Li and Wong *BioMedical Engineering OnLine* (2024) 23:80
https://doi.org/10.1186/s12938-024-01272-6

BioMedical Engineering OnLine

REVIEW

Open Access

The application of machine learning methods for predicting the progression of adolescent idiopathic scoliosis: a systematic review

Lening Li[1*] and Man-Sang Wong[1]

*Correspondence: lynn2018.li@connect.polyu.hk

[1] Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong, China

Abstract

Predicting curve progression during the initial visit is pivotal in the disease management of patients with adolescent idiopathic scoliosis (AIS)—identifying patients at high risk of progression is essential for timely and proactive interventions. Both radiological and clinical factors have been investigated as predictors of curve progression. With the evolution of machine learning technologies, the integration of multidimensional information now enables precise predictions of curve progression. This review focuses on the application of machine learning methods to predict AIS curve progression, analyzing 15 selected studies that utilize various machine learning models and the risk factors employed for predictions. Key findings indicate that machine learning models can provide higher precision in predictions compared to traditional methods, and their implementation could lead to more personalized patient management. However, due to the model interpretability and data complexity, more comprehensive and multicenter studies are needed to transition from research to clinical practice.

Keywords: Adolescent idiopathic scoliosis, Machine learning, Prediction

Introduction

Adolescent idiopathic scoliosis (AIS) manifests as a three-dimensional spinal deformity in adolescence [1]. It is widely accepted that scoliotic curvature progression in AIS patients results from the interplay between genetic predispositions and biomechanical factors. The initial spinal curvature—whose origins are not fully understood—leads to asymmetric loading on the vertebral growth plates, resulting in varied growth rates through endochondral ossification. This exacerbates lateral spinal deformity and triggers axial vertebral rotation, aligning with the Hueter–Volkmann law, which suggests that scoliosis progression is cyclical [2].

Clinically, AIS management includes regular monitoring for curves less than 25 degrees in patients with a Risser grade of 0 to 2, bracing for curves of 25 to 40 degrees, and surgical evaluation for curves exceeding 45 or 50 degrees [3]. Despite these guidelines, clinicians must develop interventions for each patient's unique conditions. In addition, based on individual variations among patients, doctors also need to predict whether patients will experience rapid progression of scoliosis in the future, necessitating more frequent clinical observation.

The progression of scoliosis is not determined by a single factor but results from the interaction of multiple factors, such as age, sex, skeletal maturity and spine curve morphology. A multifactorial analysis system can enhance the accuracy of predictions [4]. Recent studies have utilized advanced computational methods to predict the progression of AIS and assess the effectiveness of treatment strategies. In the field of AIS management, the integration of machine learning technologies has become a pivotal trend.

The purpose of this study was to conduct a systematic review of the literature on the application of AI methods for multiple risk factors to aid clinicians in the management of patients with AIS.

Methods

Search strategy

The literature search was conducted on studies published between January 1, 2016, and January 25, 2024. The initial search was carried out on January 31, 2024, with a subsequent search performed on February 28, 2024. The objective of this systematic investigation was to analyze the impact of various risk factors on spinal curvature in patients with AIS. The PubMed, Web of Science, and Google Scholar databases were utilized for the literature search. The keywords used in the search were "adolescent idiopathic scoliosis", "predict*", "deep learning" and "machine learning" in either the title or the abstract of the publications. Additionally, we supplemented the literature information not covered by the search by reading the references cited in the articles. The highlights of the analyzed studies are presented in Table 1.

Eligibility criteria

Two reviewers independently selected the literature, resolving discrepancies by consensus. The included studies were published in English, peer-reviewed, pertained to predictive factors in AIS, and focused on nongenetic scoliosis types. The exclusion criteria included reviews, animal studies, correspondences, and editorials.

Data analysis

The analysis involved extracting data on study design, publication year, author names, patient demographics, machine learning models, definitions of scoliosis progression, identified risk factors and prediction results. This review primarily analyzes the following factors: 1. the machine learning models used were as follows: 2. the input risk factors. We focused exclusively on curve prediction for nonsurgical patients with AIS, excluding articles predicting the outcomes of AIS surgeries.

