Asian Spine Journal
Review Article
Asian Spine J 2020;14(4):543-571 • https://doi.org/10.31616/asj.2020.0147

A Review on the Use of Artificial Intelligence in Spinal Diseases

Parisa Azimi[1], Taravat Yazdanian[2], Edward C. Benzel[3], Hossein Nayeb Aghaei[1], Shirzad Azhari[1], Sohrab Sadeghi[1], Ali Montazeri[4]

[1]Department of Neurosurgery, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[2]School of Medicine, Capital Medical University, Beijing, China
[3]Department of Neurosurgery, Cleveland Clinic Foundation, Cleveland, OH, USA
[4]Mental Health Research Group, Health Metrics Research Centre, Iranian Institute for Health Sciences Research, ACECR, Tehran, Iran

Artificial neural networks (ANNs) have been used in a wide variety of real-world applications and it emerges as a promising field across various branches of medicine. This review aims to identify the role of ANNs in spinal diseases. Literature were searched from electronic databases of Scopus and Medline from 1993 to 2020 with English publications reported on the application of ANNs in spinal diseases. The search strategy was set as the combinations of the following keywords: "artificial neural networks," "spine," "back pain," "prognosis," "grading," "classification," "prediction," "segmentation," "biomechanics," "deep learning," and "imaging." The main findings of the included studies were summarized, with an emphasis on the recent advances in spinal diseases and its application in the diagnostic and prognostic procedures. According to the search strategy, a set of 3,653 articles were retrieved from Medline and Scopus databases. After careful evaluation of the abstracts, the full texts of 89 eligible papers were further examined, of which 79 articles satisfied the inclusion criteria of this review. Our review indicates several applications of ANNs in the management of spinal diseases including (1) diagnosis and assessment of spinal disease progression in the patients with low back pain, perioperative complications, and readmission rate following spine surgery; (2) enhancement of the clinically relevant information extracted from radiographic images to predict Pfirrmann grades, Modic changes, and spinal stenosis grades on magnetic resonance images automatically; (3) prediction of outcomes in lumbar spinal stenosis, lumbar disc herniation and patient-reported outcomes in lumbar fusion surgery, and preoperative planning and intraoperative assistance; and (4) its application in the biomechanical assessment of spinal diseases. The evidence suggests that ANNs can be successfully used for optimizing the diagnosis, prognosis and outcome prediction in spinal diseases. Therefore, incorporation of ANNs into spine clinical practice may improve clinical decision making.

Keywords: Spine; Review; Artificial neural networks

Introduction

Artificial neural network (ANN) models represent a mathematical rendition of the human nervous system that have been broadly applied to solve various nonlinear problems in the biomedical arena [1,2]. ANN is a machine-learning technique adept at learning the relationships between specified input and output variables.

Received Apr 3, 2020; Revised Apr 10, 2020; Accepted Apr 12, 2020
Corresponding author: Parisa Azimi
Neuroscience Research Center, Shahid Beheshti University of Medical Sciences, Arabi Ave, Daneshjoo Blvd, Velenjak, Tehran 19839-63113, Iran
Tel: +98-21-2243-9770, Fax: +98-21-2243-3455, E-mail: parisa.azimi@gmail.com

Asian Spine Journal • pISSN 1976-1902 eISSN 1976-7846 • www.asianspinejournal.org

Neural networks have been used predominantly for pattern-recognition regarding prediction and classification. The history and theory of ANNs has been reported in detail elsewhere [1-4]. In addition, the advantages and disadvantages of ANN have also been previously reported by us [1,2]. ANN is a promising field with numerous applications across various branches of medicine wherein it serves as a decision support tool to provide economic solutions to time and resource management [5]. Recently, artificial intelligence and related algorithms have facilitated rapid advances in the assessment of spinal diseases [2,5]. Moreover, ANNs are applied for clinical diagnosis, prognosis, outcome prediction following spinal surgery, research, and biomechanical assessments of spinal diseases [2]. However, there has been little utilization of ANNs in spine clinical practice. Given the recent advances in the management of spinal diseases and the fundamental role of decision, this comprehensive review is conducted aiming to describe the ANN-aided decision support system for management of spinal diseases, including diagnosis, prognosis, and outcome prediction.

Methods

ANN based methodology has been reported in detail elsewhere [1,2].

1. Search strategy

A detailed search of original articles was performed on Medline (through the PubMed search engine) and Scopus (Elsevier) databases to identify the applications of ANNs diagnosis, prediction, and prognosis of spinal disease. The review is intended to include all the full-text publications in the English biomedical journals. The following combinations of keywords were searched within the titles and abstracts: "artificial neural networks," "spine," "back pain," "prognosis," "grading," "classification," "prediction," "segmentation," "biomechanics," "deep learning," and "imaging." The structural keywords were selected due to their likelihood of being mentioned in either the title or the abstract of relevant articles. Since the first study of ANNs in spine diseases published in 1993, we performed a comprehensive search covering the period 1993 to 2020. An initial search was carried out in November 2019 and updated thrice in 2020 (January, February, and March).

2. Inclusion and exclusion criteria

All research articles on ANNs in spinal diseases were screened in the Scopus and Medline databases. Each article was independently reviewed by two reviewers and disagreements were sent to each other for resolution, only the articles emphasize on the most recent advances and their application in the spinal diseases were included. Publications on other disease conditions or animal studies were excluded.

3. Data synthesis

The findings from the all identified studies were summarized in a descriptive table, including authors' names, publication year, study setting, study sample, disease conditions (if relevant data is available), and main results or conclusions. Subsequently, the findings were sorted chronologically.

Results

1. Statistics

The reviewers identified and screened 3,653 unique abstracts. After screening, 3,564 papers were found to be irrelevant. Then, the remaining 89 papers were examined and the full text were reviewed for eligibility criteria. Ultimately, we included 79 studies on qualitative analysis. The flowchart of the literature review process is illustrated in Fig. 1. Overall, we pursued four categories of studies, namely diagnosis, progression, outcome prediction, and use of ANNs in the biomechanical assessments of spinal diseases. The main findings were grouped and presented as follows [6-81].

2. Diagnosis

In spinal diseases, ANNs have been successfully tested for diagnosis of pediatric low back pain [6,9,11], normal and abnormal cervical spine vertebra [8], scoliosis spinal deformity [7,10], and identification of risk factors associated with the development of complications following posterior lumbar spine fusion [23]. Besides, artificial intelligence models have been employed for medical image analysis assessment, such as those portrayed in the Table 1.

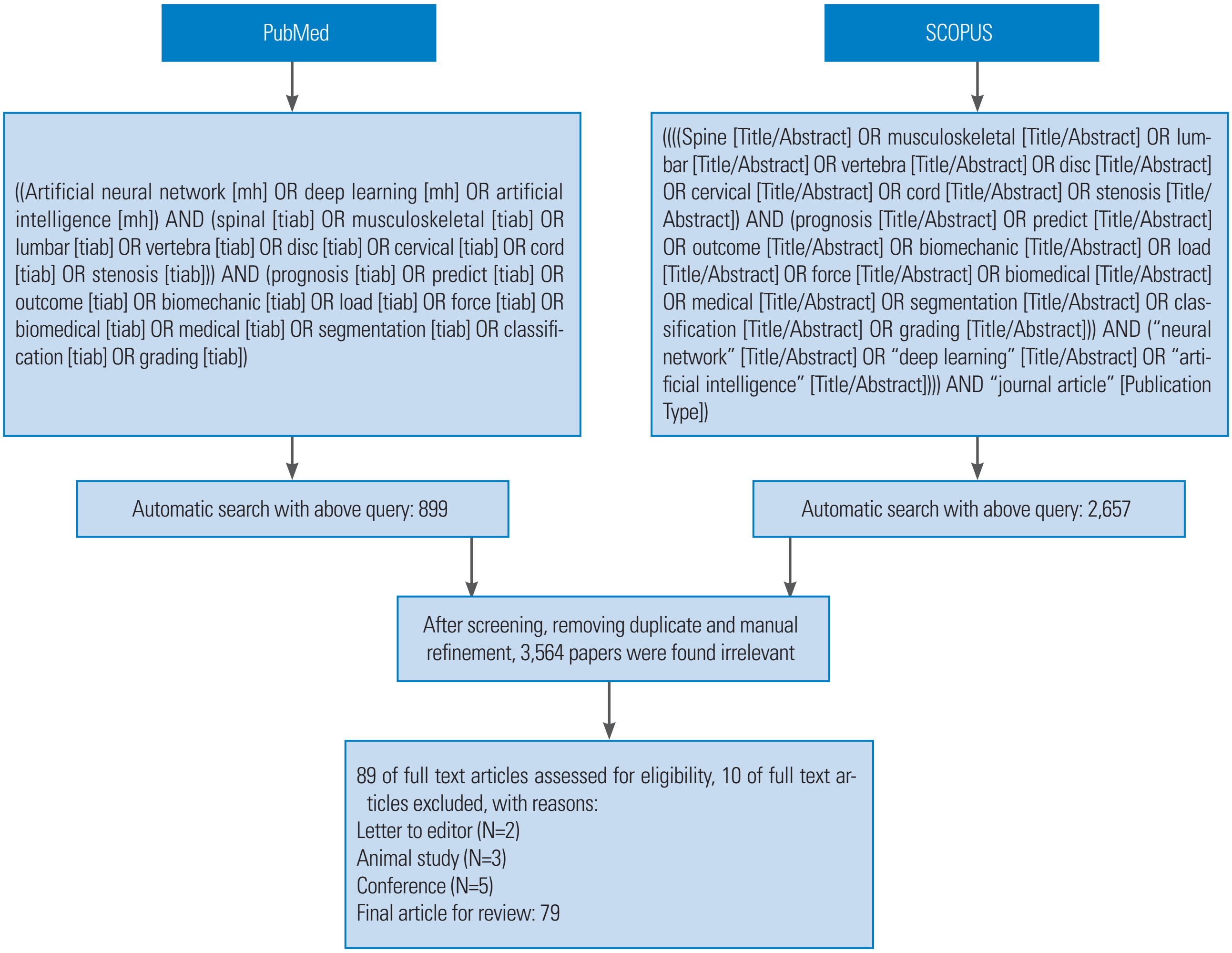

Fig. 1. Flowchart of literature search, selection, and identification.

