



Review Article

Corresponding Author

Joon-Yong Jung

<https://orcid.org/0000-0002-6909-0919>

Visual Analysis and Learning for Improved Diagnostics (VALID) Lab, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Korea
Email: jy.jung@songeui.ac.kr

Received: April 16, 2024

Revised: May 14, 2024

Accepted: May 23, 2024



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2024 by the Korean Spinal Neurosurgery Society

Artificial Intelligence in Spinal Imaging and Patient Care: A Review of Recent Advances

Sungwon Lee^{1,2}, Joon-Yong Jung^{1,2}, Akaworn Mahatthanatrakul³, Jin-Sung Kim⁴

¹Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea

²Visual Analysis and Learning for Improved Diagnostics (VALID) Lab, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea

³Department of Orthopaedics, Faculty of Medicine, Naresuan University Hospital, Phitsanulok, Thailand

⁴Spine Center, Department of Neurosurgery, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea

Artificial intelligence (AI) is transforming spinal imaging and patient care through automated analysis and enhanced decision-making. This review presents a clinical task-based evaluation, highlighting the specific impact of AI techniques on different aspects of spinal imaging and patient care. We first discuss how AI can potentially improve image quality through techniques like denoising or artifact reduction. We then explore how AI enables efficient quantification of anatomical measurements, spinal curvature parameters, vertebral segmentation, and disc grading. This facilitates objective, accurate interpretation and diagnosis. AI models now reliably detect key spinal pathologies, achieving expert-level performance in tasks like identifying fractures, stenosis, infections, and tumors. Beyond diagnosis, AI also assists surgical planning via synthetic computed tomography generation, augmented reality systems, and robotic guidance. Furthermore, AI image analysis combined with clinical data enables personalized predictions to guide treatment decisions, such as forecasting spine surgery outcomes. However, challenges still need to be addressed in implementing AI clinically, including model interpretability, generalizability, and data limitations. Multicenter collaboration using large, diverse datasets is critical to advance the field further. While adoption barriers persist, AI presents a transformative opportunity to revolutionize spinal imaging workflows, empowering clinicians to translate data into actionable insights for improved patient care.

Keywords: Artificial intelligence, Machine learning, Deep learning, Spine, Patient care, Clinical decision-making

INTRODUCTION

Providing optimal spine patient care is becoming increasingly complex due to the rapid growth of patient data, the rising number of spine patients, and expanding treatment options. The large amount of information from medical imaging and electronic health records (EMRs), combined with growing patient volumes driven by an aging population, presents a major challenge for efficient data processing and analysis. At the same

time, the increase in treatment methods, from minimally invasive procedures to personalized medicine, demands careful consideration of risks, benefits, and suitability for each individual. Traditional methods may struggle in this situation, highlighting artificial intelligence (AI)'s potential to process large amounts of data, identify patterns, and support evidence-based decision-making for more efficient, personalized, and effective spine care.^{1,2}

Medical research is undergoing a revolution. AI, powered by

machine learning (ML) and deep learning (DL), is unlocking new discoveries with the potential to improve diagnoses, treatments, and patient outcomes. While spine image research faced early challenges due to complex anatomical structures, data scarcity, and nonstandardized protocols,³ the field has witnessed a remarkable surge in research and commercially available solutions, particularly in the past 5 or 6 years. This rapid advancement is readily apparent in a PubMed search, where a staggering 86% of all AI spine research papers have been published since 2017.

While traditional research methods like regression and correlation models remain valuable tools, the field is experiencing a paradigm shift towards AI algorithms, particularly ML, DL, and generative models. Compared to regression models, which are limited to predicting simple linear relationships, ML can learn complex, nonlinear interactions and hidden patterns within data.⁴ DL builds on this by utilizing numerous layers of computations to extract more complex features from high-dimensional data, such as image data. While both ML and DL can be used to build clinical decision support systems or prognosis prediction models, DL also excels at extracting meaningful features from images, enabling applications like landmark detection and disease classification. Generative models, such as generative adversarial networks and large language models (LLMs), go even further. They learn the underlying patterns and relationships within data and then use this knowledge to create entirely new data.⁵ This opens up possibilities like synthetic computed tomography (CT) images generated from magnetic resonance imaging (MRI) images,⁶ radiology reports written by AI,⁷ and chatbots to answer frequently asked questions for patients.⁸

This review focuses on recent advancements in AI-powered spine image analysis and patient care. However, instead of focusing on the role of a specific AI algorithm, our aim is to cover this from the perspective of a clinical task-based manner (Fig. 1). To do this, we searched PubMed and Google Scholar for relevant articles published between January 2017 and March 2024. Our search query combined terms related to AI, ML, and DL with those related to the spine (including “spine” or “vertebra”) and spinal imaging modalities (including “radiograph,” “CT,” and “MR”). We then assessed the retrieved studies for their methodologies, evaluation metrics, and contributions to spine surgery or research. Additionally, we hand-searched the reference lists of included articles to identify relevant studies not captured by the initial database search. This process resulted in a final selection of 134 articles for review.

CURRENT STATUS AND CLINICAL APPLICATIONS ACROSS PATIENT CARE PROCESS

1. Image Improvement

The improvement of image quality through AI is a well-established field already implemented in clinical practice, particularly in CT and MRI. Traditional interpolation techniques based on simple mathematics artificially increase the resolution but often result in blurring. DL has taken this process to the next level, generating more realistic and sharper images.⁹ This AI-based interpolation scheme is incorporated as a part of DL-based image improvement protocols in the latest scanners.

In CT scans, the radiation dose follows the “As low as reason-

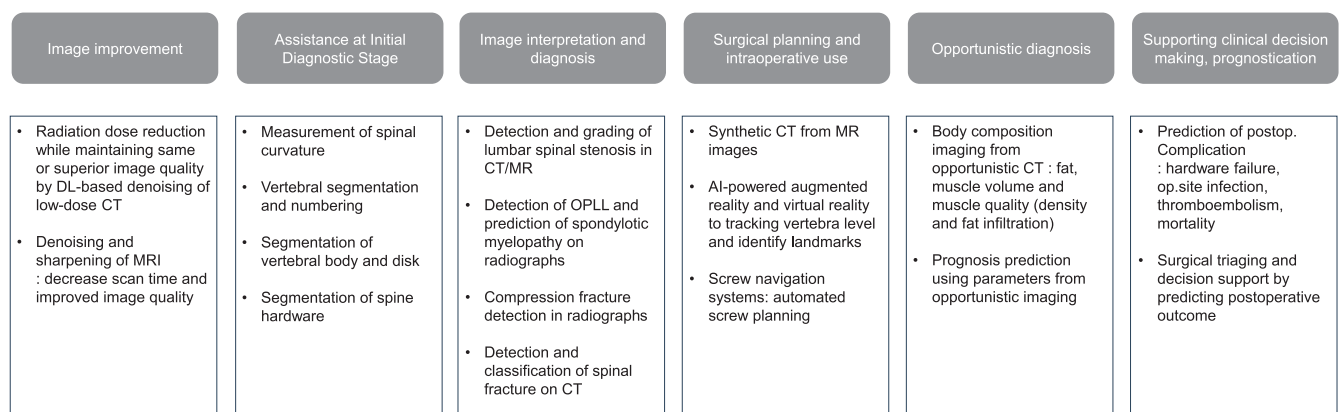


Fig. 1. Artificial intelligence (AI) applications in spinal imaging and patient care. This figure summarizes the 6 clinical applications of AI in spinal imaging and patient care, following the order they are introduced in this article. Each application contributes to various stages of patient care. DL, deep learning; CT, computed tomography; MRI, magnetic resonance imaging; MR, magnetic resonance; OPLL, ossification of posterior longitudinal ligament.