Results

The preliminary search across two databases yielded 417 publications potentially meeting the inclusion criteria. Subsequent screening of titles and abstracts narrowed the field to 93 studies for detailed analysis. After the full-text examination, 15 papers were

Table 1 Highlight of analyzed studies

Author	Study design	Prediction input	Number of patients	Machine learning model	Progression stander	Risk factors	Prediction result
Wang et al. [9]	Retrospective study	Demographic factors and X-rays in the spine and hand	810	Attention-based capsule neural network and DCNN	Cobb angle increase in the major curve of ≥ 6° between the first visit and skeletal maturity in curves that exceeded 25°	1. Patient demographics; 2. Vertebral morphology; 3. Skeletal maturity; 4. Bone quality	Model accuracy was 83.2% and AUC was 0.84
Wang et al. [12]	Retrospective study	Standing posteroanterior X-rays	490	Deep capsule network with self-attention routing	Cobb angle increase in the major curve of ≥ 6° between the first visit and skeletal maturity in curves that exceeded 25°	1. Radiomics from ROI (apical vertebrae or disc of the major curve, together with at least two adjacent vertebras above and below as well as the lateral rib articulations)	Model accuracy was 77.1% and AUC was 0.74
Yahara et al. [10]	Retrospective study	The frontal view of the total spine radiographs	58	DCNN	Cobb angle increased more than 10° within 2 years	1. Three ROI of frontal view X-rays	Model accuracy was 69%
Kadoury et al. [21]	Retrospective study	Biplanar X-rays	133	Discriminative probabilistic manifold	Difference of over 6° between the first and last visits	Geometric features from 3D models, anatomical landmarks, Intervertebral parameters, and skeletal properties	The prediction differences of 2.1° in main curve angulation
Alfraihat et al. [6]	Retrospective study	Standing and side-bending spinal radiographs, including the pelvis	193	Random forest	Cobb angle difference of 6° or more between the first and the last visit	1. Initial major Cobb angle; 2. Patient flexibility; 3. Initial lumbar lordosis angle; 4. Initial thoracic kyphosis angle; 5. Age at the last visit; 6. The number of spinal levels involved, 7. The Risser "+" stage at the initial consultation	MAE of Cobb angle between prediction and truth was 4.64
Deng et al. [31]	Retrospective study	Demographic factors and X-ray	341	Support vector machine	Future Cobb angle	1. Clinical indicators; 2. Brace usage; 3. Patient demographics; 4. Baseline clinical measurements	RMSE was 5.181

Table 1 (continued)

Author	Study design	Prediction input	Number of patients	Machine learning model	Progression stander	Risk factors	Prediction result
Chu et al. [13]	Retrospective study	Demographic factors and posteroanterior X-ray images	463	Capsule network	Cobb angle increase > 5° in 3-month follow-up	1. Sex; 2. Age; 3. Weight; 4. Sitting height; 5. Standing height; 6. Arm span; 7. Risser sign; 8. Distal radius; 9. Ulna classification; 10. Posteroanterior radiographs; 11. Bracing compliance	Accuracy of 73.9%
Guo et al. [15]	Retrospective study	Demographic factors and X-ray	1655	RNN with LSTM cells	Future Cobb angle	1. Current Cobb angle; 2. Future Cobb angle; 3. Current age; 4. Time Span; 5. Current Brace; Future Brace; 7. Change brace	RMSE was 1.229
García-Cano et al. [7]	Retrospective study	Stereoradiographic 3D reconstructions from conventional X-rays	150	Random Forest	Difference of over 6° between the first and last visits	1. 9 ICs from the 3D variability of the shapes in the posteroanterior, sagittal and apical planes	Difference between prediction and real was 1.83, 5.18, and 4.79° of Cobb angles in the proximal thoracic, main thoracic, and thoracolumbar lumbar sections, respectively
Patel et al. [25]	Retrospective study	Surface topography	38	Proportional odds logistic modeling	Cobb angle increase > 6°	1. Surface topography; 2. Age; 3. Gender; 4. scoliotic angle	Accuracy was 71%
Hong et al. [18]	Retrospective study	Surface topography	45	Decision trees	Cobb angle increase > 5°	1. Surface topography	Sensitivity and specificity were 73% and 44%
Ghaneei et al. [19]	Retrospective study	Surface topography	128	Customized k-Nearest Neighbor	Cobb angle increase > 5°	1. Surface topography	Accuracy was 93%
Zhang et al. [20]	Prospective study	Smartphone photographs of patients' backs	1780	CNNs with Attention Mechanisms	Cobb angle increase > 5° in 6-month follow-up	1. Smartphone photographs of patients' backs	AUC was 0.757

Table 1 (continued)