3. Spinal prognosis

In addition to the application of in identifying the patients with high risk of hypotension during spinal anesthesia [52], ANNS have been tested to determine the prognosis of low back pain [51,53] aiming to automatically predict (and identify risk factors for) the complications following posterior lumbar spine fusion surgery [16], and to develop and evaluate a set of predictive models for common adverse events after spine surgery [57]. Also, ANNs are useful for developing novel computational tools to predict clinical outcomes, return to work, physical disability, occurrence of complications, readmission rates, walking ability, discharge, and disposition following spine surgery [54,55,58]. Neural network techniques have also been applied to develop predictive algorithms for postoperative complications following anterior cervical discectomy and fusion [56], and to evaluate clinically relevant improvement in leg pain, back pain and functional disability after lumbar disc herniation (LDH) surgery [59], and to automatically quantify muscle fat infiltration following whiplash injury [62]. In addition, ANNs have been shown to accurately predict survival, discharge and hospital readmission rates following spinal metastasis surgery [57,60,61], to predict discharge to rehabilitation and unplanned readmissions in patients receiving spinal fusion [63], and to predict prolonged opioid prescription after surgery for LDH [64]. Last but not least, ANNs have been used to predict the survival rate following a spinopelvic chondrosarcoma diagnosis [65] and to predict the occurrence of four types of major complications, namely cardiac complications, wound complications, venous thromboembolism, and mortality in the patients undergoing spine fusion, and it has achieved better results than

