, and pelvic incidence. Performance in both coronal plane (ICCs $\geq$ 0.95; MAE $\leq$ 2.9° or 1.97 mm) and sagittal plane-based (ICCs $\geq$ 0.85; MAE $\leq$ 4.4° or 2.7 mm) parameters demonstrated excellent consistency,[43] as was also demonstrated by another similar study.[44]
- Grading of disk herniation, central canal stenosis, and nerve root compressionFeatures and

parameters upon spinal MRIs are invaluable toward diagnosis of disk degeneration and neural compression. As examples of AI-assisted diagnoses, intervertebral disk herniation, central canal stenosis, and nerve root compression may be diagnosed by analyzing the size of the herniated disk, space occupied by cerebrospinal fluid, and the spatial relationships between the disk and the nerve root. The resulting accuracy for neural compression ranges from 81% to 84%, with respective κ values comparing expected and observed accuracy ranging 0.92 to 0.98.[45,46]

Diagnosis of Spine-Related Disorders

Segmentation provides a foundational reference for detecting spinal anomalies, and thereby facilitates the diagnosis of common spinal disorders as detailed below. In addition, the introduction of automated medical reports following imaging findings has been proposed to enhance diagnostic efficiency and costs by reducing the necessity for healthcare professionals while also minimizing the chance for error.[47,48]

Vertebral fracture

AI can detect variations in vertebral height, shape, and bone density, thereby facilitating the detection of vertebral fractures.[49-52] A DL classifier for osteoporotic fractures employed spinal radiographs to identify and categorize fractures based on the amount of vertebral height loss and was proposed as a potential screening tool.[53] Furthermore, AI has the potential to classify vertebral fracture subtype by AO classification, achieving an accuracy approaching 80%.[54] The Thoracolumbar Injury Classification and Severity Score is an assessment tool commonly used to facilitate the management of traumatic thoracolumbar spine injuries. A DL model was developed to allow for automated categorization of vertebral morphology and assessment of the integrity of the posterior ligamentous complex from CT scans.[55] Vertebral fractures may also be secondary to metastasis, and by analyzing imaging features such as the presence of soft-tissue masses and bony destruction, a model was trained to detect malignant vertebral fractures and make diagnostic recommendations.[4,56]

Scoliosis

Scoliosis is defined as a coronal deformity of the spine with Cobb angles of $\geq 10°$. AI can identify adolescent idiopathic scoliosis (AIS) by analyzing whole spinal radiographic images and measuring Cobb angles.[57] Given the concerns over frequent radiographic exposure, automated segmentation of ultrasonography images has been described to facilitate scoliosis diagnosis at the clinic.[8] In another approach, DL was applied upon a light-based depth sensing technology toward radiation-free AIS screening.[15] Separately a platform used software defined radio and noncontact radio frequency signals to detect abnormal coronal and sagittal spinal curvature with an accuracy rate of 99%.[16] AI-driven assessment of skeletal maturity indices, such as the Risser sign[58] or distal radius and ulna index,[11] facilitate decision making in AIS. ML clustering was implemented to classify scoliotic patients based on age, frailty, and mental health, which helped to reduce the heterogeneity of the scoliosis population studies and contribute to the development of individualized treatment plans.[59] Automated models for AIS screening and prognostication promise to triage referral time, follow-up intervals, and identify progressive curves for preemptive treatment.[11]

LDH and lumbar spinal stenosis (LSS)

Automated diagnosis from imaging of lumbar disk bulging and spondylolisthesis achieved an accuracy of more than 90% through MRI scans.[60] Clinically, however, LDH and LSS are diagnosed through symptoms, for example, of sciatica in the former disease entity, and neurogenic claudication in the latter. To overcome the caveat of reaching diagnoses based on lumbar spine MRIs alone,[61,62] a natural language processing algorithm was proposed to distinguish between LDH and LSS from symptomatology recorded in the admission record with an accuracy exceeding 75%.[63]

Ankylosing spondylitis (AS)

AS is a chronic inflammatory disease that predominantly affects the axial skeleton, including spine, pelvis, and hips. Detection of subtle radiological changes in the sacroiliac joint facilitates early diagnosis of AS and enables prompt treatment with disease-modifying drugs. A DL-based classifier was developed to detect bony erosion upon CT scans to diagnose AS resulted on sensitivity and specificity measures surpassing that of an experienced musculoskeletal radiologist by 8.4% and 9.5%, respectively.[7]

Spinal tuberculosis

Spinal tuberculosis remains prevalent in underdeveloped nations. The scarcity of medical resources and expertise with regard to radiological and microbiological diagnosis in endemic areas facilitates disease transmission. A computer-aided diagnostic system was developed to improve the effectiveness of radiological diagnosis of tuberculosis from lumbar CT scans obtained of patients from Tibetan regions of China, resulting in an accuracy of 98.3%.[64]

Ossification of the posterior longitudinal ligament (OPLL) and DCM

Diagnosing OPLL through radiographs alone is challenging. A convolutional neural network (CNN)–based DL approach was employed to detect cervical OPLL on plain radiographs, using CT scans as the reference standard. This demonstrated a diagnostic accuracy that was 20% better than that of spine surgeons.[65] Another study used ML upon cervical MRIs toward diagnosing OPLL and achieved an accuracy of 98%.[66] Separately, DCM is the largest cause of nontraumatic cervical cord stenosis in developed countries.[67] A study proposed smartphone-based video capture of the grip-and-release test followed by DL as a viable screening tool.[19] Finger kinetics were extracted from videos to allow for binary classification of DCM with sensitivity and specificity measures approaching 90%.