Author	Study design	Prediction input	Number of patients	Machine learning model	Progression stander	Risk factors	Prediction result
Lv et al. [17]	Retrospective study	Visual inspection, Adam FBT and measurement of the angle of trunk rotation	3313	Artificial Neural Network Model	Occurrence of AIS	1. The ratio of sitting height to standing height; 2. Angle of lumbar rotation; 3. Scapular tilt; 4. Shoulder-height difference; 5. Lumbar concave; 6. Pelvic tilt;	AUC was 0.899
Yan et al. [27]	Retrospective study	Visual inspection, Adam's FBT, and measurement of the angle of trunk rotation	1779	Logistic Regression models	Occurrence of AIS	1. Angle of thoracic rotation; 2. Angle of thoracolumbar rotation; 3. Angle of lumbar rotation; 4. Scapular tilt 5. Shoulder-height difference; 6. Lumbar concave; 7. Pelvic tilt	Accuracy was 83.3%

included in the study, 14 of which were retrospective and 1 of which was prospective. The sample size ranged widely, from 38 to 1780. The papers included a total of 11,376 patients, with female patients accounting for 66.46%. Most of these patients underwent X-ray examinations, but some received radiation-free imaging or basic physical examinations. The definition of scoliosis progression primarily involves an increase in the Cobb angle between two consecutive follow-up visits, but the extent of progression varies slightly across different studies (5°–10°).

The process of using machine learning to predict AIS progression involves several steps: collecting and preprocessing a comprehensive dataset that includes radiographic images, clinical parameters, and patient demographics; extracting relevant features such as geometric data from radiographs and clinical measurements; selecting suitable machine learning models; training the chosen models with the preprocessed data; validating and testing the models on separate datasets to ensure they generalize well; and finally, using the trained models to predict scoliosis progression in new patients, providing insights and recommendations for clinical management (Fig. 1).

Types of systems and models referenced

Random Forest

The Random Forest (RF) algorithm has emerged as a cornerstone ensemble learning technique [5] used for addressing a broad spectrum of classification and regression tasks. By arranging the collective power of numerous decision trees, RFs construct a model distinguished by its superior accuracy and resilience. Each tree within the forest is cultivated from unique random subsets and features of the dataset, effectively mitigating the risk of overfitting. RF is equally proficient in dealing with discrete and continuous data types, eliminating the need for dataset normalization, and can generate a proximity matrix that sheds light on sample similarities. However, RF may incur heightened computational complexity and cost, proportional to the number of trees and their depth.

RF models were constructed by constructing numerous decision trees during training and amalgamating the predictions of these trees to forecast the final major Cobb angle in AIS patients [6]. The Sequential Backward Floating Selection (SBFS) method was utilized to pinpoint the most predictive features for curve progression, streamline model complexity and eliminate less significant features. Another study [7] leveraged independent component analysis (ICA) to distill independent components (ICs) representing primary shape variations in 3D spine models, which served as inputs to train the RF model. The predictive modeling approach involved using the RF model to estimate future changes in the spine's shape based on its initial condition. It adopts a chain of predictor strategies in which the outcome of one prediction feeds into the next, thereby chronologically simulating scoliosis progression. This innovative technique of using sequential predictors to mirror time-evolving progression underscored the model's ability to depict the spine's morphological changes over successive time frames, capturing the dynamic features of scoliosis development.

Diffusion-convolutional neural networks (DCNNs)

Traditional Convolutional Neural Networks (CNNs) encounter substantial challenges when processing structured data. The pursuit of appropriate methods for representing