Table 1. A list of papers on ANN used in spine diagnosis

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Bishop et al. [6]	1997	USA	161	22	MLP: resilient propagation neural networks, and radial basis function neural networks	LBP	Yes	To determine specific characteristics of trunk motion associated with different categories of spinal disorders and to determine whether an ANNs can be effective in distinguishing patterns.	The neural network classifier produced the best results with up to 85% accuracy on "validation" data.
Jaremko et al. [7]	2001	Canada	49	18	MLP: a three-layer back-propagation artificial neural network using the Levenberg-Marquardt algorithm	Spinal deformity	NR	To assess whether full-torso surface laser scan images can be effectively used to estimate spinal deformity with the aid of an ANNs.	The ANNs estimated the maximal Cobb angle within 6° in 63% of the test data. set and was able to distinguish a Cobb angle greater than 30° with a sensitivity of 1.0 and specificity of 0.75. ANNs of full-torso scan imaging showed promise to accurately estimate scoliotic spinal deformity in a variety of patients.
Stanley et al. [8]	2001	USA	118	118	MLP	Cervical spine vertebra	Yes	Comparing various classifiers including an ANNs, K-Means algorithm, quadratic discriminant classifier and LVQ3.	Results from those classifiers are reported for recognizing vertebrae as normal or abnormal.
Liszka-Hackzell et al. [9]	2002	Sweden	30	10	MLP	LBP	NR	To explore new techniques of patient assessment that may prospectively identify of patients experience extended chronic pain and disability at risk of developing poor outcomes.	There was a good correlation between the true and predicted values for general health (r=0.96, p<0.01) and mental health (r=0.80, p<0.01). ANNs can be applied effectively to categorizing patients with acute and chronic LBP.
Lin et al. [10]	2008	USA	25 Patterns	12 Patterns	MLP: a multilayer feed-forward, back-propagation ANN	Spinal deformity	NR	To identify the classification of unspecified patterns of the scoliosis spine models	The accuracy was within 2.0 mm. The study showed that the data do not seem to be adequate enough due to participate study were small. Nevertheless, ANNs was useful with high accuracy to identify the classification patterns of the scoliosis spinal deformity.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size: Training	No. of sample size: Testing	Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
Sari et al. [11]	2010	Türkiy	169	169	MLP: the designed ANN consisted of feed-forward back propagation, two hidden layers	LBP	NR	Comparison of ANNs and adaptive neuro-fuzzy inference system for the assessment of the LBP	The results showed that the ANNs and adaptive neuro-fuzzy inference system behave very similar to real data. The use of these systems can be used to successfully diagnose the back pain intensity.
Veronezi et al. [12]	2015	Brazil	68 Radiographies for the training stage	68 Images for tests and 70 for validation	Neural networks	Osteoarthritis of the lumbar spine	NR	For the diagnosis of osteoarthritis of the lumbar spine	The validation was carried out on the best results, achieved accuracy of 62.85%, sensitivity of 65.71%, and specificity of 60%. Although the neural network presented an average efficacy, because this was an innovative study, its results showed a potential for the use of computer-based artificial neural networks to assist and support practitioners.
Zhang et al. [13]	2017	China	235 Radiographs	105 Radiographs	DNN	Scoliosis assessment	Yes	To perform automatic measurements of Cobb angle for scoliosis assessment	The differences between the computer-aided measurement and the manual measurement by the surgeon were higher than 5 . The variability of Cobb angle measurements could be reduced if the DNN system was trained with enough vertebral patches.
Jamaludin et al. [14]	2017	UK	90% in a training set of 1,806 patients	10% in an independent sample of 203 patients	CNN	Lumbar IVDs and vertebral bodies	Yes	To automate the process of grading lumbar IVDs and vertebral bodies from MRIs.	The detection system achieved 95.6% accuracy in terms of disc detection and labeling. The model was able to produce predictions of multiple pathological grading that consistently matched those of the radiologist. The system could be beneficial in aiding clinical diagnoses in terms of objectivity of grading and the speed of analysis.
Wang et al. [15]	2017	China	A set of 26 cases	A set of 26 cases	Deep Siamese neural networks	Spinal metastasis	NR	A multi-resolution approach for spinal metastasis detection in MRI	The results showed that the proposed approach correctly detects all the spinal metastatic lesions. The results indicated that the proposed Siamese neural network method, combined with the aggregation strategy, provided a viable strategy for the automated detection of spinal metastasis in MRI images.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Kim et al. [16]	2018	USA	15,840	6,789	ANNs	Posterior lumbar spine fusion	Yes	Comparison of ANNs, LR, and ASA class to identify risk factors of developing complications following posterior lumbar spine fusion	ANN and LR both outperformed ASA class for predicting all four types of complications. ANN had greater sensitivity than LR for detecting wound complications and mortality. In summary, machine learning in the form of LR and ANNs were more accurate than benchmark ASA scores for identifying risk factors of developing complications following posterior lumbar spine fusion, suggesting they are potentially great tools for risk factor analysis in spine surgery.
Kim et al. [17]	2018	South Korea	Total training epoch was 200	The experiments were done using 5-fold cross validation and each experiment had 5 test images and 20 training images.	CNN	IVDs	Yes	To segmentation of the IVDs from MR spine images	The proposed network achieved 3% higher DSC than conventional U-net for IVD segmentation (89.44% vs. 86.44%, respectively; p<0.001). For IVD boundary segmentation, the proposed network achieved 10.46% higher DSC than conventional U-net (54.62% vs. 44.16%, respectively; p<0.001).
Kim et al. [18]	2018	South Korea	Four-fold cross validation on a patient-level independent split	Four-fold cross validation on a patient-level independent split	DCNN	Tuberculous and pyogenic spondylitis	Yes	To differentiate between tuberculous and pyogenic spondylitis on MR imaging, compared to the performance of skilled radiologists	When comparing the AUC value of the DCNN classifier (0.802) with the pooled AUC value of the three readers (0.729), there was no significant difference (p=0.079). In differentiating between tuberculous and pyogenic spondylitis using MR images, the performance of the DCNN classifier was comparable to that of three skilled radiologists.
Han et al. [19]	2018	Canada	The dataset comprises 253 lumbar scans from 253 patients	The dataset comprises 253 lumbar scans from 253 patients	Recurrent neural network	IVDs, vertebrae, and neural foraminal stenosis	NR	To perform automated segmentation and classification (i.e., normal and abnormal) of IVDs, vertebrae, and neural foramen in MRIs	Extensive experiments on MRIs of 253 patients have demonstrated that Spine-GAN achieved high pixel accuracy of 96.2%, Dice coefficient of 87.1%, sensitivity of 89.1%, and specificity of 86.0%, which revealed its effectiveness and potential as a clinical tool.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Chmelik et al. [20]	2018	Czechia	Dataset consisted of 120,000 samples in total, in 31 cases	Dataset consisted of 120,000 samples in total, in 31 cases	DCNN	Metastatic spinal lesions	Yes	To address the segmentation and classification to define metastatic spinal lesions in 3D CT data	Algorithm enables detection, segmentation and classification of small lesions greater than 1.4 mm^3 (with diameter greater than 0.7 mm) and works also with cervical vertebrae not treated in other considered methods for spinal analysis of CT scans.
Liao et al. [21]	2018	USA	242 CT scans from 125 patients are used for training	60 CT scans for testing	Deep learning, CNN, recurrent neural network, multi-task learning	Vertebrae	NR	To automatically vertebrae identification and localization in spinal CT images	The experimental results showed that approach outperforms the state-of-the-art methods by a significant margin.
Al Arif et al. [22]	2018	UK	124 X-ray images	172 Images	CNN	Cervical vertebrae	NR	To automatically framework for segmentation of cervical vertebrae in X-ray images	A Dice similarity coefficient of 0.84 and a shape error of 1.69 mm have been achieved. The framework could take an X-ray image and produce a vertebrae segmentation result without any manual intervention.
Han et al. [23]	2018	China	160 (80%)	40 (20%)	DMML-Net	LNFS	NR	To automatically pathogenesis-based diagnosis of lumbar neural foraminal stenosis	DMML-Net achieves high performance (0.845 mean average precision) on T1/T2-weighted MRI scans from 200 subjects. This method showed an efficient tool for clinical LNFS diagnosis.
Li et al. [24]	2018	China	Voxel changes for each IVD in 12 subjects within 2 time points	Voxel changes for each IVD in 12 subjects within 2 time points	FCN	IVDs	Yes	To automatically localization and segmentation of IVDs from multi-modality 3D MR data	Algorithm achieved state-of-the-art IVD segmentation performance from multi-modality images. Compared with network trained with single-scale context image, the proposed 3D multi-scale FCN could generate features with high discrimination capability.
Zhou et al. [25]	2019	China	The dataset contains 4,417 videos	The dataset contains 4,417 videos	Deep learning	Lumbar vertebras	NR	To automatically detect lumbar vertebras in MRI images	Algorithm achieved the accuracy of 98.6% and the precision of 98.9%. Most failed results were involved with wrong S1 locations or missed L5. The study demonstrated that a lumbar detection network supported by deep learning can be trained successfully without annotated MRI images.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Wang et al. [26]	2019	China	Data set of 98 spine CT scans	Data set of 98 spine CT scans	Combining deep stacked sparse autoencoder contextual features and structured regression forest	Vertebrae	Yes	To automatically vertebra localization and identification from CT	Compared with the hidden Markov model and the method based on CNN, the proposed approach could effectively and automatically locate and identify spinal targets in CT scans, and achieve higher localization accuracy, low model complexity
Lessmann et al. [27]	2019	Netherland	Five diverse datasets, including multiple modalities (CT and MR)	Five diverse datasets, including multiple modalities (CT and MR)	CNN	Vertebrae	Yes	To automatically vertebra segmentation and identification	The anatomical identification had an accuracy of 93%. Vertebrae were classified as completely or incompletely visible with an accuracy of 97%. The proposed iterative segmentation method compares favorably with state-of-the-art methods and is fast, flexible, and generalizable.
Lang et al. [28]	2019	China	A total of 61 patients with clinical spinal MRI database with a DCE sequence	A total of 61 patients (30 lung cancers and 31 non-lung cancers)	CNN	Spinal metastases originated from lung and other cancers	Yes	To differentiate metastatic lesions in the spine originated from primary lung cancer and other cancers	Classification using CNN achieved a mean accuracy of 0.71±0.043, whereas a convolutional long short-term memory improved accuracy to 0.81±0.034. DCE-MRI machine-learning analysis methods had potential to predict lung cancer metastases in the spine.
Galbusera et al. [29]	2019	Italy	443	50	Deep learning approach	To extract anatomical parameters from biplanar radiographs of the spine	NR	To automatically determine the shape of the spine	The standard errors of the estimated parameters ranged from 2.7° (for the pelvic tilt) to 11.5° (for the L1–L5 lordosis). The proposed method was able to automatically determine the spine shape in biplanar radiographs and calculate anatomical and posture parameters in a wide scenario of clinical conditions with a very good visual performance.
Hopkins et al. [30]	2019	USA	78	26	ANN	CSM	NR	(1) To predict CSM diagnosis; and (2) to predict CSM severity	Median accuracy of model was 90.00%. Machine learning provided a promising method for prediction, diagnosis, and even prognosis in patients with CSM.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size: Training	No. of sample size: Testing	Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
Horng et al. [31]	2019	Taiwan	35 Images captured from young scoliosis. The dataset consisted of 595 vertebra images	35 Images captured from young scoliosis	CNN approach	Cobb angle measurement of Spine	Yes	To automatically measure spine curvature using the anterior-posterior view spinal X-ray images	The segmentation results of the Residual U-Net were superior to the other two CNNs. The proposed system can be applied in clinical diagnosis to assist doctors for a better understanding of scoliosis severity and for clinical treatments.
Pang et al. [32]	2019	China	T1-weighted MR images of 215 subjects and T2-weighted MR images of 20 subjects	T1-weighted MR images of 215 subjects and T2-weighted MR images of 20 subjects	Cascade amplifier regression network	Spine	NR	To automatically quantitative measurement of the spine (i.e., multiple indices estimation of heights, widths, areas, and so on for the vertebral body and disc)	The proposed approach achieved impressive performance with mean absolute errors of 1.22±1.04 mm and 1.24±1.07 mm for the 30 lumbar spinal indices estimation of the T1-weighted and T2-weighted spinal MR images, respectively. The proposed method showed a great potential in clinical spinal disease diagnoses and assessments.
Li et al. [33]	2019	China	120 Cases were used for experiments	120 Cases were used for experiments	DNN	To paraspinal muscle segmentation	NR	To automatically segmentation of the paraspinal muscle in MRI	The experimental results show that the model can achieve higher predictive capability. The dice coefficient of the multifidus is as high as 0.949, and the Hausdorff distance is 4.62 mm. The proposed method can quickly calculate the cross-sectional area of the paraspinal muscles, which provides a convenient condition for doctors to screen sarcopenia and also quantify the changes of paraspinal muscles before and after lumbar spine surgery.
Chen et al. [34]	2019	China	End-to-end training at the spine level is proposed to allow the FCN to directly learn the long-range image patterns from full-size CT volumes	End-to-end training at the spine level is proposed to allow the FCN to directly learn the long-range image patterns from full-size CT volumes	FCN	Vertebrae identification and localization	NR	To automatically identification and localization of vertebrae in spinal CT imaging	The proposed framework was quantitatively evaluated on the public dataset from the MICCAI 2014 Computational Challenge on Vertebrae Localization and Identification and demonstrates an identification rate (within 20 mm) of 94.67%, a mean identification rate of 87.97%, and a mean error distance of 2.56 mm on the test set, thus achieving the highest performance reported on this dataset.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Rak et al. [35]	2019	Germany	The first whole spine images of 64 subjects were contained. The second 23.	The first whole spine images of 64 subjects were contained. The second 23.	Combining CNNs and star convex cuts	Whole spine segmentation by MRI	Yes	To automatically approach for fast vertebral body segmentation in 3D MRI of the whole spine	Complete whole spine segmentation took 32.4±1.92 seconds on average. Results were superior to those of previous works at a fraction of their run time, which illustrated the efficiency and effectiveness of their whole spine segmentation approach.
Pan et al. [36]	2019	China	Cobb angles on 248 chest X-rays were measured automatically using a computer-aided method	Cobb angles on 248 chest X-rays were measured automatically using a computer-aided method	The Cobb angle of the spinal curve was measured from the output of the Mask R-CNN models	Spine alignment assess	Yes	To automatically measure the Cobb angle and diagnose scoliosis on chest X-rays, a computer-aided method was proposed	Intraclass correlation coefficient between the computer-aided and manual methods for Cobb angle measurement was 0.854. These results indicated that the computer-aided method had good reliability for Cobb angle measurement on chest X-rays. In conclusion, the computer-aided method has potential for automatic Cobb angle measurement and scoliosis diagnosis on chest X-rays.
Weng et al. [37]	2019	Taiwan	The ResUNet was trained and evaluated on 990 standing lateral radiographs	The ResUNet was trained and evaluated on 990 standing lateral radiographs	CNN	Spine alignment assess	Yes	To develop a CNN tools for measuring the SVA from lateral radiography of whole spine for key point detection (ResUNet)	The SVA calculation takes approximately 0.2 seconds per image. The intra-class correlation coefficient of the SVA estimates between the algorithm and physicians of different years of experience ranges from 0.946 to 0.993, indicating an excellent consistency. The superior performance of the proposed method and its high consistency with physicians proved its usefulness for automatic measurement of SVA in clinical settings.
Huang et al. [38]	2019	China	50 Sets lumbar MRIs	50 Sets lumbar MRIs	DL	Vertebrae and IVDs on lumbar spine	NR	To develop a DL based program (Spine Explorer) for automated segmentation and quantification of the vertebrae and IVDs on lumbar spine MRIs	The trained Spine Explorer automatically segments and measures a lumbar MRI in half a second, with mean intersection-over-union of 94.7% and 92.6% for the vertebra and disc, respectively. Spine Explorer was an efficient, accurate, and reliable tool to acquire comprehensive quantitative measurements for lumbar vertebra and disc. Implication of such deep learning-based program can facilitate clinical studies of the lumbar spine.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Jakubicek et al. [39]	2019	Czech Republic	130 CT scans	130 CT scans	Two CNNs together with a spine tracing algorithm	Spine-ends and spine centerline delimitation assessment are important in many spine diagnostic tasks	NR	To develop a CNN to automatic spine centerline detection in CT data	Based on the evaluation of 130 CT scans including heavily distorted and complicated cases, it turned out that this new combination enables fast and robust detection with almost 90% of correctly determined spinal centerlines with computing time of fewer than 20 seconds.
Lyu et al. [40]	2019	China	75 Groups imaging data	75 Groups imaging data	CNN	To assessment of spine scoliosis by Scolioscan from 3D ultrasound	Yes	To develop a CNN to select the best ultrasound images automatically, and compare with the classification method of DenseNet.	The result showed that the proposed CNN achieves the perfect accuracy of 100% while conventional DenseNet achieved 35% only. This proves that the CNN was more suitable to highlight the best quality of ultrasound image from multiple mediocre ones.
Watanabe et al. [41]	2019	Japan	10,788 Moiré image-radiograph pairs	198 Moiré image-radiograph pairs	CNN	To assessment of spine scoliosis	NR	To create a scoliosis screening system that estimates spinal alignment, the Cobb angle, and vertebral rotation from moiré images.	The proposed method of estimating the Cobb angle and the angle of virtual reality from moiré images using a CNN was expected to enhance the accuracy of scoliosis screening.
Kök et al. [42]	2019	Türkiy	300 Individuals aged between 8 and 17 years	300 Individuals aged between 8 and 17 years	k-NN, NB, Tree, ANN, SVM, RF, and LR algorithms were used.	CVS	Yes	To determine CVS for growth and development periods by the frequently used seven artificial intelligence classifiers, and to compare the performance of these algorithms with each other	kNN and LR algorithms had the lowest accuracy values. SVM-RF-Tree and NB algorithms had varying accuracy values. ANN could be the preferred method for determining CVS.
Iriondo et al. [43]	2020	USA	38 Scans from 31 unique patients, with a total of 80 segmented slices	20 Segmented slices	CNN to segment lumbar IVDs by MRI	Lumbar IVDs	NR	To assess associations between disc degeneration, disability, and LBP	This work presented a scalable pipeline for fast, automated assessment of disc relaxation times, and voxel-based relaxometry that overcomes limitations of current region of interest-based analysis methods and may enable greater insights and associations between disc degeneration, disability, and LBP.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Lee et al. [45]	2020	South Korea	233	101	Deep convolutional networks	To identify individuals with abnormal BMD from spine X-ray images	NR	To analysis of spine X-ray features extracted by deep learning to alert high-risk osteoporosis populations	A combination of feature extraction was found, by VGGnet and classification by random forest based on the maximum BCR yielded the best performance in terms of the AUC (0.74), accuracy (0.71), sensitivity (0.81), specificity (0.60), BCR (0.70), and F1-score (0.73). Finally, the combination for the best performance in predicting high-risk populations with abnormal BMD was identified.
Won et al. [44]	2020	South Korea	542 L4–5 axial MR images	542 L4–5 axial MR images	DCNN	To identify spine stenosis grading from MRI	Yes	To compare the diagnostic agreement between the experts and trained artificial CNN classifiers	Final agreement between the expert and the model trained with the labels of the expert was 77.9% and 74.9%, and the differences between each expert and the trained models were not significant. They were concluded that automatic diagnosis using deep learning may be feasible for spinal stenosis grading.
Lee et al. [46]	2020	South Korea	280 Pairs of lumbar spine CT scans and MR T2 images	15 Pairs of lumbar spine CT scans and MR T2 images	GANs	To diagnosis of spine disease	Yes	To apply GANs, to synthesize spine MR images from CT images	The mean overall similarity of the synthetic MR T2 images evaluated by radiologists was 80.2%. Synthesis of MR images from spine CT images using GANs will improve the spine diagnostic usefulness of CT. To better inform the clinical applications of this technique, further studies are needed involving a large dataset, a variety of pathologies, and other MR sequence of the lumbar spine.
Bae et al. [47]	2020	South Korea	Patients (N=17, 1,684 slices)	Healthy controls (N=24, 3,490 slices)	CNN	Cervical spine	Yes	To identify superior and inferior vertebrae in a single slice of CT images, and a post-processing for 3D segmentation and separation of cervical vertebrae	The results demonstrated that automated method achieved comparable accuracies with inter- and intra-observer variabilities of manual segmentation by human experts, which is time consuming.