Spine tumors and spinal metastases

Distinguishing between benign and malignant vertebral fractures from CT scans can be a challenge, and an AI model demonstrated a high level of discrimination that matched or surpassed that of radiology residents (0.71 < area under curve [AUC] < 0.86).[68] In combination with clinical parameters (age, sex, tumor location), imaging can aid in differentiating benign intradural spinal cord tumors with an Area Under Curve — Receiver Operating Characteristics of 0.92.[12] In spinal oncology, precise segmentation is essential to delineate margins for therapeutic interventions such as radiation and surgery. In delineating the margins of involvement, trained models can facilitate in the grading of metastatic spinal cord compression.[3,69,70]

Osteoporosis

Dual-energy radiograph absorptiometry remains the benchmark for quantification of bone mineral density (BMD). Numerous publications have leveraged more readily available imaging modalities to extrapolate BMD.[6,71-74] A representative study from Taiwan contained 36,279 subjects who underwent dual-energy radiograph absorptiometry scans as well as hip and spine radiographs. External validation of the algorithm demonstrated its consistent and excellent ability to predict BMD from standard radiographs and classify patients with a 95% positive or negative predictive value with osteoporosis.[71] Several studies have focused on automated diagnosis of osteoporotic vertebral fractures using radiographs, achieving an accuracy of >0.9.[75-78] Automatic classification through the Assessment System of Thoracolumbar Osteoporotic Fracture has also been realized with an accuracy ranging from 69.2% to 87.5%.[79] Furthermore, MRI scans may be used as image input for the automated diagnosis of recent osteoporotic vertebral fractures (accuracy = 0.88, AUC = 0.94).[80]

Prognostication

Identifying Surgical Candidates and Predicting Outcomes

AI has been used to identify candidates with LSS who subsequently required surgery, based on axial cuts showing the severity of compression.[81] Other predictive outcomes included AIS curve progression,[11,82] blood loss in thoracolumbar burst fractures patients, spinal cord injury outcomes based on deviation in intraoperative mean arterial pressure,[83,84] fracture recurrence after vertebroplasty,[85] postoperative functional measures and mortality,[2,84,86-93] discharge timing and destination,[90,94-99] and risk of unplanned readmission.[100]

AI in perioperative risk management

In addition to predicting medical status–related postoperative complications such as prolonged hospitalization following spinal deformity surgery and postoperative delirium,[101] AI is also capable of predicting surgical complications including infection at the surgical site,[102,103] loosening of pedicle screws,[104] recurrent herniation following endoscopic disk removal,[105] and junctional breakdown.[106,107] Estimating the risk of such events all contribute toward optimizing patient care[18] and surgical decision making.[108] As examples on the application of ML to prognosticate using multimodal data, functional ambulation following incomplete spinal cord injury was predicted using demographic information, clinical measures, limb acceleration, and inertial sensors.[17,109]

AI's predictive capability in spinal metastases

Patients with spinal metastases and an unknown primary are required to undergo a full range of invasive diagnostic tests that can consume a great deal of money and time. CNN models can identify tumor features upon conventional spine MRI sequences and suggest the most likely site of the primary lesion (AUCROC > 70%).[110] AI also predicts intraoperative blood loss (AUC = 0.809)[111] as well as longer hospital stays and complications by considering preoperative medical and functional status in addition to tumor staging.[112] In addition, the capacity to regain ambulation is a factor influencing whether surgery should be offered, and AUC of a predictive model was approximately 90%.[113] AI may forecast treatment efficacy for stereotactic radiation by analyzing clinical characteristics and MRI data (AUC = 0.828).[13]

Surgical Planning and Assistance

AI assistance facilitates selection of appropriate implant size, as well as the course and length for screws.[114]

Intraoperative navigation of screw placement is also becoming common place. Their benefits include increased precision and enhanced intraoperative efficiency, markedly reducing complications and improving outcomes[115] while obstacles to more widespread integration include costs, additional preoperative investigations, and planning as well as learning curve.[116,117] Although navigation usually requires a preoperative CT scan, there has been a report of MRI-derived synthetic CT images for lumbar surgical planning as an alternative to reduce radiation exposure.[118] Postoperatively, AI can automatically perform measurements of screw trajectory for further analysis.[119]

Artificial Intelligence Learning Algorithms

Supervised Learning and Unsupervised Learning

AI learning-based approaches may be categorized into supervised and unsupervised learning. Supervised ML aims to identify patterns in a prelabeled data set to arrive at a desired outcome, whereas unsupervised ML models are given unlabeled data and allowed to identify patterns without any explicit guidance or instruction. The two approaches may also be combined in semi-supervised learning when labeled data are difficult to obtain. Most AI-based studies were conducted through labeled data sets with supervised learning (n = 100).

Machine learning and deep learning

ML is a branch of AI that applies algorithms upon data to imitate human learning and apply decision making on new and unseen data. Conventional ML commonly relies on manual or software-based feature engineering to extract analytical representations from raw data.[3] DL, a subset of ML, consists of a layered structure of artificial neural networks. Among included studies, 38 used traditional ML methods upon predefined features, whereas 67 used DL to automatically extract features.