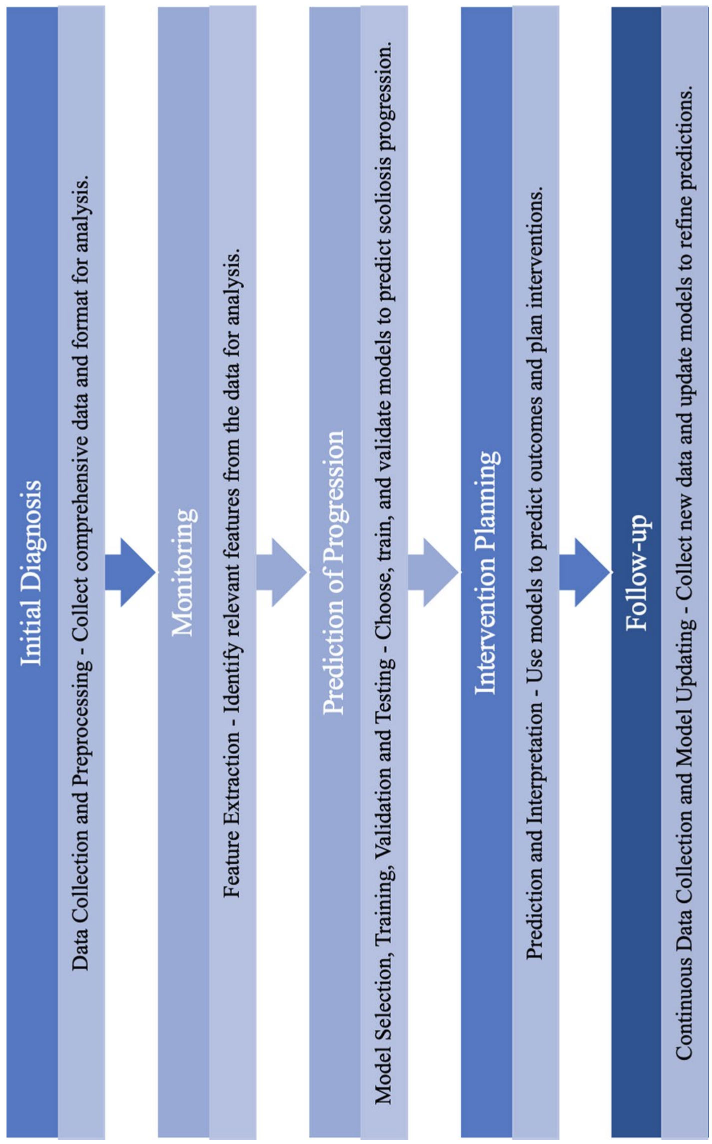

Fig. 1 Machine learning interventions in the progression and management of AIS

and exploring data structures is critical for enhancing prediction accuracy. Nonetheless, pinpointing these structures is often arduous, and incorporating them into models markedly escalates prediction complexity. The diffusion convolutional neural network (DCNN) addresses these issues by incorporating the diffusion-convolution operation within CNNs for use with graph-structured data (non-Euclidean space) [8], utilizing a universal framework that enhances predictive performance while reducing complexity. DCNNs stand out for their ability to offer a dynamic representation of graph data, capturing node features, edge features, and structural nuances with minimal preprocessing. In the realm of spinal X-ray analysis [9], DCNNs have the potential to transform imaging feature extraction for AIS prediction, leveraging deep learning to identify patterns associated with curve progression. However, the computational requirements for training and deploying DCNNs present considerable challenges, especially in resource-constrained settings. The success of DCNNs also greatly depends on the quantity and quality of the training data, with the possibility of reduced performance in instances of limited or unusual data. Additionally, research [10] has suggested that omitting key clinical parameters such as age, sex, growth rate, or markers of skeletal maturity could compromise the model's accuracy in predicting AIS progression.

Capsule neural network (CapsNet)

CapsNets are good at extracting direction-related features from X-rays, such as vertebral rotation and rib asymmetry, in addition to capturing general spatial features [11]. CNN sensitivity to local features, such as slight rotations or translations, can substantially influence the outputs. Furthermore, CNNs primarily concentrate on local features, potentially neglecting the integration of these features into coherent structures, thereby overcoming complex spatial hierarchies. In contrast, CapsNets excel in discerning complex hierarchies and spatial relationships, which are essential for numerous applications. The fundamental units of CapsNets, known as capsules, are designed to detect and encode intricate patterns and hierarchical information. Unlike CNNs, capsules produce higher-dimensional outputs and possess more elaborate internal structures, enabling a nuanced and comprehensive representation of the input data. Each capsule acts as a miniature neural network that identifies specific visual patterns, encoding probabilities and pose parameters of their presence, thus preserving extensive spatial hierarchical information.

A pivotal feature of CapsNets is the introduction of "dynamic routing", a sophisticated mechanism that facilitates information transfer among capsules. This process enhances the network's grasp of objects' internal structures and their relative spatial orientations. Compared to traditional CNN propagation techniques such as max pooling, dynamic routing offers enhanced flexibility and better information retention. Moreover, CapsNets enhance their functionality by integrating attention mechanisms within capsule routing, effectively directing the model's analysis toward the most relevant aspects of image data [9, 12]. This strategy is particularly beneficial for AIS prediction because it allows for the focused identification of features and spatial relationships crucial for indicating curve progression, such as vertebral rotation and torsion, thereby enhancing the model's accuracy and interpretability. CapsNets adeptly target the major curve apex of spinal radiographs, distinguishing curves as either progressive or nonprogressive based on key progression indicators [9, 12]. However, the model's primary reliance on vertebral rotation and torsion might not always guarantee reliable predictions of curve progression. An additional limitation arose from the model's lack of comprehensive assessment of skeletal maturity, which focused solely on the Risser sign [12]. Expanding CapsNet's utility, another study [13] successfully integrated 2D radiological images with 1D clinical data, combining spatial and clinical insights for a more encompassing prediction model.

Recurrent neural network (RNN) with LSTM cells

Recurrent neural networks (RNNs) enhance CNN designs to process data that unfold over time. They stand out for adding a "time dimension" to data analysis, enabling the model to pass information through time. This is achieved through hidden layers that store and update information based on new inputs, allowing the model to consider both new and historical data in its outputs. However, RNNs face challenges in maintaining information over long periods, a problem addressed by innovations like Long Short-Term Memory Networks (LSTMs) and Gated Recurrent Units (GRUs) [14]. These technologies introduce controls within the network that help manage how information is kept or discarded, making it easier to learn from data over extended sequences.