(Continued on next page)

Table 1. Continued

Author	Year	Country	No. of sample size: Training	No. of sample size: Testing	Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
Jakubicek et al. [48]	2020	Czech Republic	The more samples, the more accurate	The more samples, the more accurate	CNN	Incomplete spines assessment in patients with bone metastases and vertebral compression by CT imaging	NR	To localization and identification of vertebrae in 3D CT scans of possibly incomplete spines in patients with bone metastases and vertebral compressions	The proposed framework, which combined several advanced methods including also three CNNs, worked fully automatically even with incomplete spine scans and with distorted pathological cases. The achieved results allow including the presented algorithms as the first phase to the fully automated computer-aided diagnosis system for automatic spine-bone lesion analysis in oncological patients.
Kim et al. [49]	2020	South Korea	330 CT images	14 CT images	CNN for segmentation	To diagnosis of back pain	Yes	To improve diagnosis of back pain by spine segmentation in CT scans using algorithmic methods	The CNN method achieved an average dice coefficient of 90.4%, a precision of 96.81%, and an F1-score of 91.64%. The proposed CNN approach can be very practical and accurate for spine segmentation as a diagnostic method.
Rehman et al. [50]	2020	Pakistan	25 CT image data (both healthy and fractured cases)	25 CT image data (both healthy and fractured cases)	A novel combination of traditional region-based level set with deep learning framework	To diagnosis of osteoporotic fractures by vertebral bone segmentation	NR	To predict shape of vertebral bones accurately	Dice score was found to be 96.4%±0.8% without fractured cases and 92.8%±1.9% with fractured cases in dataset (lumber and thoracic). The proposed technique outperformed other state-of-the-art techniques and handled the fractured cases for the first time efficiently.

LVQ was used for quantizing the learning data to feed them to ANN.

ANN, artificial neural network; MLP, multilayer perceptron neural networks; LBP, low back pain; NR, not reported; LVQ, learning vector quantization; DNN, deep neural network; CNN, convolutional neural network; IVD, intervertebral disc; MRI, magnetic resonance imaging; LR, logistic regression; ASA, American Society for Anesthesiology; MR, magnetic resonance; DSC, Dice similarity coefficient; DCNN, deep convolutional neural network; AUC, area under the curve; 3D, three-dimensional; CT, computed tomography; DMML-Net, deep multiscale multitask learning network; LNFS, lumbar neural foraminal stenosis; FCN, fully convolutional networks; DCE, dynamic contrast enhanced; CSM, cervical spondylotic myelopathy; SVA, sagittal vertical axis; DL, deep learning; k-NN, k-nearest neighbors; NB, Naive Bayes; Tree, decision tree; SVM, support vector machine; RF, random forest; CVS, cervical vertebrae stages; BMD, bone mineral density; BCR, balanced classification rate; GANs, generative adversarial networks.

the commonly used clinical scoring methods [16]. A summary of the studies is shown in Table 2.

4. Outcome prediction

Table 3 summarizes the studies that used ANNs for outcome prediction. ANNs have been used to predict outcome in lumbar spinal stenosis (LSS) [3] and LDH [4], predict recurrent LDH [66], enhance surgical decision making for LSS [67], develop ANN algorithms for prediction of in-hospital and 90-day post-discharge mortality in spinal epidural abscess [68], predict non-routine discharge for patients undergoing surgery for lumbar disc disorders [69], assess vertebral strength and predict vertebral fracture risk in elderly patients [70], predict 30-day readmission after posterior lumbar fusion [71], and predict surgical site infections after posterior spinal fusion [72].

5. The use of artificial neural networks for the biomechanical assessments of spinal diseases

A clear understanding of biomechanical principles is important in the management of spinal disorders. There are ANNs studies focused on the biomechanics of spine via clarification of joint moments, spinal loads, and muscle forces [75-77]. Other than that, application of ANNs for optimization of the design of spinal pedicle screws [78], prediction of vertebral strength through machine learning models [70], determination of the consistency of the patient pain drawing in lumbar spine disease [73], prediction of low bone mineral density [74], recognition of low back pain patients from healthy population performed static standing tasks [79], estimation of three-dimensional whole-body posture, lumbosacral moments and spinal loads during load-handling activities [80] and automated tracking of lumbar vertebras with rotated bounding boxes in digitalized video fluoroscopic imaging, and motion and gait analysis [81] have been reported. These findings are summarized in Table 4.

Discussion

To the best of our knowledge, this is the first review devoted exclusively to an application of ANN in support of decision for management of spinal disease. Our findings offer a summary of relevant publications and a roadmap to guide future research related to the use of ANNs in spinal disease. Precisely, our findings showed that ANNs are powerful tools with the ability to improve understanding of predictive metrics, prognosis, diagnosis and biomechanical assessment in spinal diseases. Moreover, ANNs have shown consistent superiority over the traditional statistical approaches. In light of the continuous development of hardware and software methods, and advanced computational science and technology, wider consideration and broader application artificial intelligence in spinal disease is expected in the near future [2].

The number of publications on neural networks in spinal diseases has increased rapidly over the past few years, wherein a majority of the publications were in the domain of diagnosis of spinal disorders, followed by prognosis, prediction, and biomechanical for spinal applications. A number of ANN studies have focused on preoperative assessment, planning, intraoperative assistance and outcome prediction in spine surgery. Recently, Khor et al. [82] successfully developed a state-of-the-art use of a logistic regression (LR) model to predict the patient-reported outcomes in lumbar fusion surgery. They developed a clinical prediction tool model to determine the probabilities of improvement in function, back pain, and leg pain in lumbar fusion candidates at 1-year follow-up after surgery. This model showed a good accuracy in the validation cohorts. The same group also provided an online version of their prediction model for public use (https://becertain.shinyapps.io/lumbar_fusion_calculator/), where a clinician and/or patient can enter the individual demographics to predict a patient's likelihood of benefiting from a lumbar fusion procedure [83]. Besides, Karhade et al. [64] developed a machine learning tool for predicting prolonged postoperative opioid prescription in the patients undergoing LDH surgery (https://sorg-apps.shinyapps.io/lumbardiscopioid/). It is worth mentioned that preoperative prediction of opioid use could improve the risk stratification, shared decision-making, and patient counseling before LDH surgery [84]. In addition, Karhade et al. [61] developed a machine learning tool to automatically predict 90-day and 1-year mortality in spinal metastatic disease (https://sorg-apps.shinyapps.io/spinemetssurvival/) [85]. Meanwhile, the same group also developed a machine learning algorithm for predicting discharge disposition after elective inpatient surgery for lumbar disc disease [55], the model is available at https://sorg-apps.shinyapps.io/discdisposition/ [86]. Furthermore, there is also an

Table 2. A list of papers on ANN used in spine prognosis

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Dickey et al. [51]	2002	USA	157	29	MLP: three-layer ANNs were used with 32 inputs, one hidden layer and one output	LBP	Yes	To investigate the relationship between intervertebral motion, intravertebral deformation, and pain in chronic LBP patients	The neural network model showed a strong relationship between observed and predicted pain (r=0.997). ANNs are able to effectively describe relationships between pain and vertebral motion in chronic LBP.
Lin et al. [52]	2008	Taiwan	1,126	375	MLP	Spine	Yes	Comparison of ANNs and LR to identify patients with high risk of hypotension during spinal anesthesia	The ANN model had a sensitivity of 75.9% and specificity of 76.0%. The LR model had a sensitivity of 68.1% and specificity of 73.5%. The area under receiver operating characteristic curves were 0.796 and 0.748. The ANN model performed significantly better than the LR model. The prediction of clinicians had the lowest sensitivity of 28.7%, 22.2%, 21.3%, 16.1%, and 36.1%, and specificity of 76.8%, 84.3%, 83.1%, 87.0%, and 64.0%.
Parsaeian et al. [53]	2012	Iran	17,294	17,295	MLP: a three-layer perceptron with nine inputs, three hidden and one output neurons was employed	LBP	Yes	To compare empirically predictive ability of an artificial neural network with a LR in prediction of LBP	The area under the ROC curve (SE), root mean square, and -2loglikelihood of the logistic regression was 0.752 (0.004), 0.3832, and 14,769.2, respectively. The area under the ROC curve (SE), root mean square and -2log-likelihood of the artificial neural network was 0.754 (0.004), 0.3770, and 14,757.6, respectively. ANNs would give better performance than LR.
Papić et al. [54]	2016	Serbia	Data set included 145 patients, and 10-fold cross validation	10-Fold cross validation	The classification problem was approached using decision trees, SVM and MLP combined with RELIEF algorithm for feature selection.	LDH	Yes	To predict the return to work after operative treatment of LDH	MLP provided best recall of 0.86 for the class of patients not returning to work. The predictive modeling indicated at the most decisive risk factors in prolongation of work absence: psychosocial factors, mobility of the spine and structural changes of facet joints and professional factors including standing, sitting, and microclimate.