ably achievable” principle. Lower doses typically lead to noisier images. Fortunately, AI algorithms now effectively reduce this noise postacquisition, enabling high-quality images with reduced radiation doses.¹⁰

In MRI, scan time serves as a trade-off for image quality, analogous to the radiation dose in CT: Achieving high image quality often requires longer scan time, leading to patient discomfort and potential movement artifacts. Traditional techniques like parallel imaging and compressed sensing address scan time reduction but still struggle with noises and artifacts. Recent advancements in DL-based reconstruction have overcome these challenges. By combining traditional schemes with DL-based optimization, DL reconstruction significantly improves signal-to-noise ratio with significantly reduced scan times.^{11,12} In spine MRI, DL-reconstructed images have been found to be interchangeable with standard MRI for detecting various abnormalities, offering excellent image quality with a remarkable 70% reduction in scan time.¹³ Major magnetic resonance (MR) vendors have already incorporated these features, and vendor-neutral solutions for spine MR are also commercially available.¹³

However, some concerns remain regarding the potential for DL-based improvements to alter lesion details or introduce artifacts.^{13,14} Ongoing research and validation are crucial to ensure the reliability and safety of these AI-powered tools in clinical practice.

2. Assistance at the Initial Diagnostic Stage

Spinal imaging plays a critical role in the initial step of patient care, with measurements and landmark locations being one of the fundamental information. This information guides treatment planning and risk assessment, ensuring precise and objective decision-making for effective patient management. For AI-based image analysis, accurate segmentation of the vertebral body and disc spaces is a critical first step. Segmentation assists in spine numbering and data preparation for other spine analysis algorithms. This is why structure segmentation models have been a significant focus of spine AI research, with numerous public datasets available to train them.¹⁵ Additionally, recent breakthroughs in natural image segmentation techniques have nudged developments of automated segmentation tools for spine surgical hardware, achieving near-perfect Dice scores (Fig. 2).¹⁶

Assessment of kyphosis, lordosis, and scoliosis is another primary indication of spinal imaging. However, accurate assessment of spinal curvature relies on hand-measured parameters, such as Cobb angle, pelvic incidence (PI), sacral slope, pelvic tilt, lumbar lordosis (LL), and sagittal vertical axis. This is a la-

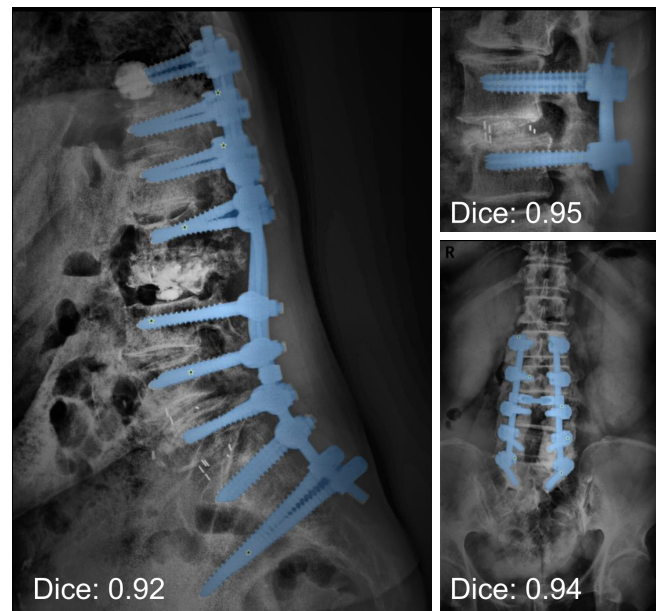


Fig. 2. Adaptation of personalized SAM (PerSAM) model for segmentation of surgical hardware (blue) in spine radiographs. This model utilizes a personalized segment anything model to automatically generate prompts (green stars) for pixel-wise segmentations of spine surgical hardware in spine radiographs. The segmentations achieved Dice scores exceeding 0.9.

borious task, prone to high interobserver agreements, and thus has been one of the early targets of spine AI research. Like human analysis, supervised DL models can be trained to identify or segment specific landmarks on anteroposterior (AP) and lateral spine radiographs, automatically connecting them to calculate spinal parameters. Advanced models can now detect up to 78 landmarks and 18 spinopelvic parameters in whole spine lateral radiographs, demonstrating excellent agreement with human measurements.¹⁷⁻¹⁹ In AP radiographs, automated Cobb angles exhibit lower mean error compared to human intra-/interobservers (2° – 6.32° vs. $\pm 9.6^\circ/\pm 11.8^\circ$), showing potential for scoliosis screening in children (sensitivity, 95.7%; specificity, 88.1%), or progression monitoring in postoperative patients.²⁰⁻²² Several commercially available software solutions have already received U.S. Food and Drug Administration (FDA) or Korean Ministry of Food and Drug Safety approval for landmark and curvature detection.^{23,24}

Vertebral segmentation and numbering within spine CT and MR images have become a well-established area of research, offering significant advancements in medical imaging analysis. This technology automatically identifies and delineates specific anatomical structures, most commonly the vertebral body, intervertebral disc, and spinal canal. Reportedly, these automated

algorithms achieve impressive accuracy, with Dice scores ranging from 89% to 95% for these key structures.²⁵ Several medical technology vendors have already received FDA approval for spine labeling software, and their incorporation into PACS (Picture Archiving and Communication Systems) is already underway.²³ The true power of automatic segmentation lies in its ability to facilitate quantitative analysis and automated diagnoses. For instance, automated vertebral segmentation enables (1) the detection of abnormal vertebral heights,²⁶ (2) the Identification of abnormal spinal curvatures,²⁷ and (3) the planning of surgical procedures and radiotherapy.²⁸ With disc segmentation, we can automatically grade disc degeneration in sagittal MRIs.²⁹ Furthermore, research with neural foramen segmentation has shown that the cross-sectional area of the neural foramen directly correlates with patient height and inversely correlates with age.³⁰ These segmentation models have opened doors for automatic diagnoses of spinal stenosis and neural foraminal stenosis.³¹⁻³³

3. Image Interpretation and Diagnosis

1) Disc herniation and degeneration

While disc and spinal canal segmentation models play a crucial role in identifying anatomical structures, their potential extends far beyond that. In MRI, AI models are making significant strides in detecting and grading lumbar spinal stenosis. These models are trained to identify stenosis in the lumbar central canal, lateral recess, and neural foramina using axial or sagittal images.^{32,34} One study reported remarkable agreement between a trained model and subspecialized radiologists in classifying stenosis severity (normal/mild vs. moderate/severe) in an external test set, achieving kappa values ranging from 0.95 to 0.96.³² AI-powered MRI analysis can significantly benefit radiologists by reducing interpretation time (124–274 seconds vs. 47–71 seconds, $p < 0.001$) and improving interobserver agreement (κ values = 0.71 and 0.70 with DL vs. 0.39 and 0.39 without DL, both $p < 0.001$).⁷ FDA-approved solutions are already emerging. For example, a lumbar spine report generation software boasts impressive sensitivity and specificity for detecting central canal stenosis, as demonstrated in 2 separate studies (92.70% and 99.04% or 77.14% and 98.95% for central canal stenosis).^{35,36}

Beyond MRI, AI-powered spinal stenosis diagnosis is also making progress in CT scans and radiographs. AI models have achieved diagnostic accuracies of 83%–88% for the spinal canal and 71%–75% for the lateral recess on axial CT scans.³⁷ This opens doors for evaluating disc herniation in CT scans or even diagnosing opportunistic disc herniation during abdomen and chest CT scans. For cervical spine radiographs, AI approaches

have demonstrated promising results in detecting ossification of the posterior longitudinal ligament (accuracy 0.88, area under the receiver operating characteristics curve, area under the curve [AUC] 0.94, surpassing the performance of spine physicians)³⁸ and spondylotic myelopathy (accuracy, 71.1%; AUC = 0.864).³⁹ Another study reports AUC values up to 90% for spinal stenosis in lumbar radiographs,⁴⁰ suggesting the potential of AI as a triage tool for further imaging.