LSTMs, in particular, are designed to solve the issue of "vanishing gradients", enabling the network to learn from long-term data dependencies. They use a system of gates that carefully manage the flow of information, making them adept at following the development of conditions such as scoliosis, which changes gradually over time. LSTMs are particularly good at remembering crucial early data, such as the initial Cobb angle or when puberty starts, and can be used to predict how the condition will progress. This feature is crucial for accurately tracking scoliosis over time. In research on the prediction of scoliosis progression [15], LSTMs have been used to analyze clinical and radiological data collected over time. This study developed two models using LSTMs: one classifies the severity of scoliosis as mild, moderate, or severe based on current clinical data, and the other predicts how the spinal curvature will change in the future by using changes in clinical indicators to estimate variations in the Cobb angle.

Convolutional neural networks with attention mechanisms

Traditional CNN models employ fixed-size filters to analyze input images, a technique that falls short when encountering objects or scenes varying in size. Moreover, these models may overlook the complex relationships between different parts of an image, potentially missing crucial information. To overcome these challenges, attention mechanisms have become a critical advancement, allowing models to focus selectively on specific parts of the input data. This focus is achieved by adding extra parameters to the CNN architecture, which helps evaluate the importance of each data point and adjust the model's attention and processing power accordingly [16].

Attention mechanisms greatly improve a model's ability to identify and prioritize key details in images, enhancing both its performance and efficiency. The advantages of implementing attention mechanisms include a better understanding of the input data, increased effectiveness of the model, and improved efficiency. However, these benefits include heightened computational requirements and increased memory consumption. Within the realm of AIS analysis using smartphone photographs [17], CNNs enhanced with attention mechanisms rigorously identify and analyze patterns and features, pinpointing indicators essential for determining scoliosis severity and curve types. Focusing on areas that show significant curvature or asymmetry on the patient's back, attention mechanisms significantly improved the model's accuracy and clarity.

Table 2 illustrates the stages of AIS progression and where machine learning interventions can be applied to predict and manage the disease.

Types of factors referenced

In this retrospective article, the prediction models primarily incorporate two types of predictive factor inputs: 1D numerical and 2D imaging. The numerical category included patient demographic information such as sex, age, weight, sitting height, standing height, arm span, shoulder height difference, and bracing compliance, as well as parameters obtained through physical exams for AIS screening. These parameters included the thoracic rotation angle, the thoracolumbar rotation angle, the lumbar rotation angle, the scapular tilt, the shoulder height difference, the lumbar concavity, and the pelvic tilt. Additionally, parameters calculated from X-ray images, such as the Cobb angle, initial lumbar lordosis angle, initial thoracic kyphosis angle, Risser sign, distal radius, and ulna classification, are included. The imaging category involves using whole radiology images, such as biplanar X-rays and surface topography, as input for model training.

Types of evaluation parameters referenced

1. ROC curve: Also known as the "Receiver Operating Characteristic Curve" or "Iso-Sensitivity Curve", the ROC Curve is primarily used to assess the accuracy of predictions. The area under the ROC curve is known as the AUC, which is used to measure the quality of a model. The value of this area usually ranges from 0.5 to 1, where 0.5 indicates random judgment, and 1 represents a perfect model.

Table 2 Advantages and disadvantages of machine learning models

Model type	Advantages	Disadvantages	Contexts of best performance
Random Forest	High accuracy, robustness to overfitting, handles both numerical and categorical data	High computational cost, less interpretable	Works well with structured data and mixed types
DCNNs	Effective with structured data, high predictive performance, good at recognizing spatial patterns	High computational requirements, needs large datasets	Suitable for image-based tasks
CapsNet	Captures spatial hierarchies, robust to variations, dynamic routing for better feature selection	Complex architecture, difficult to train, high computational resources needed	Effective for image recognition with spatial relationships
RNNs with LSTM	Handles sequential data, retains long-term dependencies, effective for time-series data	Computationally intensive, requires large training data, difficult to interpret	Best for time-series predictions and data with temporal dependencies
CNNs	Excellent at feature extraction and image analysis, scalable, handles high-dimensional data well	Requires large datasets, susceptible to overfitting if not properly regularized	Ideal for image-based tasks and spatial data

2. Root mean square error (RMSE): This is calculated by taking the square of the difference between the actual and predicted values, summing these squares, averaging them, and finally taking the square root. It is commonly used to evaluate the prediction accuracy of regression models. A smaller prediction error results in a lower RMSE, indicating better model prediction performance.
3. Accuracy: This refers to the proportion of correctly predicted samples to the total number of samples in the test set.