(Continued on next page)

Table 2. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Kim et al. [16]	2018	USA	15,840 (70%)	6,789 (30%)	ML	Posterior lumbar spine fusion surgery	Yeas	To automatically predict (identify risk factors for) complications following posterior lumbar spine fusion and compared with regression model (LR)	Though ML and LR had comparable AUC values for predicting all types of complications as cardiac complications, wound complications, venous thromboembolism, and mortality. However, ANN had greater sensitivity than LR for detecting wound complications and mortality. ML and LR were more accurate than benchmark ASA scores
Karhade et al. [55]	2018	USA	21,091	5,273	ML algorithms	Lumbar degenerative disc	NR	To use ML to develop an open-access web application for preoperative prediction of nonroutine discharges in surgery for elective inpatient lumbar degenerative disc disorders	The rate of nonroutine discharge for 26,364 patients who underwent elective inpatient surgery for lumbar degenerative disc disorders was 9.28%. Machine learning algorithms showed promising results on internal validation for preoperative prediction of nonroutine discharges.
Arvind et al. [56]	2018	USA	14,615 Patients	6,264	ANN, LR, SVM, and RF models	Cervical discectomy	Yes	To develop predictive algorithms for postoperative complications following anterior cervical discectomy and fusion	The SVM and RF models were no better than random chance at predicting any of the postoperative complications (p<0.05). ANN and LR algorithms outperform ASA physical status classification for predicting individual postoperative complications. Additionally, neural networks have greater sensitivity than LR when predicting mortality and wound complications.
Han et al. [57]	2019	USA	355,607 (70%)	152,403 (30%)	MLA	Spine surgery	Yes	To develop and evaluate a set of predictive models for common adverse events after spine surgery	The predictive models for adverse events following spine surgery built based on this data showed greater accuracy versus the previous models, with AUC ranging between 0.7 and 0.76, which account for patient-, diagnosis-, and procedure-related factors.

(Continued on next page)

Table 2. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
DeVries et al. [58]	2019	Canada	862 Patients included that walk (n=323) not walk (n=318)	862 Patients included	MLA	tSCI	Yes	To automatically prognosticate walking recovery in patients with tSCI and compared with LR	MLAs had comparable prognostication as the previously reported models. Overall, no relevant differences were found between the models suggesting that using a more sophisticated MLA and a greater amount of neurological data does not improve the prediction of walking recovery in tSCI patients.
Staartjes et al. [59]	2019	Netherland	A total of 422 were included and data training, sets was 60%.	Data validation, and test sets was in a 20%/20% ratio.	Deep learning-based analytics	LDH	Yes	To evaluate a clinically relevant improvement in leg pain, back pain, and functional disability after LDH surgery by deep learning and compared with regression model	After 1 year, 337 (80%), 219 (52%), and 337 (80%) patients reported a clinically relevant improvement in leg pain, back pain, and functional disability, respectively. The regression models provided inferior performance measures for each of the outcomes. The study demonstrated that generating personalized and robust deep learning-based analytics for outcome prediction was feasible even with limited amounts of center-specific data.
Karhade et al. [60]	2019	USA	1,432 (80%)	358 (20%)	MLA	Spinal metastatic disease	NR	To automatically predict 30-day mortality of patients undergoing surgery for spinal metastatic disease	The 30-day mortality for the 1,790 patients undergoing surgery for spinal metastatic disease was 8.49%. MLAs were promising for prediction of postoperative outcomes in spinal oncology and these algorithms could be integrated into clinically useful decision tools.
Karhade et al. [61]	2019	USA	587 (80%)	145 (20%)	Five models (penalized logistic regression, random forest, stochastic gradient boosting, neural network, and support vector machine)	To develop predictive algorithms for spinal metastatic disease	NR	To automatically predict 90-day and 1-year mortality in spinal metastatic disease	Overall, 732 patients were identified with 90-day and 1-year mortality rates of 181 (25.1%) and 385 (54.3%), respectively. The final models were incorporated into an open access web application able to provide predictions as well as patient-specific explanations of the results generated by the algorithms.

(Continued on next page)

Table 2. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Weber et al. [62]	2019	USA	Train and test a CNN for muscle segmentation and automatic money flow index calculation were performed using high resolution fat-water images from 39 participants	Train and test were performed using high resolution fat-water images from 39 participants	Deep learning CNN models	Muscle fat infiltration following whiplash injury in cervical spine	NR	To automatically quantification of muscle fat infiltration following whiplash injury	Overall, CNN's may improve d the efficiency and objectivity of muscle measures allowing for the quantitative monitoring of muscle properties in disorders of and beyond the cervical spine.
Goyal et al. [63]	2019	USA	A total of 59,145 cases were analyzed. The best combination selected by a 10-fold cross-validation procedure.	10-Fold cross-validation procedure	ML algorithms	Spinal fusion surgery	NR	To develop algorithms to predict discharge to rehabilitation and unplanned readmissions in patients receiving spinal fusion	The incidence rates of discharge to nonhome facility and 30-day unplanned readmission were 12.6% and 4.5%, respectively. All classification algorithms showed excellent discrimination (AUC >0.80; range, 0.85–0.87) for predicting nonhome discharge. Multiple ML algorithms were found to reliably predict nonhome discharge with modest performance noted for unplanned readmissions
Karhade et al. [64]	2019	USA	4,331 (80%)	1,082 (20%)	ML algorithms	LDH	Yes	To develop algorithms for prediction of prolonged opioid prescription after surgery for LDH	Overall, 5,413 patients were identified, with sustained postoperative opioid prescription of 416 (7.7%) at 90 to 180 days after surgery. The elastic-net penalized logistic regression model had the best discrimination (c-statistic 0.81) and good calibration and overall performance. They showed that preoperative prediction of prolonged postoperative opioid prescription with this model can help identify candidates for increased surveillance after surgery.

(Continued on next page)

Table 2. Continued

Author	Year	Country	No. of sample size: Training	No. of sample size: Testing	Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
Ryu et al. [65]	2020	South Korea	870	218	RED_SNN: final network consists of embedding layer, long short-term memory layer, four fully connected layers.	Spino-pelvic chondrosarcoma	Yes	To predict survival following a spino-pelvic chondrosarcoma diagnosis	The median c-index of the five validation sets was 0.84 (95% confidence interval, 0.79–0.87). Risk estimate distance survival neural network (RED_SNN) was a valid method to predict survival for spinal and pelvic chondrosarcoma, and it appears to be comparable to other methods.

ANN, artificial neural network; MLP, multilayer perceptron neural networks; LBP, low back pain; LR, logistic regression; ROC, receiver operating characteristic; SE, standard error; SVM, support vector machine; LDH, lumbar disc herniation; ML, machine learning; NR, not reported; AUC, area under the curve; ASA, American Society for Anesthesiology; NR, not reported; RF, random forest decision tree; MLA, machine learning algorithms; tSCI, traumatic spinal cord injury; CNN, convolutional neural network; RED_SNN, risk estimate distance survival neural network.

application at https://sorg-apps.shinyapps.io/spinemets/ [87] which allows prediction of 30-day mortality after surgery for spinal metastatic disease [60]. Nonetheless, these machine learning tools are fitted for general educational purposes and they are not capable of substituting the professional medical advice, consultation, diagnosis, or treatment [82-87]. One might inquire about how the ANNs can assist in the clinical decision-making process? Spine neural network tools will never replace human experts, but it helps in screening and can be used by the experts to validate their diagnosis, prognosis, and prediction. More importantly, ANNs can be used to identify the variables that experts may not observe, thus enhancing the diagnostic acumen of experts. As aforementioned, the currently available web applications are not a good fit with the clinical practice setting. However, development of this software tool is a prerequisite for an international consultancy group to satisfy the diverse needs as randomized clinical trials (RCT) data exists that specifically examines this tool. Nevertheless, ANNs will never replace human expert decision-making, but it can assist in validating the routine decision-making process [2].

Few neural network studies have focused on medical imaging analysis. There are a wide variety of medical imaging modalities and magnetic resonance imaging (MRI) is majorly applied for clinical diagnosis and prognosis. Recently, some studies have demonstrated successful application of artificial intelligence algorithms for spine medical image segmentation [17,19,20,22,24,27,33,35,38,47,49], computer-aided spine diagnosis [84-87], and disease detection and classification [10,45]. In other words, spinal images could be analyzed, processed, and categorized by using neural network. By selecting a suitable training set and learning process, neural networks is appropriate for recognition of unusual images [88]. Artificial intelligence will play a vital role in the development of medical image analysis methods. However, deep learning architecture requires a large amount of training data and computational power. Currently, there is also a rising interest with respect to the digital image analysis solutions with artificial intelligence for clinical applications. Such applications aim to increase diagnostic and prognostic accuracy, reliability, and efficiency by enabling quantitative image analysis. For instance, Oxford SpineNet software system, a machine learning based system for automated analysis of spinal T2 MRI scans acquired from a DICOM (Digital Imaging and Communications in Medicine) file,

Table 3. A list of papers on ANN used in spine outcome prediction

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Azimi et al. [3]	2014	Iran	84	84	ANN model	LSS	Yes	To develop an ANN model for predicting 2-year surgical satisfaction, and to compare the new model with traditional predictive tools in patients with lumbar spinal canal stenosis	The ANN model displayed a better accuracy rate in 96.9% of patients, a better Hosmer-Lemeshow statistic in 42.4% of patients, and a better receiver operating characteristic-AUC in 80% of patients, compared with the LR model. ANNs can predict 2-year surgical satisfaction in LSS patients with a high level of accuracy.
Azimi et al. [66]	2015	Iran	201	201	ANN model	Recurrent LDH	Yes	To develop an ANN model to predict recurrent LDH	Compared with the LR model, the ANN model was associated with superior results: accuracy rate, 94.1%; H-L statistic, 40.2%; and AUC, 0.83% of patients. ANNs could be used to predict the diagnostic statues of recurrent and nonrecurrent group of LDH patients before the first or index microdiscectomy.
Azimi et al. [4]	2016	Iran	102	101	ANN model	LDH	Yes	To develop an ANNs model for predict successful surgery outcome in LDH	Compared to the LR model, the ANN model showed better results: accuracy rate, 95.8%; H-L statistic, 41.5%; and AUC, 0.82% of patients. ANNs can predict successful surgery outcome with a high level of accuracy in LDH patients.
Azimi et al. [67]	2017	Iran	174	86	ANN model	LSCS	Yes	To accurately select patients for surgery or non-surgical options and to compare such with the traditional clinical decision-making approach in LSCS patients	The ANN model displayed better accuracy rate (97.8%), a better H-L statistic (41.1%) which represented a good-fit calibration, and a better AUC (89.0%), compared to the LR model. ANN model could predict the optimal treatment choice for LSCS patients in clinical setting and is superior to LR model.
Karhade et al. [68]	2019	USA	844 (80%)	209 (20%)	ML algorithm	SEA	NR	To develop ML algorithms for prediction of in-hospital and 90-day postdischarge mortality in SEA	Overall, 1,053 SEA patients were identified in the study, with 134 (12.7%) experiencing in-hospital or 90-day postdischarge mortality. The stochastic gradient boosting model achieved the best performance across discrimination, c-statistic=0.89, calibration, and decision curve analysis. ML algorithms showed promise on internal validation for prediction of 90-day mortality in SEA.