2) Fracture

Compression fracture, the most common type of spinal fracture, has been one of the first targets for fracture detection in radiographs. Current DL models achieve impressive accuracy, reaching around 90% in compression fracture detections⁴¹⁻⁴³ and even in differentiating between old and new compression fractures (AUC = 0.80).⁴⁴ One model reports the performance was comparable to human readers (accuracy of 93%, $p < 0.001$) with lung markings as the primary source of false positives.⁴⁵ Further advancing the field, an FDA-approved universal fracture detector is already in widespread use. It efficiently detects fractures in radiographs of all limbs, spine, and ribs. For spine fracture specifically, a recent study compared human, AI-only, and AI-assisted human detection. The results revealed a clear advantage of AI-assisted human interpretation, with sensitivity/specificity of 94.5%/100% compared to 92.4%/98.4% for humans alone and 89.1%/62.2% for AI alone.⁴⁶

More recently, advancements in fracture detection have extended to CT scans. In 2022, the winners of the cervical spine fracture detection challenge⁴⁷ achieved an AUC of 0.96 (95% confidence interval [CI], 0.95–0.96), sensitivity of 88% (95% CI, 86%–90%), and specificity of 94% (95% CI, 93%–96%).⁴⁸ This same year also saw the introduction of an FDA-approved cervical spine fracture detector. Research further extends to thoracolumbar CT scans, with some studies demonstrating the ability to categorize fractures into 4 types (no injury, compression, burst, translational/rotational, and distraction) with accuracies ranging from 68.6% (burst) to 89.3% (distraction).⁴⁹

3) Inflammation and tumors

Identifying and classifying inflammatory diseases, infections, and tumors are actively researched fields within spine imaging, each too vast to explore fully here. However, AI models are already demonstrating impressive capabilities, achieving expert-level performance in several tasks. For example, in pelvic radiographs, DL models can diagnose sacroiliitis with accuracy that is on par with experts (Cohen kappa = 0.79).⁵⁰ Similarly, they

can accurately quantify inflammatory sacroiliitis in MRI^{51,52} and differentiate between tuberculous and pyogenic spondylitis with an AUC of 0.802 (compared to 0.729 for human experts, $p = 0.079$).⁵³

Spinal oncology has also seen exciting advancements in radiomics and DL. Algorithms can now detect metastatic lesions in CT scans (sensitivity, 75%–90%^{54,55}) and be applied to reduce interrater variability.⁵⁶ In MRI, DL models outperformed fourth-year residents in differentiating malignant vertebral fractures, achieving 90% sensitivity and 79% specificity.⁵⁷ Furthermore, ML and DL can discriminate between normal and pathologic bone marrow patterns in MRI^{58,59} and distinguish spine metastases from lung cancer versus other primary origins.⁶⁰

4. Surgical Planning and Intraoperative Use

CT aids surgical planning by detailing bone anatomy and pathology and guiding surgical approaches. Moreover, navigation systems using preoperative CT images improve screw placement accuracy. Synthetic CT scans are computer-generated images resembling conventional CT scans but are created using other imaging modalities like MR images through generative models or other DL algorithms.⁶¹ This technology has several advantages over CT scans, such as elimination of radiation exposure (radiation dose of an average lumbar spine CT scan: 3.5 mSv),² improved visualization of metal structures and the peripheral field of view, and high-resolution depiction of soft tissue structures like intraosseous hemangiomas.⁶² Studies comparing visualization of body structures, artifacts, and geometrical measurements between synthetic and traditional CT scans have found the synthetic versions to be non-inferior.⁶²

However, synthetic CT is still a relatively young field with limited clinical use. While some applications like presurgical and radiotherapy planning are emerging,^{61,63,64} caution is advised when using measurements from synthetic spine CT scans. Studies have shown inaccuracies in pedicle measurements performed in the axial plane, with relative errors reaching up to 34%.⁶⁵

Augmented reality (AR) and virtual reality (VR) are emerging technologies demonstrating promising benefits in various healthcare fields, including robotic surgery, laparoscopic surgery, and, notably, orthopedic surgery for the spine.⁶⁶ One of their main applications is as a navigation tool in the operating room. AR and VR systems use computer vision techniques to process preoperative or intraoperative images (radiographs, CT scans, or MRIs) and overlay relevant anatomical structures, potential screw trajectories, or ideal screw locations onto the surgical field, guiding surgeons with real-time visualization.⁶⁷

The integration of DL into AR and VR systems further enhances their capabilities, particularly in object and landmark detection within images. For example, DL has been used to track a specific vertebra of interest in fluoroscopic images with high accuracy (mean error of 2.27%)⁶⁸ and identify 7 anatomical landmarks on intraoperative lumbar spine CT scans with minimal error.⁶⁹ Additionally, DL has shown promise in improving robotic screw placement^{1,70} and identifying bone drill breakthroughs during surgery.⁷¹ Classification models powered by DL can even differentiate various types of pedicle screws⁷² and anterior cervical fusion systems⁷³ in radiographs.

Screw navigation systems are getting a boost from neural networks. These AI tools automate screw planning, including screw size and trajectory. One study reports a dramatic 90% reduction in workflow time, with just 3 out of 130 screws requiring manual adjustment.⁷⁴ Neural networks can also personalize screw placement for each patient's bone structure. This customization helps maximize pull-out force and reduce screw failure, which will be an essential benefit for patients with osteoporosis.⁷⁵

We anticipate that AI will play a crucial role in addressing some current limitations of AR and VR, such as low image resolution and steep learning curves. AI-powered AR and VR solutions hold potential, offering features like determining ideal spine alignment and implant size, compensating for motion, and even facilitating 3D printing.^{66,67}

5. Opportunistic Diagnosis

Opportunistic screening refers to usage of incidental information from existing medical images acquired for a different purpose.⁷⁶ For example, an abdominal radiographs and CT scans intended for other diagnoses may incidentally reveal compression fractures.^{77,78} One of the most promising areas for this approach lies in body composition imaging, particularly for the assessment of bone mineral density (BMD).