Discussion

In the management of AIS, the most crucial aspect is predicting the progression of the curve during the initial visit. Most patients, when first diagnosed, are mild and do not reach the level requiring brace management. However, the ability to identify which AIS patients may significantly worsen during puberty is currently limited. Being able to predict this risk in advance would facilitate clinical decision-making. For patients at greater risk of progression, early intervention treatment can be initiated, while for those at lower risk, a longer follow-up visit can be arranged to avoid unnecessary ionizing radiation. Most current analyses based on curve prediction are retrospective, categorizing patients into progressive and nonprogressive groups based on changes in the Cobb angle between two follow-ups, generally defined as 5 degrees [13, 18–20]. Some studies define an increase of 25 degrees in the Cobb angle from disease onset to skeletal maturity as the threshold for progression or nonprogression [9, 12].

Neural networks have made further technical improvements over traditional deep learning, capturing more features in images to enhance the accuracy of prediction models. Therefore, compared to considering one-dimensional numerical factors, machine learning methods that directly use two-dimensional images as inputs are more common [10, 12, 21]. In [21], Bayesian modeling of input priors was performed using a previously reconstructed 3D spinal set obtained from longitudinal assessments of P-type and NP-type AIS patients, training a discriminative manifold that achieved a classification rate of 81% between P and NP patients, with predicted main curve angle differences within 2.1°. In [7], independent component analysis (ICA) was used to extract 9 independent components (ICs) representing the main directions of shape change from a dataset of 150 AIS patients, with prediction results showing deviations of 1.83°, 5.18°, and 4.79° for the proximal thoracic, main thoracic, and thoracolumbar/lumbar segments from the actual spinal curvature, respectively. In spinal detection using X-rays, anatomical priors can be used to enhance the accuracy of machine learning models. These priors include the relative positions, orientations, and shapes of vertebrae, as well as their biomechanical interactions. Incorporating anatomical priors into machine learning models can improve their ability to detect and analyze spinal deformities [9, 22].

Considering the need for close clinical observation in AIS patients, excessive radiation doses from imaging may pose a cancer risk [23]. Therefore, some studies have utilized radiation-free imaging methods for disease management, including direct methods such as ultrasound and magnetic resonance imaging (MRI), as well as indirect methods like surface topography (ST). ST, a nonradiative back imaging technique, indirectly reflects the condition of the bones of the back, and its correlation with 2D radiology has been confirmed [24]. ST allows for real-time data collection, is generally less expensive than other imaging alternatives, requires almost no expertise beyond marker placement, and has the potential to capture patients' postures while walking. Progress in research based on the ST has evolved from basic determinations of progression/nonprogression (P/NP) to predicting Cobb angle progression in progressive patients. While the image acquisition methods may vary, the core goal is to obtain the images. The three surface topography image acquisition methods aim to capture detailed 3D images for clinical spine and torso assessment using advanced imaging technologies. The DIERS Formetric 4D system [25] employs light grid projection and body markers, requiring subjects to wear an apron and stand on a treadmill for image capture, focusing on analyzing a set of 85 formetric parameters. In contrast, the VIVID 910 3D laser scanner methods [18, 26] involve patients standing in a frame to ensure minimal movement, capturing images from all sides without the need for an apron or specific markers, with the data processed using Geomagic Control software to create a comprehensive 3D model.

The latest method involves using smartphones to take pictures of the back [20], which is the simplest of all collection methods. However, current analyses are mostly based on single ST images and very few clinical parameters. Additionally, since ST indirectly reflects skeletal conditions, factors such as BMI and muscle imbalance can also affect prediction results [18, 26], leading to a wide variation in accuracy rates for predicting spinal progression, ranging from 41 to 92%. This does not fully demonstrate the advantages of multifactor integration with Deep Learning. Nonetheless, it is undeniable that ST provides a radiation-free follow-up method for patients, particularly in mild cases, potentially reducing the radiation dose by 31% [18] and 74% [19], respectively.

Studies [17, 27] based on large-scale population screening data for AIS have identified risk factors for AIS using different machine learning methods based on parameters obtained from basic physical exams. Common factors include the lumbar rotation angle, scapular tilt, shoulder height difference, lumbar concavity, and pelvic tilt. However, the machine learning models used vary, with one based on artificial neural networks and the other on logistic regression.