(Continued on next page)

Table 3. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Stopa et al. [69]	2019	USA	144 Patients	144 Patients	ML algorithm	Lumbar disc disorders surgery	NR	To predict nonroutine discharge for patients undergoing surgery for lumbar disc disorders	A nonroutine discharge rate of 6.9% (n=10). The neural network algorithm generalized well to the institutional data, with a c-statistic of 0.89. ML showed that a reliable method for identifying patients with lumbar disc disorder at risk for nonroutine discharge,
Zhang et al. [70]	2019	China	58	22	ML	Lumbar vertebral strength of elderly men	NR	To predict vertebral strength based on clinical quantitative computed tomography images by using machine learning	High accuracy was achieved to predict vertebral strength. This study provided an effective approach to predict vertebral strength and showed that it may have great potential in clinical applications for noninvasive assessment of vertebral fracture risk.
Hopkins et al. [71]	2019	USA	17,448	5,816	DNN	Spinal fusions	NR	To develop an AI model to predict 30-day readmissions after posterior lumbar fusion	Mean positive predictive value was 78.5%. Mean negative predictive value was 97%. The DNN model was able to predict those patients who would not require readmission.
Hopkins et al. [72]	2020	USA	3,034	1,012	DNN	Spinal fusions	NR	To develop an AI model for predict surgical site infections after posterior spinal fusions	The five highest weighted variables were congestive heart failure, chronic pulmonary failure, hemiplegia/paraplegia, multilevel fusion, and cerebrovascular disease, respectively. Notable factors that were protective against infection were intensive care unit admission, increasing Charlson Comorbidity Index score, race (White), and being male. They reported that AI was relevant and impressive tools that should be employed in the clinical decision making for patients.

ANN, artificial neural network; LSS, lumbar spinal stenosis; AUC, area under the curve; LR, logistic regression; LDH, lumbar disk herniation; H-L statistic, Hosmer-Lemeshow statistic; LSCS, lumbar spinal canal stenosis; ML, machine learning; SEA, spinal epidural abscess; NR, not reported; DNN, deep neural network; AI, artificial intelligence.

Table 4. A list of papers on ANN used in the biomechanical assessments of spinal disease

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Mann et al. [73]	1993	USA	The more samples, the more accurate	The more samples, the more accurate	MLP	Lumbar spine disorder	NR	To determine the reliability of the patient pain drawing when diagnosing low-back disorders and to delineate the pain mark patterns particular to each disorder by comparing physicians with computerized methods	The physicians averaged 51% accuracy with individual preferences for certain disorder groups. The computerized methods demonstrated comparable accuracy (48%) and more agreement in classification. ANNs was useful to clinicians for making accurate predictions of diagnosis from pain drawings.
Ongphiphadhanakul et al. [74]	1997	Thailand	100	29	MLP	Low BMD	NR	To evaluate the risk factors associated with low BMD and assess the prediction of low BMD using an ANN compared to a LR	There was no significant difference in terms of accuracy, sensitivity, and specificity in the prediction of low BMD at the lumbar spine or the femoral neck between ANN model and LR model. Results showed that ANN did not perform better than convention statistical methods in the prediction of low BMD.
Nussbaum et al. [75]	1997	USA	The more samples, the more accurate	The more samples, the more accurate	MLP	Lumbar muscle recruitment during static loading	NR	To examine inter-individual differences in the patterns of torso muscle recruitment during 3D static moment loading of the lumbar spine.	It was speculated that inter individual muscle recruitment differences may be important for assessing individual musculoskeletal risk.
Wang et al. [76]	2002	USA	The EMG signals of 10 flexor and extensor muscles	The EMG signals of 10 flexor and extensor muscles	MLP	Joint moments	NR	To determine muscle activations from EMG signals.	The results showed that the neural network model can be used to represent the relationship between EMG signals and joint moments well.
Arjmand et al. [77]	2013	Iran	5,220 Load positions and the more samples, the more accurate	The more samples, the more accurate	Five-layer, feed-forward neural network model	Spinal loads and muscle forces	Yes	Two ANNs was constructed, trained, and tested to map inputs of a complex trunk FE model to its outputs for spinal loads and muscle forces and compared to regression equations.	Results indicated that the ANNs were more accurate in mapping input-output relationships of the FE model (RMSE=20.7 N for spinal loads and RMSE=4.7 N for muscle forces) as compared to regression equations (RMSE=120.4 N for spinal loads and RMSE=43.2 N for muscle forces). Using these user-friendly tools, spine loads and trunk muscle can be easily estimated.

(Continued on next page)

Table 4. Continued

Author	Year	Country	No. of sample size: Training	No. of sample size: Testing	Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
Amaritsakul et al. [78]	2013	Taiwan	25 Screw designs were used as the learning set.	10 Randomly selected screw designs	MLP: a three-layered ANN	Optimization design of spinal pedicle screws	Yes	Using the 3D FE analytical results based on an L25 orthogonal array, bending and pullout objective functions were developed by an ANN algorithm, and the trade-off solutions known as Pareto optima were explored by a GA.	Multi-objective optimization study of spinal pedicle screws using the hybrid of ANN and GA could achieve an ideal with high bending and pullout performances simultaneously.
Hu et al. [79]	2018	USA	44 Chronic LBP and healthy individuals	44 Chronic LBP and healthy individuals	Deep neural networks	LBP	NR	To recognize LBP patients from healthy population performed static standing tasks, while their spine kinematics and center of pressure were recorded.	Results indicated that deep neural networks could recognize LBP populations with precision up to 97.2%. Results showed a deep learning network can solve the above classification problem with both promising precision and recall performance.
Aghazadeh et al. [80]	2019	Iran	15 Individuals each performed 135 load-handling activities	15 Individuals each performed 135 load-handling activities	Coupled ANNs	The risk of spine injury during manual material handling	NR	To estimate 3D whole-body posture, lumbosacral moments, and spinal loads during load-handling activities	The results showed outputs of the coupled ANNs for L4–L5 IDPs during a number of activities were in agreement with measured IDPs. Hence, coupled ANNs were found to be robust tools to evaluate posture, lumbosacral moments, spinal loads, and thus risk of injury during load-handling activities.
Liu et al. [81]	2019	China	The model was trained for 20 epochs with a mini-batch size of 16.	The model was trained for 20 epochs with a mini-batch size of 16.	CSNN	Tracking the motion of the lumbar spine	NR	To automatically track lumbar vertebras with rotated bounding boxes in digitalized video fluoroscopic imaging, sequences.	Results indicated that the proposed tracking method can track the lumbar vertebra steadily and robustly. The study demonstrated that the lumbar tracker based on CSNN can be trained successfully without annotated lumbar sequences.

(Continued on next page)

Table 4. Continued

Author	Year	Country	No. of sample size		Model type	Conditions	Comparison with non-ANN models	Main focus	Results/conclusion(s)
			Training	Testing					
Zhang et al. [70]	2019	China	80 Subjects with QCT data of lumbar spine were randomly selected	80 Subjects with QCT data of lumbar spine were randomly selected	Machine learning models	To predict vertebral strength	NR	The parameters extracted from QCT images were used to predict vertebral strength through machine learning models.	The 58 parameters were simplified to five features and nine PCs. High accuracy was achieved by using the five features or the nine PCs to predict vertebral strength. This study provided an effective approach to predict vertebral strength and showed that it may have great potential in clinical applications for noninvasive assessment of vertebral fracture risk.

ANN, artificial neural network; MLP, multilayer perceptron neural networks; NR, not reported; BMD, bone mineral density; LR, logistic regression; 3D, three-dimensional; EMG, electromyogram; FE, finite element; RMSE, root mean square error; GA, genetic algorithm; LBP, low back pain; IDP, intradiscal pressure; CSNN, convolutional siamese neural network; QCT, quantitative computed tomography; PCs, principal components.

introduced by Jamaludin et al. [14,89] helps in clinical and algorithmic research [90]. This system can extract a wide range of relevant measurements from magnetic resonance (MR) images automatically including Pffirrmann grades, Modic changes, spinal stenosis, grading LSS into four grades, and disc herniation. The software needs to be able to learn without human input and classify multiple radiological features simultaneously. Therefore, the SpineNet software system adopts a neural network which allows it to learn and classify multiple scores at the same time. The SpineNet software tool is available online at http://zeus.robots.ox.ac.uk/spinenet/. Meanwhile, the authors stated that this is neither a diagnostics tool nor a medical device; hence, it should be used for research only [90]. In the future, ANN tools for the spinal workflow may help to empower the spine surgeon in management of spinal diseases, positively engage the patients, reduce the probability of error and improve spine surgeon efficiency.

ANNs have been effectively applied for biomechanics of spine, such as estimation of loads and stresses [80], estimation of the material properties of biological tissues [74], and analysis of the motion and gait [81]. A clear understanding of biomechanical principles is vital for the treatment of spinal disorders [91]. In view of the direct and invasive *in vivo* measurements of spinal loads and muscle forces, researchers have decided to apply ANNs for spinal biomechanical assessment by using different possible models [77]. Yet, the use of artificial intelligence techniques in clinical biomechanics of the spine is still in its infancy. However, continued assessment using ANN methods and understanding of mathematical and bioengineering principles will ensure the discovery of new solutions and more effective ways of helping spine patients, especially spine surgery [75-78]. In general, at the present time, there is no clinical application available for biomechanics of spine. Nonetheless, the aforementioned studies clearly demonstrate the potential of neural network in this arena.