While dual-energy x-ray absorptiometry (DXA) remains the standard for BMD measurement, CT scans offer valuable insights, especially for individuals with obesity, severe degenerative spine disease, or postoperative spine conditions where DXA accuracy can be compromised. Quantitative CT using reference phantoms or dual-energy CT has been used for these cases, but their accessibility remains limited. However, advancements in ML are now enabling the extraction of reliable BMD information from various CT scans, including the abdomen, chest, and lumbar spine, transforming them into valuable opportunistic screening tools.⁷⁹⁻⁸¹

Opportunistic CT scans are also revealing valuable insights into sarcopenia, a condition linked to vertebral compression fractures and increased mortality across various diseases.^{82,83} Sarcopenia is strongly associated with the cross-sectional area and mean density of muscles like the psoas or abdominal wall in CT scans.⁷⁶ AI algorithms can now automatically segment muscles, subcutaneous and visceral fat, and vertebral bodies for sarcopenia and BMD measurement.^{84,85} Body composition measurements from opportunistic imaging could even potentially aid in the risk stratification of patients undergoing spinal surgery or predict future fractures.^{86,87}

6. Clinical Decision-Making, Prediction, and Prognostication

ML and DL algorithms are revolutionizing the way we predict patient outcomes in spine surgery. While still in their early stages of clinical implementation, these powerful tools hold immense promise for personalized care and improved decision-making.

While the traditional studies with well-designed large cohorts or linear regression models have achieved success,⁸⁸ ML algorithms can capture both linear and nonlinear relationships between diverse factors and outcomes, requiring less human intervention in model development.⁸⁹ This leads to superior predictive performance compared to traditional methods.^{90,91} Furthermore, ML enables clustering of patients with similar data patterns. This paves the way for prognostication and treatment optimization tailored to specific groups of patients.^{92,93}

However, it's important to note that successful outcome prediction using ML requires high-quality, well-structured datasets with minimal missing data and clear outcome measures.

1) Herniated disc disease surgery

Herniated disc disease is a common spine condition where ML has shown utility in surgical decision support and postoperative prognosis. One key area of focus involves predicting recurrent herniation after surgery. Several ML studies have identified high-risk patients based on patient demographics, clinical parameters, and pre- and postoperative pain scores.^{90,94} Significant factors reported to be associated with reherniation included pain scores, Oswestry Disability Index (ODI), PI-LL mismatch, body mass index, coronal angulation, duration of symptoms, and age.⁹⁴ Additionally, incorporating radiographic features such as facet orientation, herniation type, Modic changes, and disc calcification has further enhanced prediction accuracy for recurrent lumbar disc herniation.⁹⁵

The decision to proceed with surgery for herniated disc dis-

ease involves complex considerations. Models trained on patient demographics, questionnaire data, and MRI results exhibit promising potential in surgical triaging, predicting surgical referrals with AUCs ranging from 0.68 to 0.88.^{96,97} Similar models have also been used to predict improvements in quality of life after surgery (AUCs up to 0.78),⁹⁸ or after conservative therapy (100% accuracy for 12 scale ODI).⁹¹ Mourad et al.⁹⁹ built a surgical recommendation model based on clinical symptoms, MRI findings, and patient demographic factors. The root mean square error between model predictions and ground truth was 0.0964, with agreement being higher than agreement between individual doctors.

Moreover, ML models trained with demographic information, comorbidities, and preoperative/intraoperative findings have shown efficacy in predicting hospital length of stay or readmission after spine surgeries like anterior cervical discectomy and fusion,¹⁰⁰⁻¹⁰² or lumbar single-level laminectomy.¹⁰³ This could aid in determining the appropriate care setting (outpatient vs. inpatient) or optimize hospital resource allocations.

2) Spinal deformity

Adult spinal deformity surgery is complex and carries a high risk of complications. To improve outcomes and reduce these complications, researchers have developed algorithms to aid clinical decision-making.¹⁰⁴ For instance, a study by the International Spine Study Group used ML-based clustering and found 4 prognostic phenotypes.¹⁰⁵ This approach identified that younger, more resilient patients with good mental health were less likely to need repeat surgery compared to older, frail patients with poorer mental health. Another study by Scheer et al. employed decision tree ensembles to predict proximal junctional kyphosis and pseudoarthrosis with an AUC of 0.896¹⁰⁶ and 0.947.¹⁰⁷

3) Postoperative complication

ML models are demonstrating remarkable potential in predicting major postoperative complications following spine surgery, such as wound infections, thromboembolism, and mortality.^{108,109} These models leverage diverse demographic and clinical parameters, often exceeding the performance of traditional logistic regression models and established risk scores like the American Society of Anesthesiologists (ASA) physical status classification score.¹¹⁰

For instance, Hopkins et al. developed an ML model that predicts postoperative surgical site infection after spinal fusion with a median AUC of 0.787.¹¹¹ Their analysis revealed factors like

congestive heart failure, chronic pulmonary failure, hemiplegia/paraplegia, and multilevel fusion as the most influential variables, providing valuable insights for risk stratification. Additionally, another study identified 10 key predictors, including age, gender, ASA physical status classification grade, surgical approach, and preoperative laboratory values, demonstrating the broad range of factors that ML models can incorporate for comprehensive risk assessment.¹¹²

Hardware failure is another concern, especially for patients with known risk factors such as osteoporosis, long fixation length, and certain fixation end points.¹¹³ You only look once v5 (YO-LOv5), a type of convolutional neural network model specialized for both detection and classification, can help detect hard-

ware failure in postoperative radiographs (Fig. 3).¹¹⁴

DISCUSSION

1. Limitations and Challenges

While AI holds immense promise for spinal imaging, several limitations require careful consideration. The “black box” nature of certain models and the lack of interpretability in decision-making raise concerns for clinical adoption. Fortunately, researchers are developing techniques to make AI models more interpretable. These techniques, like Gradient-weighted class activation mapping or Shapley additive explanations, can create heatmaps or graphs that highlight the focus of the model when making predictions. This can help explain, for instance, how the model identified spinal stenosis¹¹⁵ or reherniation.⁹⁴ Even more recent advancements involve using LLMs to explain AI model predictions in a more comprehensive way, further improving interpretability.¹¹⁶

Another crucial hurdle is ensuring generalizability.¹¹⁷ Recent studies suggest external validation may not adequately assess true model performance. They recommend that researchers prioritize recurrent local validation across diverse datasets to ensure real-world applicability.¹¹⁸ The focus should shift from chasing perfect metrics to demonstrating tangible clinical value, such as reducing inter-reader variability or mitigating human errors.

Furthermore, the scarcity of high-quality labeled datasets poses a significant challenge. Compared to other radiology fields, large public benchmarks for spinal imaging remain inadequate. Existing public datasets (Table 1) tend to be small in size (< 1,000 images), fragmented across institutions, and heterogeneous in acquisition and populations, leading to overfitting and limited generalizability. While current models perform well for normal

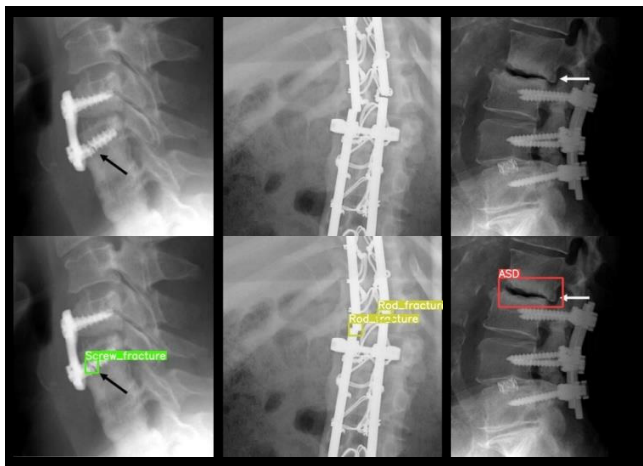


Fig. 3. Hardware failure predictions on spine radiographs. An off-the-shelf model was trained to predict 8 types of hardware failures: adjacent segment disease, loosening, migration, pseudoarthrosis, protrusion, rod fracture, screw fracture, and subsidence. The arrows were embedded in the original images to indicate screw fracture (black arrow) and adjacent segment disease (white arrow).