ML tasks and algorithms

Twenty-eight studies concerned classification tasks for prognosis prediction, whereas five concerned classification tasks for spinal disease diagnosis or screening. Four ML-based studies regarded prognosis forecasting through regression with continuous (not binary or categorical) outputs. As an example, an ensemble ML model reported spinal cord independence measure using features present at the time of admission for rehabilitation.[92] One study with a mixed task definition reported, respectively, combining classification and regression for DCM screening.[111] One study reported using ML for spinal CT segmentation to reduce computing load with DL methods.[35]

ML-based studies commonly used multiple algorithms, reporting the one with the best performance. Algorithms with best prediction performance included random forest/decision tree (n = 13), gradient boost/XGBoost (n = 6), support vector machine (n = 5), and logistic regression (n = 3). Ensemble ML methods incorporate multiple algorithms to enhance performance and were reported in three studies,[13,92,113] whereas eight studies were based on other ML algorithms.

DL tasks and algorithms

For DL-based literature, 25 studies were defined as classification tasks, 30 studies as MIC tasks (landmark detection, objective detection, and segmentation), 10 studies as composite tasks, and two studies as medical image generation tasks. Among classification tasks, 23 studies concerned diagnosis, whereas two studies concerned prognosis for AIS curve deterioration.[11,82] In studies concerning DL-based MIC, 24 studies were for image segmentation, including 13 with resultant automated measurements. Five studies were there concerning landmark detection on radiographic images for automated measurements, and one study for object detection to identify LDH.[5]

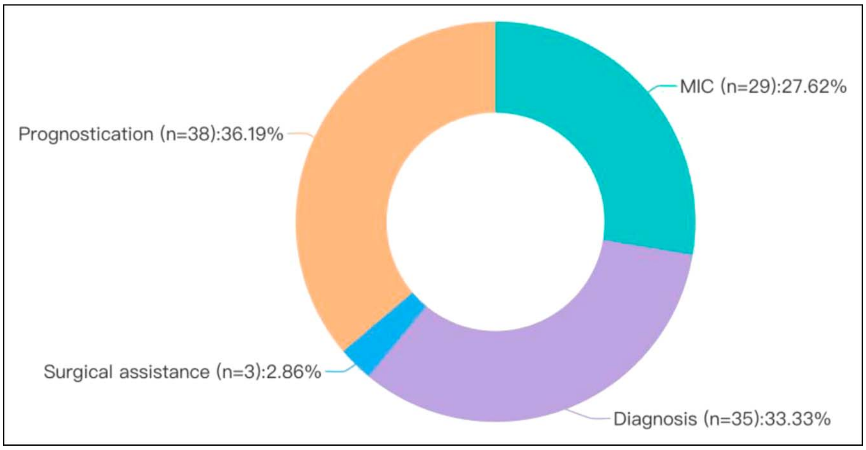
A natural language processing study used long short-term memory method to diagnose LDH and LSS from clinical records.[63] Otherwise, all other DL-based literature conducting medical image analysis with computer vision methods selected CNNs as the basic structure. As more than one DL structure may have been implemented, we quantified commonly reported DL structures. Among included studies, 39 studies used CNN structures and its variants, including Visual Geometry Group, ResNet, HR-Net, graph CNNs and attention-based CNNs. The U-Net family (n = 21) and Mask-RCNN (n = 3) were the most often used DL structures for segmentation tasks, whereas You Only Look Once (n = 5) and Fast/Faster-RCNN (n = 5) were used for object detection tasks. Vision transformer (n = 2) and generative adversarial networks (n = 2) were reported in the literature to address image segmentation and synthesis tasks. Four studies were there concerning ensemble learning to combine multiple DL models or composite inputs for enhanced prediction performance.

Discussion

The Present State of Artificial Intelligence in Spine Surgery

As summarized in Figure 5, the literature concerning AI in spine surgery predominantly concerned prognosis

Figure 5



Donut chart showingmode of application of artificial intelligence in spine surgery.

prediction (36.2%), diagnostic assistance (33.3%), and MIC (27.6%). Conversely, AI's application in surgical assistance appears underrepresented in academic literature (2.9%), likely because progress is of greater relevance to the industry sector and kept proprietary instead being disclosed.

With the focus on patient management, AI-facilitated diagnosis promises to create highly individualized treatment plans, which will be refined when models are developed upon larger, multicenter data sets.[120] From the standpoint of the healthcare provider, an automated workflow will diminish burden on the medical staff, permitting healthcare providers to allocate their expertise more judiciously. AI-facilitated prognostication promises to improve case selection for surgery and, in so doing, improve perioperative outcomes and ensure optimal resource utilization.[121] Intraoperative navigation, robotics, and augmented reality technologies will minimize procedural errors and enhance patient safety; however, they may best be selectively applied to complex cases to reduce expenses, in the hands of a surgical team familiarized with surgery of the relevant technology and its limitations.[122]

Assessing the Quality of Artificial Intelligence–Related Research

The exponential increase in scientific research concerning the application of ML across biology and medical disciplines has led to the imparting of recommendations to improve upon ML assessment and reproducibility.[123] Broadly speaking, these encompass data management (sample size and quality), optimization of the ML model (fitting, parameter tuning), model (interpretation of model, access to source code), and evaluation (provision of performance measures and comparison to other methods).

The AUC upon ROC curves, Brier scores, and prediction squared error are common means to describe the accuracy and discriminatory value of ML models. Calibration and reliability of the model are other performance measures to be cognizant of, which would otherwise result in underestimation or overestimation of risk.[124] Despite the "black box" nature of model training, biological feasibility with regard to the model input, architecture, and results also add weight to clinical application.