There was considerable heterogeneity in the modeling methods used. Studies varied with regard to the inclusion criteria, input and output variables used, machine learning techniques applied, and models performance for the evaluation of the four categories of issues discussed above. Yet, LR, neural networks, and support vector machines (SVM) are the most commonly applied computer-based algorithms. However, these algorithms are developing and require further improvement. As an example, the following machine learning techniques are the most recommended to be employed: Bi-directional Long Short-Term Memory, Deep Learning and Neural Network, Extreme Machine Learning, SVM and TBasts, Random Forest, and Boosting [92]. On the other hand, the comparison of different neural network algorithms could assist in the selection of the best model for making sure that it has the best performance. However, it is also crucial to consider the advantages and disadvantages of these algorithms based on the dataset details and planed application to make the best choice [92]. Hence, there is a need for method standardization in order to apply ANNs spine clinical practice.Although the current research highlights the promise and potential of neural networks in spine disorders, there is continuing debate whether neural networks could or should be used in routine clinical practice. At present, there is a paucity of literature on the clinical application of neural networks in spine clinical decision-making. In addition, there are no specific consensus recommendations for designing an optimal ANN for clinical practice [2]. Hence, in order to develop an ANN tool for spine clinical practice, it is recommended that an international center need to consider the following subjects: (1) to create a multidisciplinary team of spine clinicians, engineers and data scientists to evaluate ANN tools; (2) to design methods of data collection for training and testing, and standardization for each spinal disorders; (3) to select the best neural network algorithm; (4) to develop and improve user-friendly software environment; (5) to validate ("model testing"); (6) to assess ANNs as tools for performing RCT; and (7) to provide up-to-date ANN models which could handle higher patient data (big data) entry for augmenting augment decision-making efficacy [2]. In general, the preliminary results from tools are satisfactory. However, it still requires many efforts to make the ANN tools more accurate and to make the idea practically achievable.

ANN technology has been attracting substantial attention in spinal disease, but there are challenges to be implemented in clinical setting. Several limitations exist in most of the studies on ANNs in spinal diseases, including largely heterogeneous study design, data analysis, modeling technique, training and testing features applied, algorithms employed, and end points. Hence, a focused synthesis of the literature cannot be provided. In addition, the search strategy was limited to the keywords in the titles or abstracts of publications. Thus, we might have missed some papers. Secondly, this work restricted the query search for articles in PubMed. Thirdly, non-English publications were not considered in this study. We believe that the research regarding the application of ANNs in spinal diseases have also been published in other languages. Fourthly, most ANNs algorithms in these studies were validated with their dataset, it may be lack of external validation and generalizability of their results. Fifthly, some studies did not compare ANNs with conventional statistical analysis; hence, a comparison between any two models is limited [2].

Conclusions

The evidence suggests that ANNs can be successfully used for spinal disease to manage its diagnosis, prognosis and outcome prediction. Further ANNs algorithm retraining, generalizability of models, data standardization in neural networks, and focus on the application of ANNs as a tool in clinical spine practice, will augment decision-making efficacy.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

The authors thank the staff of the Neurosurgery Unit at Imam-Hossain Hospital, Tehran, Iran.

References

1. Azimi P, Mohammadi HR. Predicting endoscopic third ventriculostomy success in childhood hydrocephalus: an artificial neural network analysis. J Neurosurg Pediatr 2014;13:426-32.
2. Azimi P, Mohammadi HR, Benzel EC, Shahzadi S, Azhari S, Montazeri A. Artificial neural networks in neurosurgery. J Neurol Neurosurg Psychiatry 2015;86:251-6.
3. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR. Use of artificial neural networks to predict surgical satisfaction in patients with lumbar spinal

canal stenosis: clinical article. J Neurosurg Spine 2014;20:300-5.

4. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR. The prediction of successful surgery outcome in lumbar disc herniation based on artificial neural networks. J Neurosurg Sci 2016;60:173-7.
5. Nolting J. Developing a neural network model for health care. AMIA Annu Symp Proc 2006;2006:1049.
6. Bishop JB, Szpalski M, Ananthraman SK, McIntyre DR, Pope MH. Classification of low back pain from dynamic motion characteristics using an artificial neural network. Spine (Phila Pa 1976) 1997;22:2991-8.
7. Jaremko JL, Poncet P, Ronsky J, et al. Estimation of spinal deformity in scoliosis from torso surface cross sections. Spine (Phila Pa 1976) 2001;26:1583-91.
8. Stanley RJ, Long R. A radius of curvature-based approach to cervical spine vertebra image analysis. Biomed Sci Instrum 2001;37:385-90.
9. Liszka-Hackzell JJ, Martin DP. Categorization and analysis of pain and activity in patients with low back pain using a neural network technique. J Med Syst 2002;26:337-47.
10. Lin H. Identification of spinal deformity classification with total curvature analysis and artificial neural network. IEEE Trans Biomed Eng 2008;55:376-82.
11. Sari M, Gulbandilar E, Cimbiz A. Prediction of low back pain with two expert systems. J Med Syst 2012;36:1523-7.
12. Veronezi CC, de Azevedo Simoes PW, Dos Santos RL, et al. Computational analysis based on artificial neural networks for aiding in diagnosing osteoarthritis of the lumbar spine. Rev Bras Ortop 2015;46:195-9.
13. Zhang J, Li H, Lv L, Zhang Y. Computer-aided cobb measurement based on automatic detection of vertebral slopes using deep neural network. Int J Biomed Imaging 2017;2017:9083916.
14. Jamaludin A, Lootus M, Kadir T, et al. Issls Prize in Bioengineering Science 2017: automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. Eur Spine J 2017;26:1374-83.
15. Wang J, Fang Z, Lang N, Yuan H, Su MY, Baldi P. A multi-resolution approach for spinal metastasis detection using deep Siamese neural networks. Comput Biol Med 2017;84:137-46.
16. 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.
17. Kim S, Bae WC, Masuda K, Chung CB, Hwang D. Fine-grain segmentation of the intervertebral discs from MR spine images using deep convolutional neural networks: BSU-Net. Appl Sci (Basel) 2018;8:1656.
18. Kim K, Kim S, Lee YH, Lee SH, Lee HS, Kim S. 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.
19. Han Z, Wei B, Mercado A, Leung S, Li S. Spine-GAN: semantic segmentation of multiple spinal structures. Med Image Anal 2018;50:23-35.
20. Chmelik J, Jakubicek R, Walek P, et al. Deep convolutional neural network-based segmentation and classification of difficult to define metastatic spinal lesions in 3D CT data. Med Image Anal 2018;49:76-88.
21. Liao H, Mesfin A, Luo J. Joint vertebrae identification and localization in spinal CT images by combining short- and long-range contextual information. IEEE Trans Med Imaging 2018;37:1266-75.
22. Al Arif SMMR, Knapp K, Slabaugh G. Fully automatic cervical vertebrae segmentation framework for X-ray images. Comput Methods Programs Biomed 2018;157:95-111.
23. Han Z, Wei B, Leung S, Nachum IB, Laidley D, Li S. Automated pathogenesis-based diagnosis of lumbar neural foraminal stenosis via deep multiscale multitask learning. Neuroinformatics 2018;16:325-37.
24. Li X, Dou Q, Chen H, et al. 3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images. Med Image Anal 2018;45:41-54.
25. Zhou Y, Liu Y, Chen Q, Gu G, Sui X. Automatic lumbar MRI detection and identification based on deep learning. J Digit Imaging 2019;32:513-20.
26. Wang X, Zhai S, Niu Y. Automatic vertebrae localization and identification by combining deep SSAE contextual features and structured regression forest. J Digit Imaging 2019;32:336-48.
27. Lessmann N, van Ginneken B, de Jong PA, Isgum

I. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. Med Image Anal 2019;53:142-55.
28. 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.
29. 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.
30. Hopkins BS, Weber KA 2nd, Kesavabhotla K, Paliwal M, Cantrell DR, Smith ZA. Machine learning for the prediction of cervical spondylotic myelopathy: a post hoc pilot study of 28 participants. World Neurosurg 2019;127:e436-42.
31. Horng MH, Kuok CP, Fu MJ, Lin CJ, Sun YN. Cobb angle measurement of spine from X-ray images using convolutional neural network. Comput Math Methods Med 2019;2019:6357171.
32. Pang S, Su Z, Leung S, et al. Direct automated quantitative measurement of spine by cascade amplifier regression network with manifold regularization. Med Image Anal 2019;55:103-15.
33. Li H, Luo H, Liu Y. Paraspinal muscle segmentation based on deep neural network. Sensors (Basel) 2019;19:2650.
34. Chen Y, Gao Y, Li K, Zhao L, Zhao J. vertebrae identification and localization utilizing fully convolutional networks and a hidden Markov model. IEEE Trans Med Imaging 2020;39:387-99.
35. Rak M, Steffen J, Meyer A, Hansen C, Tonnies KD. Combining convolutional neural networks and star convex cuts for fast whole spine vertebra segmentation in MRI. Comput Methods Programs Biomed 2019;177:47-56.
36. Pan Y, Chen Q, Chen T, et al. Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays. Eur Spine J 2019;28:3035-43.
37. Weng CH, Wang CL, Huang YJ, et al. Artificial intelligence for automatic measurement of sagittal vertical axis using ResUNet framework. J Clin Med 2019;8:1826.
38. 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;20:590-9.
39. Jakubicek R, Chmelik J, Ourednicek P, Jan J. Deep-learning-based fully automatic spine centerline detection in CT data. Conf Proc IEEE Eng Med Biol Soc 2019;2019:2407-10.
40. Lyu J, Ling SH, Banerjee S, et al. 3D ultrasound spine image selection using convolution learning-to-rank algorithm. Conf Proc IEEE Eng Med Biol Soc 2019;2019:4799-802.
41. Watanabe K, Aoki Y, Matsumoto M. An application of artificial intelligence to diagnostic imaging of spine disease: estimating spinal alignment from moire images. Neurospine 2019;16:697-702.
42. Kok H, Acilar AM, Izgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. Prog Orthod 2019;20:41.
43. Iriondo C, Pedoia V, Majumdar S. Lumbar intervertebral disc characterization through quantitative MRI analysis: an automatic voxel-based relaxometry approach. Magn Reson Med 2020 Feb 14 [Epub]. https://doi.org/10.1002/mrm.28210.
44. Won D, Lee HJ, Lee SJ, Park SH. Spinal stenosis grading in magnetic resonance imaging using deep convolutional neural networks. Spine (Phila Pa 1976) 2020 Jan 20 [Epub]. https://doi.org/10.1097/BRS.0000000000003384.
45. Lee S, Choe EK, Kang HY, Yoon JW, Kim HS. The exploration of feature extraction and machine learning for predicting bone density from simple spine X-ray images in a Korean population. Skeletal Radiol 2020;49:613-8.
46. Lee JH, Han IH, Kim DH, et al. Spine computed tomography to magnetic resonance image synthesis using generative adversarial networks: a preliminary study. J Korean Neurosurg Soc 2020 Jan 14 [Epub]. https://doi.org/10.3340/jkns.2019.0084.
47. Bae HJ, Hyun H, Byeon Y, et al. Fully automated 3D segmentation and separation of multiple cervical vertebrae in CT images using a 2D convolutional neural network. Comput Methods Programs Biomed 2020;184:105119.
48. Jakubicek R, Chmelik J, Jan J, Ourednicek P, Lambert L, Gavelli G. Learning-based vertebra localization and labeling in 3D CT data of possibly incomplete and pathological spines. Comput Methods Programs Biomed 2020;183:105081.