Table 1. Public spine radiology image datasets

Source	Modality	Region of interest	Segmentation/labels	Data size	Released date
SPIDER ¹³¹	MR	Lumbar	Segmentations of the vertebrae, disc, and spinal canal	447 Sagittal series (218 patients)	2023
Mendeley data ¹³²	MR	Lower Lumbar	Radiology report	48,345 Axial slices (515 patients)	2019
RSNA cervical fracture ⁴⁷	CT	Cervical	Label and segmentations of vertebrae and fractures	3,112 Patients	2023
CTSpine1K ¹³³	CT	Whole spine	Segmentation of vertebra	1,005 CT	2021
VerSe ¹³⁴	CT	Whole spine	Segmentations of the vertebrae	374 Sagittal series (355 patients)	2020
BUU' spine ¹³⁵	XR	Lumbar	Labels of 4 disease entities	400	2023
VinDr-SpineXR ¹³⁶	XR	Whole spine	Labels of 13 disease entities	10,466	2021

MR, magnetic resonance; CT, computed tomography; XR, x-ray.

spinal anatomy, there are very few models trained on datasets with normal variants,¹¹⁹ fractures,¹²⁰ or models predicting various disease entities.^{121,122}

2. Future Directions

Despite the challenges, the future of AI in spinal imaging is bubbling with potential. Researchers are increasingly leveraging prospectively gathered data from clinical trials and multi-centered datasets, yielding promising results.¹⁵ For instance, studies utilizing the American college of surgeons-national surgical quality improvement program database have successfully generated ML models for predicting 30-day readmissions^{111,123} or discharge to nonhome facility¹²⁴ after lumbar fusion using pre-discharge information.

Another exciting approach involves directly incorporating EMR or imaging data into DL models. Natural language processing shows promise in predicting intraoperative vascular injuries with superior accuracy compared to traditional methods.¹²⁵ Similarly, DL models trained directly on images can identify hidden features invisible to the human eye. For example, DL models utilizing preoperative sagittal MRIs of the cervical spine as inputs have demonstrated higher accuracy in predicting early onset adjacent segment disease in cervical fusion patients compared to models using only preoperative clinical data.^{126,127} Additionally, AI models trained on spinal radiographs and CT scans are outperforming established risk assessment tools in predicting fractures.^{128,129} Finally, encouraging results have been achieved in predicting the early progression of scoliosis using spinal radiographs.¹³⁰

CONCLUSION

In conclusion, this review highlights the remarkable progress and potential of AI in advancing spinal imaging and patient care. From automated measurements to surgical planning, AI is transforming workflow efficiency, accuracy, and reliability across the spectrum of spine imaging. However, thoughtfulness is required to ensure real-world validity, utility, and adoption of AI tools.

NOTES

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This work was supported by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare,

the Ministry of Food and Drug Safety) (Project Number: 17111 96789, RS-2023-00252804).

Author Contribution: Conceptualization: SL, JYJ, AM, JSK; Data curation: SL, JYJ; Formal analysis: SL, JYJ; Funding acquisition: JYJ; Methodology: SL, JYJ, AM, JSK; Project administration: JYJ, JSK; Visualization: SL, JYJ; Writing – original draft: SL, JYJ; Writing – review & editing: SL, JYJ, AM, JSK.

ORCID

Sungwon Lee: 0000-0002-1684-2996

Joon-Yong Jung: 0000-0002-6909-0919

Akaworn Mahatthanatrakul: 0000-0003-4216-8541

Jin-Sung Kim: 0000-0001-5086-0875

REFERENCES

1. Benzakour A, Altsitzioglou P, Lemée JM, et al. Artificial intelligence in spine surgery. *Int Orthop* 2023;47:457-65.
2. Jin M, Schröder M, Staartjes VE. Artificial intelligence and machine learning in spine surgery. In: Veeravagu A, Wang MY, editors. *Robotic and navigated spine surgery*. New Delhi (India): Elsevier; 2023. p. 213-29.
3. Hirschmann A, Cyriac J, Stieltjes B, et al. Artificial intelligence in musculoskeletal imaging: review of current literature, challenges, and trends. *Semin Musculoskelet Radiol* 2019;23:304-11.
4. Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: a primer for radiologists. *Radiographics* 2017;37:2113-31.
5. Wolterink JM, Mukhopadhyay A, Leiner T, et al. Generative adversarial networks: a primer for radiologists. *Radiographics* 2021;41:840-57.
6. Morbée L, Chen M, Herregods N, et al. MRI-based synthetic CT of the lumbar spine: geometric measurements for surgery planning in comparison with CT. *Eur J Radiol* 2021; 144:109999.
7. Lim DSW, Makmur A, Zhu L, et al. Improved productivity using deep learning–assisted reporting for lumbar spine MRI. *Radiology* 2022;305:160-6.
8. Subramanian T, Shahi P, Araghi K, et al. Using artificial intelligence to answer common patient-focused questions in minimally invasive spine surgery. *JBSJ* 2023;105:1649-53.
9. Chaudhari AS, Fang Z, Kogan F, et al. Super-resolution musculoskeletal MRI using deep learning. *Magn Reson Med* 2018;80:2139-54.
10. Yeoh H, Hong SH, Ahn C, et al. Deep learning algorithm for simultaneous noise reduction and edge sharpening in