As a means to standardize reporting frameworks, guidelines, such as CONSORT-AI, have been proposed for prospective trials, and this is likely the way forward.[125] Although the clinical question and application would be best defined by a clinician, close collaboration with computer science experts in project planning, execution, and interpretation would be essential. The challenges of big data classification, processing massive data sets, caveats in supervised and unsupervised ML approaches, feature selection, model selection, and consideration of external validity have been summarized elsewhere.[126]

Limitations and Outlook

Like all innovations, AI's integration into spine surgery also has its challenges. The efficacy of an AI model is dependent on the quality and comprehensiveness of its training data set,[127] which would be incorporating a standardized pipeline for data collection and storage within routine clinical practice. Most studies included in this scoping review are retrospective. Furthermore, multicenter validation was lacking, with most data sources for AI training coming from a single hospital. The heterogeneity of input data offers a challenge for AI, especially when the clinical context, population, or disease is rare thereby presenting a data bottleneck.[128]

Conclusion

An increased incorporation of AI into clinical practice is inevitable and exciting. The present generation of clinicians needs to be aware of the potential and pitfalls of AI-related technology to selectively and effectively harness its benefits.

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 2009;6:e1000097.

2. Song J, Li J, Zhao R, Chu X: Developing predictive models for surgical outcomes in patients with degenerative cervical myelopathy: A comparison of statistical and machine learning approaches. *Spine J* 2024;24:57-67.

3. 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;127:518-525.

4. 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: 5069-5076.

5. Tsai J-Y, Hung IY-J, Guo YL, et al: Lumbar disc herniation automatic detection in magnetic resonance imaging based on deep learning. *Front Bioeng Biotechnol* 2021;9:708137.

6. Zhang B, Chen Z, Yan R, et al: Development and validation of a feature-based broad-learning system for opportunistic osteoporosis screening using lumbar spine radiographs. *Acad Radiol* 2024;31:84-92.

7. Castro-Zunti R, Park EH, Choi Y, Jin GY, Ko SB: Early detection of ankylosing spondylitis using texture features and statistical machine learning, and deep learning, with some patient age analysis. *Comput Med Imaging Graph* 2020;82:101718.

8. Jiang W, Mei F, Xie Q: Novel automated spinal ultrasound segmentation approach for scoliosis visualization. *Front Physiol* 2022;13:1051808.

9. Huang Y, Jiao J, Yu J, Zheng Y, Wang Y: Si-MSPDNet: A multiscale Siamese network with parallel partial decoders for the 3-D measurement of spines in 3D ultrasonic images. *Comput Med Imaging Graph* 2023;108: 102262.

10. Hong N, Cho SW, Shin S, et al: Deep-learning-based detection of vertebral fracture and osteoporosis using lateral spine X-ray radiography. *J Bone Mineral Res* 2023;38:887-895.

11. Wang H, Zhang T, Zhang C, et al: An intelligent composite model incorporating global/regional X-rays and clinical parameters to predict progressive adolescent idiopathic scoliosis curvatures and facilitate population screening. *EBioMedicine* 2023;95:104768.

12. Kita K, Fujimori T, Suzuki Y, et al: Bimodal artificial intelligence using TabNet for differentiating spinal cord tumors-Integration of patient background information and images. *iScience* 2023;26:107900.

13. Chen Y, Qin S, Zhao W, et al: MRI feature-based radiomics models to predict treatment outcome after stereotactic body radiotherapy for spinal metastases. *Insights Imaging* 2023;14:169.

14. Staartjes VE, Seevinck PR, Vandertop WP, van Stralen M, Schröder ML: Magnetic resonance imaging–based synthetic computed tomography of the lumbar spine for surgical planning: A clinical proof-of-concept. *Neurosurg Focus* 2021;50:E13.

15. Meng N, Wong KYK, Zhao M, Cheung JPY, Zhang T: Radiograph-comparable image synthesis for spine alignment analysis using deep learning with prospective clinical validation. *EClinicalMedicine* 2023;61: 102050.

16. Mustafa A, Ullah F, Rehman MU, et al: Non-intrusive RF sensing for early diagnosis of spinal curvature syndrome disorders. *Comput Biol Med* 2023;155:106614.

17. Werner C, Gönel M, Lerch I, Curt A, Demkó L: Data-driven characterization of walking after a spinal cord injury using inertial sensors. *J Neuroeng Rehabil* 2023;20:55.

18. Zheng S, Liang G, Chen J, Duan Q, Chang Y: Severity assessment of cervical spondylotic myelopathy based on intelligent video analysis. *IEEE J Biomed Health Inform* 2022;26:4486-4496.

19. Ibara T, Matsui R, Koyama T, et al: Screening for degenerative cervical myelopathy with the 10-second grip-and-release test using a smartphone and machine learning: A pilot study. *Digit Health* 2023;9: 20552076231179030.

20. Chen X, Deng Q, Wang Q, et al: Image quality control in lumbar spine radiography using enhanced U-net neural networks. *Front Public Health* 2022;10:891766.

21. Kim KC, Cho HC, Jang TJ, Choi JM, Seo JK: Automatic detection and segmentation of lumbar vertebrae from X-ray images for compression fracture evaluation. *Comput Methods Programs Biomed* 2021;200: 105833.

22. Tao R, Liu W, Zheng G: Spine-transformers: Vertebra labeling and segmentation in arbitrary field-of-view spine CTs via 3D transformers. *Med Image Anal* 2022;75:102258.

23. Da Mutten R, Zanier O, Theiler S, et al: Whole spine segmentation using object detection and semantic segmentation. *Neurospine* 2024;21:57-67.

24. Netherton TJ, Nguyen C, Cardenas CE, et al: An automated treatment planning framework for spinal radiation therapy and vertebral-level second check. *Int J Radiat Oncol Biol Phys* 2022;114:516-528.