49. Kim YJ, Ganbold B, Kim KG. Web-based spine segmentation using deep learning in computed tomography images. Healthc Inform Res 2020;26:61-7.
50. Rehman F, Ali Shah SI, Riaz MN, Gilani SO, R F. A region-based deep level set formulation for vertebral bone segmentation of osteoporotic fractures. J Digit Imaging 2020;33:191-203.
51. Dickey JP, Pierrynowski MR, Bednar DA, Yang SX. Relationship between pain and vertebral motion in chronic low-back pain subjects. Clin Biomech (Bristol, Avon) 2002;17:345-52.
52. Lin CS, Chiu JS, Hsieh MH, Mok MS, Li YC, Chiu HW. Predicting hypotensive episodes during spinal anesthesia with the application of artificial neural networks. Comput Methods Programs Biomed 2008;92:193-7.
53. Parsaeian M, Mohammad K, Mahmoudi M, Zeraati H. Comparison of logistic regression and artificial neural network in low back pain prediction: second national health survey. Iran J Public Health 2012;41:86-92.
54. Papic M, Brdar S, Papic V, Loncar-Turukalo T. Return to work after lumbar microdiscectomy: personalizing approach through predictive modeling. Stud Health Technol Inform 2016;224:181-3.
55. Karhade AV, Ogink P, Thio Q, et al. Development of machine learning algorithms for prediction of discharge disposition after elective inpatient surgery for lumbar degenerative disc disorders. Neurosurg Focus 2018;45:E6.
56. Arvind V, Kim JS, Oermann EK, Kaji D, Cho SK. Predicting surgical complications in adult patients undergoing anterior cervical discectomy and fusion using machine learning. Neurospine 2018;15:329-37.
57. Han SS, Azad TD, Suarez PA, Ratliff JK. A machine learning approach for predictive models of adverse events following spine surgery. Spine J 2019;19:1772-81.
58. DeVries Z, Hoda M, Rivers CS, et al. Development of an unsupervised machine learning algorithm for the prognostication of walking ability in spinal cord injury patients. Spine J 2020;20:213-24.
59. Staartjes VE, de Wispelaere MP, Vandertop WP, Schroder ML. Deep learning-based preoperative predictive analytics for patient-reported outcomes following lumbar discectomy: feasibility of center-specific modeling. Spine J 2019;19:853-61.
60. Karhade AV, Thio QCBS, Ogink PT, et al. Development of machine learning algorithms for prediction of 30-day mortality after surgery for spinal metastasis. Neurosurgery 2019;85:E83-91.
61. Karhade AV, Thio QC, Ogink PT, et al. Predicting 90-day and 1-year mortality in spinal metastatic disease: development and internal validation. Neurosurgery 2019;85:E671-81.
62. Weber KA, Smith AC, Wasielewski M, et al. Deep learning convolutional neural networks for the automatic quantification of muscle fat infiltration following whiplash injury. Sci Rep 2019;9:7973.
63. Goyal A, Ngufor C, Kerezoudis P, McCutcheon B, Storlie C, Bydon M. 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 2019:1-11.
64. Karhade AV, Ogink PT, Thio QCBS, et al. Development of machine learning algorithms for prediction of prolonged opioid prescription after surgery for lumbar disc herniation. Spine J 2019;19:1764-71.
65. Ryu SM, Seo SW, Lee SH. Novel prognostication of patients with spinal and pelvic chondrosarcoma using deep survival neural networks. BMC Med Inform Decis Mak 2020;20:3.
66. Azimi P, Mohammadi HR, Benzel EC, Shahzadi S, Azhari S. Use of artificial neural networks to predict recurrent lumbar disk herniation. J Spinal Disord Tech 2015;28:E161-5.
67. Azimi P, Mohammadi HR, Benzel EC, Shahzadi S, Azhari S. Use of artificial neural networks to decision making in patients with lumbar spinal canal stenosis. J Neurosurg Sci 2017;61:603-11.
68. Karhade AV, Shah AA, Bono CM, et al. Development of machine learning algorithms for prediction of mortality in spinal epidural abscess. Spine J 2019;19:1950-9.
69. Stopa BM, Robertson FC, Karhade AV, et al. Predicting nonroutine discharge after elective spine surgery: external validation of machine learning algorithms. J Neurosurg Spine 2019:1-6.
70. Zhang M, Gong H, Zhang K, Zhang M. Prediction of lumbar vertebral strength of elderly men based on quantitative computed tomography images using machine learning. Osteoporos Int 2019;30:2271-82.
71. Hopkins BS, Yamaguchi JT, Garcia R, et al. Using ma-

chine learning to predict 30-day readmissions after posterior lumbar fusion: an NSQIP study involving 23,264 patients. J Neurosurg Spine 2019:1-8.

72. Hopkins BS, Mazmudar A, Driscoll C, et al. Using artificial intelligence (AI) to predict postoperative surgical site infection: a retrospective cohort of 4046 posterior spinal fusions. Clin Neurol Neurosurg 2020;192:105718.
73. Mann NH 3rd, Brown MD, Hertz DB, Enger I, Tompkins J. Initial-impression diagnosis using low-back pain patient pain drawings. Spine (Phila Pa 1976) 1993;18:41-53.
74. Ongphiphadhanakul B, Rajatanavin R, Chailurkit L, et al. Prediction of low bone mineral density in postmenopausal women by artificial neural network model compared to logistic regression model. J Med Assoc Thai 1997;80:508-15.
75. Nussbaum MA, Chaffin DB. Pattern classification reveals intersubject group differences in lumbar muscle recruitment during static loading. Clin Biomech (Bristol, Avon) 1997;12:97-106.
76. Wang L, Buchanan TS. Prediction of joint moments using a neural network model of muscle activations from EMG signals. IEEE Trans Neural Syst Rehabil Eng 2002;10:30-7.
77. Arjmand N, Ekrami O, Shirazi-Adl A, Plamondon A, Parnianpour M. Relative performances of artificial neural network and regression mapping tools in evaluation of spinal loads and muscle forces during static lifting. J Biomech 2013;46:1454-62.
78. Amaritsakul Y, Chao CK, Lin J. Multiobjective optimization design of spinal pedicle screws using neural networks and genetic algorithm: mathematical models and mechanical validation. Comput Math Methods Med 2013;2013:462875.
79. Hu B, Kim C, Ning X, Xu X. Using a deep learning network to recognise low back pain in static standing. Ergonomics 2018;61:1374-81.
80. Aghazadeh F, Arjmand N, Nasrabadi AM. Coupled artificial neural networks to estimate 3D whole-body posture, lumbosacral moments, and spinal loads during load-handling activities. J Biomech 2020;102:109332.
81. Liu Y, Sui X, Liu C, Kuang X, Hu Y. Automatic lumbar spine tracking based on Siamese convolutional network. J Digit Imaging 2020;33:423-30.
82. Khor S, Lavallee D, Cizik AM, et al. Development and validation of a prediction model for pain and functional outcomes after lumbar spine surgery. JAMA Surg 2018;153:634-42.
83. Comparative Effectiveness Research Translation Network. Spine lumbar fusion outcomes calculator [Internet]. Seattle (WA): Comparative Effectiveness Research Translation Network; c2017 [cited 2020 Mar 20]. Available from: https://becertain.shinyapps.io/lumbar_fusion_calculator/.
84. Skeletal Oncology Research Group. Lumbar disc disorder prolonged opioid prescription [Internet]. Boston (MA): Skeletal Oncology Research Group [cited 2020 Mar 20]. Available from: https://sorg-apps.shinyapps.io/lumbardiscopioid/.
85. Skeletal Oncology Research Group. Spinal metastatic disease survival predictions [Internet]. Boston (MA): Skeletal Oncology Research Group [cited 2020 Mar 20]. Available from: https://sorg-apps.shinyapps.io/spinemetssurvival/.
86. Skeletal Oncology Research Group. Discharge disposition: elective inpatient lumbar degenerative disc disorders [Internet]. Boston (MA): Skeletal Oncology Research Group [cited 2020 Mar 20]. Available from: https://sorg-apps.shinyapps.io/discdisposition/.
87. Skeletal Oncology Research Group. Thirty-day mortality spinal metastatic disease [Internet]. Boston (MA): Skeletal Oncology Research Group [cited 2020 Mar 20]. Available from: https://sorg-apps.shinyapps.io/spinemets/.88. Jiang J, Trundle P, Ren J. Medical image analysis with artificial neural networks. Comput Med Imaging Graph 2010;34:617-31.
89. Jamaludin A, Kadir T, Zisserman A. SpineNet: automated classification and evidence visualization in spinal MRIs. Med Image Anal 2017;41:63-73.
90. Visual Geometry Group, University of Oxford. SpineNet online demo [Internet]. Oxford: Visual Geometry Group, University of Oxford [cited 2020 Mar 20]. Available from: http://zeus.robots.ox.ac.uk/spinenet/.
91. White AA III, Panjabi MM. Clinical biomechanics of the spine. 2nd ed. Baltimore (MD): Lippincott Williams & Wilkins; 1990.
92. Barone AV, Helcl J, Sennrich R, Haddow B, Birch A. Deep architectures for neural machine translation [Internet]. Ithaca (NY): arXiv, Cornell University; 2017 [cited 2020 Mar 20]. Available from: https://arxiv.org/pdf/1707.07631.pdf.