- low-dose CT images: a pilot study using lumbar spine CT. *Korean J Radiol* 2021;22:1850-7.
11. Lin DJ, Johnson PM, Knoll F, et al. Artificial intelligence for MR image reconstruction: an overview for clinicians. *J Magn Reson Imaging* 2021;53:1015-28.
 12. Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med* 2018;79:3055-71.
 13. Almansour H, Herrmann J, Gassenmaier S, et al. Deep learning reconstruction for accelerated spine MRI: prospective analysis of interchangeability. *Radiology* 2023;306:e212922.
 14. Antun V, Renna F, Poon C, et al. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc Natl Acad Sci U S A* 2020;117:30088-95.
 15. Constant C, Aubin CE, Kremers HM, et al. The use of deep learning in medical imaging to improve spine care: a scoping review of current literature and clinical applications. *N Am Spine Soc J* 2023;15:100236.
 16. Ma J, He Y, Li F, et al. Segment anything in medical images. *Nat Commun* 2024;15:654.
 17. Galbusera F, Niemeyer F, Wilke HJ, et al. Fully automated radiological analysis of spinal disorders and deformities: a deep learning approach. *Eur Spine J* 2019;28:951-60.
 18. Yeh YC, Weng CH, Huang YJ, et al. Deep learning approach for automatic landmark detection and alignment analysis in whole-spine lateral radiographs. *Scie Rep* 2021;11:7618.
 19. Song SY, Seo MS, Kim CW, et al. AI-driven segmentation and automated analysis of the whole sagittal spine from x-ray images for spinopelvic parameter evaluation. *Bioengineering (Basel)* 2023;10:1229.
 20. Loder RT, Urquhart A, Steen Hu, et al. Variability in Cobb angle measurements in children with congenital scoliosis. *J Bone Joint Surg Br* 1995;77:768-70.
 21. Suri A, Tang S, Kargilis D, et al. Conquering the Cobb angle: a deep learning algorithm for automated, hardware-invariant measurement of Cobb angle on radiographs in patients with scoliosis. *Radiol Artif Intell* 2023;5:e220158.
 22. 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.
 23. U.S. Food and Drug Administration. Artificial intelligence and machine learning (AI/ML)-enabled medical devices. Silver Spring (MD): U.S. Food and Drug Administration; 2024 [cited 2024 Jan 11]. Available from: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>.
 24. Ministry of Food and Drug Safety. Medical device and quality control standards [Internet]. Cheongju (Korea): Ministry of Food and Drug Safety; 2024 [cited 2024 Jan 11]. Available from: <https://emedi.mfds.go.kr/msismext/emd/ifm/gmpSrngStatusView.do>.
 25. D'Antoni F, Russo F, Ambrosio L, et al. Artificial intelligence and computer vision in low back pain: a systematic review. *Int J Environ Res Public Health* 2021;18:10909.
 26. Suri A, Jones BC, Ng G, et al. Vertebral deformity measurements at MRI, CT, and radiography using deep learning. *Radiol Artif Intell* 2021;4:e210015.
 27. Alukaev D, Kiselev S, Mustafaev T, et al. A deep learning framework for vertebral morphometry and Cobb angle measurement with external validation. *Eur Spine J* 2022;31:2115-24.
 28. Diniz JOB, Ferreira JL, Diniz PHB, et al. A deep learning method with residual blocks for automatic spinal cord segmentation in planning CT. *Biomed Signal Process Control* 2022;71(Pt A):103074.
 29. Zheng HD, Sun YL, Kong DW, et al. Deep learning-based high-accuracy quantitation for lumbar intervertebral disc degeneration from MRI. *Nat Commun* 2022;13:841.
 30. Gaonkar B, Beckett J, Villaroman D, et al. Quantitative analysis of neural foramina in the lumbar spine: an imaging informatics and machine learning study. *Radiol Artif Intell* 2019;1:180037.
 31. Al-Kafri AS, Sudirman S, Hussain A, et al. Boundary delineation of MRI images for lumbar spinal stenosis detection through semantic segmentation using deep neural networks. *IEEE Access* 2019;7:43487-501.
 32. Hallinan JTPD, Zhu L, Yang K, et al. Deep learning model for automated detection and classification of central canal, lateral recess, and neural foraminal stenosis at lumbar spine MRI. *Radiology* 2021;300:130-8.
 33. Tumko V, Kim J, Uspenskaia N, et al. A neural network model for detection and classification of lumbar spinal stenosis on MRI. *Eur Spine J* 2024;33:941-8.
 34. 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;33:3435-43.
 35. Georgiev R, Novakova M, Bliznakova K. Clinical assessment of CoLumbo deep learning system for central canal stenosis diagnostics. *EJMO* 2023;7:42-8.

36. Lehnen NC, Haase R, Faber J, et al. Detection of degenerative changes on MR images of the lumbar spine with a convolutional neural network: a feasibility study. *Diagnostics (Basel)* 2021;11:902.
37. Li KY, Weng JJ, Li HL, et al. Development of a deep learning model for diagnosing lumbar spinal stenosis based on CT images. *Spine (Phila Pa 1976)* 2023 Dec 19. doi: 10.1097/BRS.0000000000004903. [Epub].
38. Tamai K, Terai H, Hoshino M, et al. A deep learning algorithm to identify cervical ossification of posterior longitudinal ligaments on radiography. *Sci Rep* 2022;12:2113.
39. Lee GW, Shin H, Chang MC. Deep learning algorithm to evaluate cervical spondylotic myelopathy using lateral cervical spine radiograph. *BMC Neurol* 2022;22:147.
40. Kim T, Kim YG, Park S, et al. Diagnostic triage in patients with central lumbar spinal stenosis using a deep learning system of radiographs. *J Neurosurg Spine* 2022;37:104-11.
41. Murata K, Endo K, Aihara T, et al. Artificial intelligence for the detection of vertebral fractures on plain spinal radiography. *Sci Rep* 2020;10:20031.
42. Kim DH, Jeong JG, Kim YJ, et al. Automated vertebral segmentation and measurement of vertebral compression ratio based on deep learning in X-ray images. *J Digit Imaging* 2021;34:853-61.
43. Kim KC, Cho HC, Jang TJ, et al. Automatic detection and segmentation of lumbar vertebrae from X-ray images for compression fracture evaluation. *Comput Methods Programs Biomed* 2021;200:105833.
44. Chen W, Liu X, Li K, et al. A deep-learning model for identifying fresh vertebral compression fractures on digital radiography. *Eur Radiol* 2022;32:1496-505.
45. Li YC, Chen HH, Horng-Shing Lu H, et al. Can a deep-learning model for the automated detection of vertebral fractures approach the performance level of human subspecialists? *Clin Orthop Relat Res* 2021;479:1598-612.
46. Oppenheimer J, Lücken S, Hamm B, et al. A prospective approach to integration of AI fracture detection software in radiographs into clinical workflow. *Life (Basel)* 2023; 13:223.
47. Lin HM, Colak E, Richards T, et al. The RSNA Cervical Spine Fracture CT Dataset. *Radiol Artif Intell* 2023;5:e230034.
48. Lee GR, Flanders AE, Richards T, et al. Performance of the winning algorithms of the RSNA 2022 cervical spine fracture detection challenge. *Radiol Artif Intell* 2024;6:e230256.
49. 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.
50. Bressemer KK, Vahldiek JL, Adams L, et al. Deep learning for detection of radiographic sacroiliitis: achieving expert-level performance. *Arthritis Res Ther* 2021;23:106.
51. Lin KYY, Peng C, Lee KH, et al. Deep learning algorithms for magnetic resonance imaging of inflammatory sacroiliitis in axial spondyloarthritis. *Rheumatology* 2022;61:4198-206.
52. Bressemer KK, Adams LC, Proft F, et al. Deep learning detects changes indicative of axial spondyloarthritis at MRI of sacroiliac joints. *Radiology* 2022;305:655-65.
53. Kim K, Kim S, Lee YH, et al. Performance of the deep convolutional neural network based magnetic resonance image scoring algorithm for differentiating between tuberculous and pyogenic spondylitis. *Sci Rep* 2018;8:13124.
54. Gilberg L, Teodorescu B, Maerkisch L, et al. Deep learning enhances radiologists' detection of potential spinal malignancies in CT scans. *Appl Sci* 2023;13:8140.
55. Li Y, Zhang Y, Zhang E, et al. Differential diagnosis of benign and malignant vertebral fracture on CT using deep learning. *Eur Radiol* 2021;31:9612-9.
56. Hallinan JTPD, Zhu L, Zhang W, et al. Deep learning assessment compared to radiologist reporting for metastatic spinal cord compression on CT. *Front Oncol* 2023;13: 1151073.
57. Yeh LR, Zhang Y, Chen JH, et al. A deep learning-based method for the diagnosis of vertebral fractures on spine MRI: retrospective training and validation of ResNet. *Eur Spine J* 2022;31:2022-30.
58. Hwang EJ, Jung JY, Lee SK, et al. Machine learning for diagnosis of hematologic diseases in magnetic resonance imaging of lumbar spines. *Sci Rep* 2019;9:6046.
59. Hwang EJ, Kim S, Jung JY. Bone marrow radiomics of T1-weighted lumbar spinal MRI to identify diffuse hematologic marrow diseases: comparison with human readings. *IEEE Access* 2020;8:133321-9.
60. Lang N, Zhang Y, Zhang E, et al. Differentiation of spinal metastases originated from lung and other cancers using radiomics and deep learning based on DCE-MRI. *Magn Reson Imaging* 2019;64:4-12.
61. Staartjes VE, Seevinck PR, Vandertop WP, et al. Magnetic resonance imaging-based synthetic computed tomography of the lumbar spine for surgical planning: a clinical proof-of-concept. *Neurosurg Focus* 2021;50:E13.
62. van der Kolk B, Slotman DJ, Nijholt IM, et al. Bone vi-