25. Zhang Y, Shi Z, Wang H, et al: LumVertCancNet: A novel 3D lumbar vertebral body cancellous bone location and segmentation method based on hybrid Swin-transformer. *Comput Biol Med* 2024;171:108237.

26. Pang S, Pang C, Su Z, et al: DGMSNet: Spine segmentation for MR image by a detection-guided mixed-supervised segmentation network. *Med Image Anal* 2022;75:102261.

27. Zhang X, Yang Y, Shen YW, et al: SeUneter: Channel attentive U-Net for instance segmentation of the cervical spine MRI medical image. *Front Physiol* 2022;13:1081441.

28. Huang J, Shen H, Wu J, et al: Spine explorer: A deep learning based fully automated program for efficient and reliable quantifications of the vertebrae and discs on sagittal lumbar spine MR images. *Spine J* 2020;20: 590-599.

29. Das P, Pal C, Acharyya A, Chakrabarti A, Basu S: Deep neural network for automated simultaneous intervertebral disc (IVDs) identification and segmentation of multi-modal MR images. *Comput Methods Programs Biomed* 2021;205:106074.

30. Sáenz-Gamboa JJ, Domenech J, Alonso-Manjarrés A, Gómez JA, de la Iglesia-Vayá M: Automatic semantic segmentation of the lumbar spine: Clinical applicability in a multi-parametric and multi-center study on magnetic resonance images. *Artif Intell Med* 2023;140:102559.

31. Shen H, Huang J, Zheng Q, et al: A deep-learning-based, fully automated program to segment and quantify major spinal components on axial lumbar spine magnetic resonance images. *Phys Ther* 2021;101:pzab041.

32. Zhou J, Damasceno PF, Chachad R, et al: Automatic vertebral body segmentation based on deep learning of Dixon images for bone marrow fat fraction quantification. *Front Endocrinol (Lausanne)* 2020;11:612.

33. Gao KT, Tibrewala R, Hess M, et al: Automatic detection and voxel-wise mapping of lumbar spine Modic changes with deep learning. *JOR Spine* 2022;5:e1204.

34. Shea KG, Stevens PM, Nelson M, Smith JT, Masters KS, Yandow S: A comparison of manual versus computer-assisted radiographic measurement. Intraobserver measurement variability for Cobb angles. *Spine (Phila Pa 1976)* 1998;23:551-555.

35. Jimenez-Pastor A, Alberich-Bayarri A, Fos-Guarinos B, et al: Automated vertebrae localization and identification by decision forests and image-based refinement on real-world CT data. *Radiol Med* 2020;125:48-56.

36. Wang Y, Huang L, Wu M, Liu S, Jiao J, Bai T: Multi-input adaptive neural network for automatic detection of cervical vertebral landmarks on X-rays. *Comput Biol Med* 2022;146:105576.

37. Wills BP, Auerbach JD, Zhu X, et al: Comparison of Cobb angle measurement of scoliosis radiographs with preselected end vertebrae: Traditional versus digital acquisition. *Spine (Phila Pa 1976)* 2007;32:98-105.

38. Liu J, Yuan C, Sun X, Sun L, Dong H, Peng Y: The measurement of Cobb angle based on spine X-ray images using multi-scale convolutional neural network. *Phys Eng Sci Med* 2021;44:809-821.

39. 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 Inform* 2023;27:3002-3013.

40. van der Graaf JW, van Hooff ML, van Ginneken B, et al: Development and validation of AI-based automatic measurement of coronal Cobb angles in degenerative scoliosis using sagittal lumbar MRI. *Eur Radiol* 2024;34:5748-5757.

41. Meng N, Cheung JPY, Wong KYK, et al: An artificial intelligence powered platform for auto-analyses of spine alignment irrespective of image quality with prospective validation. *EClinicalMedicine* 2022;43:101252.

42. Yuh WT, Khil EK, Yoon YS, et al: Deep learning-assisted quantitative measurement of thoracolumbar fracture features on lateral radiographs. *Neurospine* 2024;21:30-43.

43. Zerouali M, Parpaleix A, Benbakoura M, Rigault C, Champsaur P, Guenoun D: Automatic deep learning-based assessment of spinopelvic coronal and sagittal alignment. *Diagn Interv Imaging* 2023;104:343-350.

44. Kim YT, Jeong TS, Kim YJ, Kim WS, Kim KG, Yee GT: Automatic spine segmentation and parameter measurement for radiological analysis of whole-spine lateral radiographs using deep learning and computer vision. *J Digit Imaging* 2023;36:1447-1459.

45. Hallinan J, 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;300:130-138.

46. Su ZH, Liu J, Yang MS, et al: Automatic grading of disc herniation, central canal stenosis and nerve roots compression in lumbar magnetic resonance image diagnosis. *Front Endocrinol (Lausanne)* 2022;13:890371.

47. Han Z, Wei B, Xi X, Chen B, Yin Y, Li S: Unifying neural learning and symbolic reasoning for spinal medical report generation. *Med Image Anal* 2021;67:101872.

48. Lim DSW, Makmur A, Zhu L, et al: Improved productivity using deep learning-assisted reporting for lumbar spine MRI. *Radiology* 2022;305:160-166.

49. Xu F, Xiong Y, Ye G, et al: Deep learning-based artificial intelligence model for classification of vertebral compression fractures: A multicenter diagnostic study. *Front Endocrinol (Lausanne)* 2023;14:1025749.

50. Germann C, Meyer AN, Staib M, Sutter R, Fritz B: Performance of a deep convolutional neural network for MRI-based vertebral body measurements and insufficiency fracture detection. *Eur Radiol* 2023;33:3188-3199.