Traditional methods, such as the SOSORT guidelines and the Lonstein and Carlson method, have been the cornerstone of clinical practice due to their reliability and validation. The SOSORT guidelines provide a comprehensive approach to conservative management and prediction of scoliosis progression, incorporating clinical and radiological assessments [28]. The Lonstein and Carlson method, developed in 1984, calculates the risk of curve progression during growth based on the initial Cobb angle and the patient's Risser sign, among other factors [29]. This method has proven effective in clinical settings, allowing for timely interventions. However, with the advent of machine learning and AI, there is an opportunity to enhance these traditional methods by integrating a wider array of data points and continuously updating prediction models with new patient data. AI-based prediction models offer several advantages over traditional methods. They can integrate multiple data types, including imaging and clinical parameters, leading to potentially higher precision in predictions [9]. AI models can also be continuously updated with new data, improving their accuracy over time [12, 21]. Furthermore, AI can provide individualized predictions based on a comprehensive analysis of each patient's unique characteristics [30]. However, the limitation cannot be ignored.

Like many deep learning models, numerous models in this retrospective review exhibit "black box" characteristics, such as CapsuleNet [9, 12, 13], LSTM [14], and even RF [6, 7], posing challenges in interpreting predictions and potentially hindering their acceptance and use by medical professionals. However, this review also incorporates interpretable models, including Decision Tree models, k-nearest neighbor models [19] and Logistic Regression (LR) models [27]. Decision Trees [21] split data based on certain conditions or feature thresholds, with each node representing an easily interpretable decision rule (e.g., "Is the Cobb angle > 25 degrees?"). This made it straightforward to understand how the model achieved specific predictions. The customized k-NN algorithm [19] excelled in handling unique characteristics of AIS progression and severity classification, utilizing specific features from surface morphology analysis, such as RMS and MaxDev. Integrating domain-specific knowledge into the model's decision-making process and maintaining transparency through clear principles of neighbor selection and features used for classification enhanced performance for specific tasks. LR models [27] were used to identify influential factors for AIS and develop predictive models with various adjusted weights. LR allowed for a clear understanding of how predictive factors influence the model's forecasts, offering insights into the relationships between various physical indicators and the risk of AIS progression. Another study [20] avoided the "black box" issue by using techniques such as Score-CAM to provide visual heatmaps that highlight the areas of the image most influential in the model's decision-making process.

The implementation of machine learning models, while promising, also encounters limitations related to data availability, model transparency, and interpretability. The effectiveness of these models depends on the quality and comprehensiveness of the datasets used for training, which may not always encompass the wide range of variability seen in clinical practice. Furthermore, the studies reviewed predominantly focused on specific subsets of the AIS population, such as those with curve types or stages of skeletal maturity. This focus may limit the extrapolation of findings to the broader AIS community, necessitating further research across a wider spectrum of patient profiles. Finally, the current body of research underscores a significant gap in long-term outcome studies. The dynamic nature of AIS and its progression over time calls for extended follow-up periods to truly understand the impact of various treatment modalities on patient outcomes and quality of life.

Conclusion

This systematic review demonstrates the potential of machine learning models in predicting the progression of AIS. By integrating clinical and radiological data, these models offer a promising tool for enhancing prediction accuracy and personalizing patient management. However, further research is needed to address data availability, model interpretability, and integration into clinical workflows. With continuous advancements in machine learning, it is hopeful that these technologies will become integral to the clinical management of AIS.

Author contribution

L.L. and MS.W. completed the literature search, L.L. wrote the main manuscript text and all authors reviewed the manuscript.

Funding
Not applicable.

Availability of data and materials
Not applicable

Declarations

Ethical approval and consent to participate
Not applicable.

Competing interests
The authors declare no competing interests concerning the subject of this study.