- sualization of the cervical spine with deep learning-based synthetic CT compared to conventional CT: a single-center noninferiority study on image quality. *Eur J Radiol* 2022; 154:110414.
63. Spadea MF, Maspero M, Zaffino P, et al. Deep learning based synthetic-CT generation in radiotherapy and PET: a review. *Med Phys* 2021;48:6537-66.
64. Minnema J, Ernst A, van Eijnatten M, et al. A review on the application of deep learning for CT reconstruction, bone segmentation and surgical planning in oral and maxillofacial surgery. *Dentomaxillofac Radiol* 2022;51:20210437.
65. Roberts M, Hinton G, Wells AJ, et al. Imaging evaluation of a proposed 3D generative model for MRI to CT translation in the lumbar spine. *Spine J* 2023;23:1602-12.
66. Ghaednia H, Fourman MS, Lans A, et al. Augmented and virtual reality in spine surgery, current applications and future potentials. *Spine J* 2021;21:1617-25.
67. Yamout T, Orosz LD, Good CR, et al. Technological advances in spine surgery: navigation, robotics, and augmented reality. *Orthop Clin* 2023;54:237-46.
68. Liu Y, Sui X, Liu C, et al. Automatic lumbar spine tracking based on siamese convolutional network. *J Digit Imaging* 2020;33:423-30.
69. Siemionow K, Luciano C, Forsthoefel C, et al. Autonomous image segmentation and identification of anatomical landmarks from lumbar spine intraoperative computed tomography scans using machine learning: a validation study. *J Craniovertebr Junction Spine* 2020;11:99-103.
70. 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.
71. Seibold M, Maurer S, Hoch A, et al. Real-time acoustic sensing and artificial intelligence for error prevention in orthopedic surgery. *Sci Rep* 2021;11:3993.
72. Yang HS, Kim KR, Kim S, et al. Deep learning application in spinal implant identification. *Spine (Phila Pa 1976)* 2021; 46:E318-24.
73. Huang KT, Silva MA, See AP, et al. A computer vision approach to identifying the manufacturer and model of anterior cervical spinal hardware. *J Neurosurg Spine* 2019;31: 844-50.
74. 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;22:1666-76.
75. Ma C, Zou D, Qi H, et al. A novel surgical planning system using an AI model to optimize planning of pedicle screw trajectories with highest bone mineral density and strongest pull-out force. *Neurosurg Focus* 2022;52:E10.
76. Boutin RD, Lenchik L. Value-added opportunistic CT: insights into osteoporosis and sarcopenia. *AJR Am J Roentgenol* 2020;215:582-94.
77. Burns JE, Yao J, Summers RM. Vertebral body compression fractures and bone density: automated detection and classification on CT images. *Radiology* 2017;284:788-97.
78. Chen HY, Hsu BW, Yin YK, et al. Application of deep learning algorithm to detect and visualize vertebral fractures on plain frontal radiographs. *PLoS One* 2021;16:e0245992.
79. Yasaka K, Akai H, Kunimatsu A, et al. Prediction of bone mineral density from computed tomography: application of deep learning with a convolutional neural network. *Eur Radiol* 2020;30:3549-57.
80. Nam KH, Seo I, Kim DH, et al. Machine learning model to predict osteoporotic spine with hounsfield units on lumbar computed tomography. *J Korean Neurosurg Soc* 2019;62: 442-9.
81. Pan Y, Zhao F, Cheng G, et al. Automated vertebral bone mineral density measurement with phantomless internal calibration in chest LDCT scans using deep learning. *Br J Radiol* 2023;96:20230047.
82. Anand A, Shetty AP, Renjith K, et al. Does sarcopenia increase the risk for fresh vertebral fragility fractures?: a case-control study. *Asian Spine J* 2020;14:17-24.
83. Beaudart C, Zaaria M, Pasleau F, et al. Health outcomes of sarcopenia: a systematic review and meta-analysis. *PLoS one* 2017;12:e0169548.
84. Ha J, Park T, Kim HK, et al. Development of a fully automatic deep learning system for L3 selection and body composition assessment on computed tomography. *Sci Rep* 2021;11:21656.
85. Nowak S, Theis M, Wichtmann BD, et al. End-to-end automated body composition analyses with integrated quality control for opportunistic assessment of sarcopenia in CT. *Eur Radiol* 2022;32:3142-51.
86. Moskven E, Bourassa-Moreau E, Charest-Morin R, et al. The impact of frailty and sarcopenia on postoperative outcomes in adult spine surgery. A systematic review of the literature. *Spine J* 2018;18:2354-69.
87. Pickhardt PJ, Perez AA, Garrett JW, et al. Fully automated deep learning tool for sarcopenia assessment on CT: L1