51. Nicolaes J, Skjødt MK, Raeymaeckers S, et al: Towards improved identification of vertebral fractures in routine computed tomography (CT) scans: Development and external validation of a machine learning algorithm. *J Bone Miner Res* 2023;38:1856-1866.

52. Choi E, Park D, Son G, et al: Weakly supervised deep learning for diagnosis of multiple vertebral compression fractures in CT. *Eur Radiol* 2023;34:3750-3760.

53. Dong Q, Luo G, Lane NE, et al: Deep learning classification of spinal osteoporotic compression fractures on radiographs using an adaptation of the genant semiquantitative criteria. *Acad Radiol* 2022;29:1819-1832.

54. Zhang J, Liu F, Xu J, et al: Automated detection and classification of acute vertebral body fractures using a convolutional neural network on computed tomography. *Front Endocrinol (Lausanne)* 2023;14:1132725.

55. Doerr SA, Weber-Levine C, Hersh AM, et al: Automated prediction of the Thoracolumbar Injury Classification and Severity Score from CT using a novel deep learning algorithm. *Neurosurg Focus* 2022;52:E5.

56. Li Y, Zhang Y, Zhang E, et al: Differential diagnosis of benign and malignant vertebral fracture on CT using deep learning. *Eur Radiol* 2021;31:9612-9619.

57. Wu C, Meng G, Lian J, et al: A multi-stage ensemble network system to diagnose adolescent idiopathic scoliosis. *Eur Radiol* 2022;32:5880-5889.

58. Magnide E, Tchaha GW, Joncas J, et al: Automatic bone maturity grading from EOS radiographs in Adolescent Idiopathic Scoliosis. *Comput Biol Med* 2021;136:104681.

59. Mohanty S, Hassan FM, Lenke LG, et al: Machine learning clustering of adult spinal deformity patients identifies four prognostic phenotypes: A multicenter prospective cohort analysis with single surgeon external validation. *Spine J* 2024;24:1095-1108.

60. Xuan J, Ke B, Ma W, Liang Y, Hu W: Spinal disease diagnosis assistant based on MRI images using deep transfer learning methods. *Front Public Health* 2023;11:1044525.

61. Tsai JY, Hung IY, Guo YL, et al: Lumbar disc herniation automatic detection in magnetic resonance imaging based on deep learning. *Front Bioeng Biotechnol* 2021;9:708137.

62. Bharadwaj UU, Christine M, Li S, et al: Deep learning for automated, interpretable classification of lumbar spinal stenosis and facet arthropathy from axial MRI. *Eur Radiol* 2023;33:3435-3443.

63. Ren G, Yu K, Xie Z, et al: Differentiation of lumbar disc herniation and lumbar spinal stenosis using natural language processing–based machine learning based on positive symptoms. *Neurosurg Focus* 2022;52:E7.

64. Li Z, Wu F, Hong F, et al: Computer-aided diagnosis of spinal Tuberculosis from CT images based on deep learning with multimodal feature fusion. *Front Microbiol* 2022;13:823324.

65. Ogawa T, Yoshii T, Oyama J, et al: Detecting ossification of the posterior longitudinal ligament on plain radiographs using a deep convolutional neural network: A pilot study. *Spine J* 2022;22:934-940.

66. Shemesh S, Kimchi G, Yaniv G, Harel R: MRI-based detection of cervical ossification of the posterior longitudinal ligament using a novel automated machine learning diagnostic tool. *Neurosurg Focus* 2023;54:E11.

67. Badhiwala JH, Ahuja CS, Akbar MA, et al: Degenerative cervical myelopathy - update and future directions. *Nat Rev Neurol* 2020;16:108-124.

68. Foreman SC, Schinz D, El Husseini M, et al: Deep learning to differentiate benign and malignant vertebral fractures at multidetector CT. *Radiology* 2024;310:e231429.

69. Hallinan J, Zhu L, Zhang W, et al: Deep learning model for grading metastatic epidural spinal cord compression on staging CT. *Cancers (Basel)* 2022;14:3219.

70. Liu H, Jiao M, Yuan Y, et al: Benign and malignant diagnosis of spinal tumors based on deep learning and weighted fusion framework on MRI. *Insights Imaging* 2022;13:87.

71. Hsieh CI, Zheng K, Lin C, et al: Automated bone mineral density prediction and fracture risk assessment using plain radiographs via deep learning. *Nat Commun* 2021;12:5472.

72. Mao L, Xia Z, Pan L, et al: Deep learning for screening primary osteopenia and osteoporosis using spine radiographs and patient clinical covariates in a Chinese population. *Front Endocrinol (Lausanne)* 2022;13: 971877.

73. Shen L, Gao C, Hu S, et al: Using artificial intelligence to diagnose osteoporotic vertebral fractures on plain radiographs. *J Bone Miner Res* 2023;38:1278-1287.

74. Xue L, Qin G, Chang S, et al: Osteoporosis prediction in lumbar spine X-ray images using the multi-scale weighted fusion contextual transformer network. *Artif Intell Med* 2023;143:102639.

75. Li YC, Chen HH, Horng-Shing Lu H, Hondar Wu HT, Chang MC, Chou PH: Can a deep-learning model for the automated detection of vertebral fractures approach the performance level of human subspecialists? *Clin Orthop Relat Res* 2021;479:1598-1612.

76. Hong N, Cho SW, Shin S, et al: Deep-learning-based detection of vertebral fracture and osteoporosis using lateral spine X-ray radiography. *J Bone Miner Res* 2023;38:887-895.