- versus L3 vertebral level muscle measurements for opportunistic prediction of adverse clinical outcomes. *AJR Am J Roentgenol* 2022;218:124-31.
88. Bernstein DN, Keswani A, Chi D, et al. Development and validation of risk-adjustment models for elective, single-level posterior lumbar spinal fusions. *J Spine Surg* 2019;5: 46-57.
 89. Chang M, Canseco JA, Nicholson KJ, et al. The role of machine learning in spine surgery: the future is now. *Front Surg* 2020;7:54.
 90. Azimi P, Mohammadi HR, Benzel EC, et al. Use of artificial neural networks to predict recurrent lumbar disk herniation. *Clin Spine Surg* 2015;28:E161-5.
 91. Wirries A, Geiger F, Hammad A, et al. Artificial intelligence facilitates decision-making in the treatment of lumbar disc herniations. *Eur Spine J* 2021;30:2176-84.
 92. Ames CP, Smith JS, Pellisé F, et al. Artificial intelligence based hierarchical clustering of patient types and intervention categories in adult spinal deformity surgery: towards a new classification scheme that predicts quality and value. *Spine (Phila Pa 1976)* 2019;44:915-26.
 93. Tee JW, Rivers CS, Fallah N, et al. Decision tree analysis to better control treatment effects in spinal cord injury clinical research. *J Neurosurg Spine* 2019;31:464-72.
 94. Harada GK, Siyaji ZK, Mallow GM, et al. Artificial intelligence predicts disk re-herniation following lumbar microdiscectomy: development of the “RAD” risk profile. *Eur Spine J* 2021;30:2167-75.
 95. Ren G, Liu L, Zhang P, et al. Machine learning predicts recurrent lumbar disc herniation following percutaneous endoscopic lumbar discectomy. *Global Spine J* 2024;14: 146-52.
 96. Wilson B, Gaonkar B, Yoo B, et al. Predicting spinal surgery candidacy from imaging data using machine learning. *Neurosurgery* 2021;89:116-21.
 97. 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.
 98. Khan O, Badhiwala JH, Witiw CD, et al. Machine learning algorithms for prediction of health-related quality-of-life after surgery for mild degenerative cervical myelopathy. *Spine J* 2021;21:1659-69.
 99. Mourad R, Kolisnyk S, Baiun Y, et al. Performance of hybrid artificial intelligence in determining candidacy for lumbar stenosis surgery. *Eur Spine J* 2022;31:2149-55.
 100. 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: 758-66.
 101. 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;91:322-30.
 102. Shah AA, Devana SK, Lee C, et al. Development and external validation of a risk calculator for prediction of major complications and readmission after anterior cervical discectomy and fusion. *Spine (Phila Pa 1976)* 2023;48:460-7.
 103. Li Q, Zhong H, Girardi FP, et al. Machine learning approaches to define candidates for ambulatory single level laminectomy surgery. *Global Spine J* 2022;12:1363-8.
 104. Joshi RS, Haddad AF, Lau D, et al. Artificial intelligence for adult spinal deformity. *Neurospine* 2019;16:686-94.
 105. 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;24:1095-108.
 106. Scheer JK, Osorio JA, Smith JS, et al. Development of validated computer-based preoperative predictive model for proximal junction failure (PJF) or clinically significant PJK with 86% accuracy based on 510 ASD patients with 2-year follow-up. *Spine (Phila Pa 1976)* 2016;41:E1328-35.
 107. Scheer JK, Oh T, Smith JS, et al. Development of a validated computer-based preoperative predictive model for pseudarthrosis with 91% accuracy in 336 adult spinal deformity patients. *Neurosurg Focus* 2018;45:E11.
 108. Stephens ME, O’Neal CM, Westrup AM, et al. Utility of machine learning algorithms in degenerative cervical and lumbar spine disease: a systematic review. *Neurosurg Rev* 2022;45:965-78.
 109. Kim JS, Merrill RK, Arvind V, et al. Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion. *Spine (Phila Pa 1976)* 2018;43:853-60.
 110. Arvind V, Kim JS, Oermann EK, et al. Predicting surgical complications in adult patients undergoing anterior cervical discectomy and fusion using machine learning. *Neurospine* 2018;15:329-37.
 111. Hopkins BS, Yamaguchi JT, Garcia R, et al. Using machine learning to predict 30-day readmissions after posterior lumbar fusion: an NSQIP study involving 23,264 patients. *J Neurosurg Spine* 2019;32:399-406.

112. Fatima N, Zheng H, Massaad E, et al. Development and validation of machine learning algorithms for predicting adverse events after surgery for lumbar degenerative spondylolisthesis. *World Neurosurgery* 2020;140:627-41.
113. Chen Z, Lv G, Zhang O, et al. Risk factors of instrumentation failure after laminectomy and posterior cervical fusions (PCF). *BMC Musculoskelet Disord* 2024;25:1.
114. Lee JW, Chun KS, Lee S, et al. Hardware failure detection in spine radiography using AI. Paper presented at Radiological Society of North America; 2023 Nov 29; Chicago, USA.
115. Park S, Kim JH, Ahn Y, et al. Multi-pose-based convolutional neural network model for diagnosis of patients with central lumbar spinal stenosis. *Sci Rep* 2024;14:203.
116. Bhattacharjee A, Moraffah R, Garland J, et al. LLMs as counterfactual explanation modules: can ChatGPT explain black-box text classifiers? *arXiv:2309.13340v2* [Preprint]. 2024 [2024 Jan 11]. Available from: <https://doi.org/10.48550/arXiv.2309.13340>.
117. Voter A, Larson M, Garrett J, et al. 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.
118. Youssef A, Pencina M, Thakur A, et al. External validation of AI models in health should be replaced with recurring local validation. *Nat Med* 2023;29:2686-7.
119. 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.
120. Jakubicek R, Chmelik J, Jan J, et al. Learning-based vertebra localization and labeling in 3D CT data of possibly incomplete and pathological spines. *Comput Methods Programs Biomed* 2020;183:105081.
121. Zerunian M, Pucciarelli F, Caruso D, et al. Fast high-quality MRI protocol of the lumbar spine with deep learning-based algorithm: an image quality and scanning time comparison with standard protocol. *Skeletal Radiol* 2024;53:151-9.
122. Haim O, Agur A, Gabay S, et al. Differentiating spinal pathologies by deep learning approach. *Spine J* 2024;24:297-303.
123. Kalagara S, Eltorai AE, Durand WM, et al. Machine learning modeling for predicting hospital readmission following lumbar laminectomy. *J Neurosurg Spine* 2018;30:344-52.
124. Goyal A, Ngufor C, Kerezoudis P, et al. Can machine learning algorithms accurately predict discharge to nonhome facility and early unplanned readmissions following spinal fusion? Analysis of a national surgical registry. *J Neurosurg Spine* 2019;31:568-78.
125. Karhade AV, Bongers ME, 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.
126. Rudisill SS, Hornung AL, Barajas JN, et al. Artificial intelligence in predicting early-onset adjacent segment degeneration following anterior cervical discectomy and fusion. *Eur Spine J* 2022;31:2104-14.
127. Goedmakers CM, Lak AM, Duey AH, et al. Deep learning for adjacent segment disease at preoperative MRI for cervical radiculopathy. *Radiology* 2021;301:664-71.
128. Kong SH, Lee JW, Bae BU, et al. Development of a spine X-ray-based fracture prediction model using a deep learning algorithm. *Endocrinol Metab (Seoul)* 2022;37:674-83.
129. Hu X, Zhu Y, Qian Y, et al. Prediction of subsequent osteoporotic vertebral compression fracture on CT radiography via deep learning. *View* 2022;3:20220012.
130. Wang H, Zhang T, Cheung KM, et al. Application of deep learning upon spinal radiographs to predict progression in adolescent idiopathic scoliosis at first clinic visit. *EClinicalMedicine* 2021;42:101220.
131. van der Graaf JW, van Hooft ML, Buckens CFM, et al. Lumbar spine segmentation in MR images: a dataset and a public benchmark. *Sci Data* 2024;11:264.
132. Sudirman S, Al Kafri A, Natalia F, et al. Lumbar spine MRI dataset. *Mendeley Data* 2019;2:2019.
133. Deng Y, Wang C, Hui Y, et al. Ctspine1k: a large-scale dataset for spinal vertebrae segmentation in computed tomography. *arXiv:2105.14711v3* [Preprint]. 2021 [cited 2014 Jan 11]. Available from: <https://doi.org/10.48550/arXiv.2105.14711>.
134. Sekuboyina A, Hussein ME, Bayat A, et al. VerSe: a vertebrae labelling and segmentation benchmark for multi-detector CT images. *Med Image Anal* 2021;73:102166.
135. Klinwicht P, Yookwan W, Limchareon S, et al. BUU-LSPINE: a thai open lumbar spine dataset for spondylolisthesis detection. *Appl Sci* 2023;13:8646.
136. Nguyen HT, Pham HH, Nguyen HT, et al. VinDr-SpineXR: a large annotated medical image dataset for spinal lesions detection and classification from radiographs. *arXiv:2106.12930v1* [Preprint]. 2021 [cited 2014 Jan 11]. Available from: <https://doi.org/10.48550/arXiv.2106.12930>.