77. Dong Q, Luo G, Lane NE, et al: Generalizability of deep learning classification of spinal osteoporotic compression fractures on radiographs using an adaptation of the modified-2 algorithm-based qualitative criteria. *Acad Radiol* 2023;30:2973-2987.

78. Ryu SM, Lee S, Jang M, et al: Diagnosis of osteoporotic vertebral compression fractures and fracture level detection using multitask learning with U-Net in lumbar spine lateral radiographs. *Comput Struct Biotechnol J* 2023;21:3452-3458.

79. Zhang J, Xia L, Liu J, et al: Exploring deep learning radiomics for classifying osteoporotic vertebral fractures in X-ray images. *Front Endocrinol (Lausanne)* 2024;15:1370838.

80. Yabu A, Hoshino M, Tabuchi H, et al: Using artificial intelligence to diagnose fresh osteoporotic vertebral fractures on magnetic resonance images. *Spine J* 2021;21:1652-1658.

81. Wilson B, Gaonkar B, Yoo B, et al: Predicting spinal surgery candidacy from imaging data using machine learning. *Neurosurgery* 2021;89: 116-121.

82. Wang H, Zhang T, Cheung KM, Shea GK: Application of deep learning upon spinal radiographs to predict progression in adolescent idiopathic scoliosis at first clinic visit. *EClinicalMedicine* 2021;42:101220.

83. Yang B, Gao L, Wang X, et al: Application of supervised machine learning algorithms to predict the risk of hidden blood loss during the perioperative period in thoracolumbar burst fracture patients complicated with neurological compromise. *Front Public Health* 2022;10:969919.

84. Torres-Espín A, Haefeli J, Ehsanian R, et al: Topological network analysis of patient similarity for precision management of acute blood pressure in spinal cord injury. *Elife* 2021:10:e68015.

85. Dong ST, Zhu J, Yang H, Huang G, Zhao C, Yuan B: Development and internal validation of supervised machine learning algorithm for predicting the risk of recollapse following minimally invasive kyphoplasty in osteoporotic vertebral compression fractures. *Front Public Health* 2022; 10:874672.

86. 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: 1659-1669.

87. Khan O, Badhiwala JH, Akbar MA, Fehlings MG: Prediction of worse functional status after surgery for degenerative cervical myelopathy: A machine learning approach. *Neurosurgery* 2021;88:584-591.

88. Yagi M, Michikawa T, Yamamoto T, et al: Development and validation of machine learning-based predictive model for clinical outcome of decompression surgery for lumbar spinal canal stenosis. *Spine J* 2022;22: 1768-1777.

89. Toyoda H, Terai H, Yamada K, et al: A decision tree analysis to predict clinical outcome of minimally invasive lumbar decompression surgery for lumbar spinal stenosis with and without coexisting spondylolisthesis and scoliosis. *Spine J* 2023;23:973-981.

90. Karabacak M, Margetis K: Precision medicine for traumatic cervical spinal cord injuries: Accessible and interpretable machine learning models to predict individualized in-hospital outcomes. *Spine J* 2023;23: 1750-1763.

91. Schonnagel L, Caffard T, Vu-Han TL, et al: Predicting postoperative outcomes in lumbar spinal fusion: Development of a machine learning model. *Spine J* 2024;24:239-249.

92. Kato C, Uemura O, Sato Y, Tsuji T: Functional outcome prediction after spinal cord injury using ensemble machine learning. *Arch Phys Med Rehabil* 2024;105:95-100.

93. Karabacak M, Schupper AJ, Carr MT, Bhimani AD, Steinberger J, Margetis K: Development and internal validation of machine learning models for personalized survival predictions in spinal cord glioma patients. *Spine J* 2024;24:1065-1076.

94. Russo GS, Canseco JA, Chang M, et al: A novel scoring system to predict length of stay after anterior cervical discectomy and fusion. *J Am Acad Orthop Surg* 2021;29:758-766.

95. Valliani AA, Feng R, Martini ML, et al: Pragmatic prediction of excessive length of stay after cervical spine surgery with machine learning and validation on a national scale. *Neurosurgery* 2022;91:322-330.

96. Etzel CM, Veeramani A, Zhang AS, et al: Supervised machine learning for predicting length of stay after lumbar arthrodesis: A comprehensive artificial intelligence approach. *J Am Acad Orthop Surg* 2022;30:125-132.

97. Shahrestani S, Chan AK, Bisson EF, et al: Developing nonlinear k-nearest neighbors classification algorithms to identify patients at high risk of increased length of hospital stay following spine surgery. *Neurosurg Focus* 2023;54:E7.

98. Kato C, Uemura O, Sato Y, Tsuji T: Decision tree analysis accurately predicts discharge destination after spinal cord injury rehabilitation. *Arch Phys Med Rehabil* 2024;105:88-94.

99. Arora A, Demb J, Cummins DD, Callahan M, Clark AJ, Theologis AA: Predictive models to assess risk of extended length of stay in adults with spinal deformity and lumbar degenerative pathology: Development and internal validation. *Spine J* 2023;23:457-466.

100. Martini ML, Neifert SN, Oermann EK, et al: Machine learning with feature domains elucidates candidate drivers of hospital readmission following spine surgery in a large single-center patient cohort. *Neurosurgery* 2020;87:E500-e510.

101. Zhang Y, Wan DH, Chen M, et al: Automated machine learning-based model for the prediction of delirium in patients after surgery for degenerative spinal disease. *CNS Neurosci Ther* 2023;29:282-295.

102. Zhang Q, Chen G, Zhu Q, et al: Construct validation of machine learning for accurately predicting the risk of postoperative surgical site infection following spine surgery. *J Hosp Infect* 2024;146:232-241.

103. Lu K, Tu Y, Su S, et al: Machine learning application for prediction of surgical site infection after posterior cervical surgery. *Int Wound J* 2024;21: e14607.

104. Jiang F, Li X, Liu L, Xie Z, Wu X, Wang Y: Automated machine learning-based model for the prediction of pedicle screw loosening after degenerative lumbar fusion surgery. *Biosci Trends* 2024;18:83-93.

105. Jia M, Lai J, Li K, et al: Optimizing prediction accuracy for early recurrent lumbar disc herniation with a directional mutation-guided SVM model. *Comput Biol Med* 2024;173:108297.

106. Lee CH, Jo DJ, Oh JK, et al: Development and validation of an online calculator to predict proximal junctional kyphosis after adult spinal deformity surgery using machine learning. *Neurospine* 2023;20:1272-1280.

107. Peng L, Lan L, Xiu P, et al: Prediction of proximal junctional kyphosis after posterior scoliosis surgery with machine learning in the lenke 5 adolescent idiopathic scoliosis patient. *Front Bioeng Biotechnol* 2020;8:559387.

108. Karhade AV, Fogel HA, Cha TD, et al: Development of prediction models for clinically meaningful improvement in PROMIS scores after lumbar decompression. *Spine J* 2021;21:397-404.

109. Rigot SK, Boninger ML, Ding D, et al: Toward improving the prediction of functional ambulation after spinal cord injury through the inclusion of limb accelerations during sleep and personal factors. *Arch Phys Med Rehabil* 2022;103:676-687.e6.

110. Liu K, Qin S, Ning J, et al: Prediction of primary tumor sites in spinal metastases using a ResNet-50 convolutional neural network based on MRI. *Cancers (Basel)* 2023;15:2974.

111. Shi X, Cui Y, Wang S, Pan Y, Wang B, Lei M: Development and validation of a web-based artificial intelligence prediction model to assess massive intraoperative blood loss for metastatic spinal disease using machine learning techniques. *Spine J* 2024;24:146-160.

112. Karabacak M, Margetis K: A machine learning-based online prediction tool for predicting short-term postoperative outcomes following spinal tumor resections. *Cancers (Basel)* 2023;15:812.

113. Cui Y, Shi X, Qin Y, et al: Establishment and validation of an interactive artificial intelligence platform to predict postoperative ambulatory status for patients with metastatic spinal disease: A multicenter analysis. *Int J Surg* 2024;110:2738-2756.

114. Scherer M, Kausch L, Ishak B, et al: Development and validation of an automated planning tool for navigated lumbosacral pedicle screws using a convolutional neural network. *Spine J* 2022;22:1666-1676.

115. Sun WX, Huang WQ, Li HY, et al: Clinical efficacy of robotic spine surgery: An updated systematic review of 20 randomized controlled trials. *EFORT Open Rev* 2023;8:841-853.

116. Suarez-Ahedo C, Lopez-Reyes A, Martinez-Armenta C, et al: Revolutionizing orthopedics: A comprehensive review of robot-assisted surgery, clinical outcomes, and the future of patient care. *J Robot Surg* 2023;17:2575-2581.

117. Passias PG, Brown AE, Alas H, et al: A cost benefit analysis of increasing surgical technology in lumbar spine fusion. *Spine J* 2021;21:193-201.

118. Staartjes VE, Seevinck PR, Vandertop WP, van Stralen M, Schröder ML: Magnetic resonance imaging-based synthetic computed tomography of the lumbar spine for surgical planning: A clinical proof-of-concept. *Neurosurg Focus* 2021;50:E13.

119. Mulford KL, Regan C, Nolte CP Jr, Pinter ZW, Milbrandt TA, Larson AN: Automated measurements of interscrew angles in vertebral body tethering patients with deep learning. *Spine J* 2024;24:333-339.

120. Rajkomar A, Dean J, Kohane I: Machine learning in medicine. *N Engl J Med* 2019;380:1347-1358.

121. Kunze KN, Karhade AV, Polce EM, Schwab JH, Levine BR: Development and internal validation of machine learning algorithms for predicting complications after primary total hip arthroplasty. *Arch Orthop Trauma Surg* 2023;143:2181-2188.

122. Molliqaj G, Schatlo B, Alaid A, et al: Accuracy of robot-guided versus freehand fluoroscopy-assisted pedicle screw insertion in thoracolumbar spinal surgery. *Neurosurg Focus* 2017;42:E14.

123. Walsh I, Fishman D, Garcia-Gasulla D, et al: DOME: Recommendations for supervised machine learning validation in biology. *Nat Methods* 2021;18:1122-1127.

124. Huang C, Li SX, Caraballo C, et al: Performance metrics for the comparative analysis of clinical risk prediction models employing machine learning. *Circ Cardiovasc Qual Outcomes* 2021;14:e007526.

125. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group: Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nat Med* 2020/09/01 2020;26:1364-1374.

126. Shehab M, Abualigah L, Shambour Q, et al: Machine learning in medical applications: A review of state-of-the-art methods. *Comput Biol Med* 2022;145:105458.

127. Shah NH, Milstein A, Bagley PhD SC: Making machine learning models clinically useful. *JAMA* 2019;322:1351-1352.

128. Topol EJ: High-performance medicine: The convergence of human and artificial intelligence. *Nat Med* 2019;25:44-56.