Received: 22 April 2024 Accepted: 22 July 2024
Published online: 08 August 2024

References

1. Altaf F, Gibson A, Dannawi Z, Noordeen H. Adolescent idiopathic scoliosis. BMJ Br Med J. 2013;346:f2508. https://doi.org/10.1136/bmj.f2508.
2. Stokes IA. Mechanical effects on skeletal growth. J Musculoskelet Neuronal Interact. 2002;2(3):277–80.
3. Weiss HR, Goodall D. The treatment of adolescent idiopathic scoliosis (AIS) according to present evidence. A systematic review. Eur J Phys Rehabil Med. 2008;44(2):177–93.
4. Manzetti M, Ruffilli A, Barile F, Viroli G, Traversari M, Vita F, Cerasoli T, Arceri A, Artioli E, Mazzotti A, Faldini C. Is there a skeletal age index that can predict accurate curve progression in adolescent idiopathic scoliosis? A systematic review. Pediatr Radiol. 2024;54(2):299–315. https://doi.org/10.1007/s00247-023-05834-z.
5. Breiman L. Random forests. Mach Learn. 2001;45:5–32.
6. Alfraihat A, Samdani AF, Balasubramanian S. Predicting curve progression for adolescent idiopathic scoliosis using random forest model. PLoS ONE. 2022;17(8): e0273002. https://doi.org/10.1371/journal.pone.0273002.
7. García-Cano E, Arámbula Cosío F, Duong L, Bellefleur C, Roy-Beaudry M, Joncas J, Parent S, Labelle H. Prediction of spinal curve progression in adolescent idiopathic scoliosis using random forest regression. Comput Biol Med. 2018;103:34–43. https://doi.org/10.1016/j.compbiomed.2018.09.029.
8. Atwood J, Towsley D. Diffusion-convolutional neural networks. Adv Neural Inf Process Syst. 2015. https://doi.org/10.48550/arxiv.1511.02136.
9. Wang H, Zhang T, Zhang C, Shi L, Ng SY-L, Yan H-C, Yeung KC-M, Wong JS-H, Cheung KM-C, Shea GK-H. An intelligent composite model incorporating global/regional X-rays and clinical parameters to predict progressive adolescent idiopathic scoliosis curvatures and facilitate population screening. EBioMedicine. 2023. https://doi.org/10.1016/j.ebiom.2023.104768.
10. Yahara Y, Tamura M, Seki S, Kondo Y, Makino H, Watanabe K, Kamei K, Futakawa H, Kawaguchi Y. A deep convolutional neural network to predict the curve progression of adolescent idiopathic scoliosis: a pilot study. BMC Musculoskelet Disord. 2022;23(1):610. https://doi.org/10.1186/s12891-022-05565-6.
11. Xinyi Z, Chen L. *Capsule graph neural network*. In: *International conference on learning representations*. 2018.
12. Wang H, Zhang T, Cheung KM, Shea GK. Application of deep learning upon spinal radiographs to predict progression in adolescent idiopathic scoliosis at first clinic visit. EClinicalMedicine. 2021;42:101220. https://doi.org/10.1016/j.eclinm.2021.101220.
13. Chu K, Kuang X, Cheung PWH, Li S, Zhang T, Cheung JPY. Predicting progression in adolescent idiopathic scoliosis at the first visit by integrating 2D imaging and 1D clinical information. Glob Spin J. 2023. https://doi.org/10.1177/21925682231211273.
14. Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput. 2019;31(7):1235–70.
15. Guo X, Xu S, Wang Y, Cheung JPY, Hu Y. Prediction model of scoliosis progression bases on deep learning. Singapore: Springer Singapore; 2019.
16. Soydaner D. Attention mechanism in neural networks: where it comes and where it goes. Neural Comput Appl. 2022;34(16):13371–85.
17. Lv Z, Lv W, Wang L, Ou J. Development and validation of machine learning-based models for prediction of adolescent idiopathic scoliosis: a retrospective study. Medicine. 2023;102(14): e33441. https://doi.org/10.1097/md.0000000000033441.
18. Hong A, Jaswal N, Westover L, Parent EC, Moreau M, Hedden D, Adeeb S. Surface topography classification trees for assessing severity and monitoring progression in adolescent idiopathic scoliosis. Spine. 2017;42(13):E781-e787. https://doi.org/10.1097/brs.0000000000001971.
19. Ghaneei M, Ekyalimpa R, Westover L, Parent EC, Adeeb S. Customized k-nearest neighbourhood analysis in the management of adolescent idiopathic scoliosis using 3D markerless asymmetry analysis. Comput Methods Biomech Biomed Engin. 2019;22(7):696–705. https://doi.org/10.1080/10255842.2019.1584795.
20. Zhang T, Zhu C, Zhao Y, Zhao M, Wang Z, Song R, Meng N, Sial A, Diwan A, Liu J, Cheung JPY. Deep learning model to classify and monitor idiopathic scoliosis in adolescents using a single smartphone photograph. JAMA Netw Open. 2023;6(8):e2330617–e2330617. https://doi.org/10.1001/jamanetworkopen.2023.30617.

21. Kadoury S, Mandel W, Roy-Beaudry M, Nault ML, Parent S. 3-D morphology prediction of progressive spinal deformities from probabilistic modeling of discriminant manifolds. IEEE Trans Med Imaging. 2017;36(5):1194–204. https://doi.org/10.1109/tmi.2017.2657225.
22. Burns JE, Yao J, Muñoz H, Summers RM. Automated detection, localization, and classification of traumatic vertebral body fractures in the thoracic and lumbar spine at CT. Radiology. 2016;278(1):64–73. https://doi.org/10.1148/radiol.2015142346.
23. Simony A, Hansen EJ, Christensen SB, Carreon LY, Andersen MO. Incidence of cancer in adolescent idiopathic scoliosis patients treated 25 years previously. Eur Spine J. 2016;25(10):3366–70. https://doi.org/10.1007/s00586-016-4747-2.
24. Komeili A, Westover LM, Parent EC, Moreau M, El-Rich M, Adeeb S. Surface topography asymmetry maps categorizing external deformity in scoliosis. Spine J. 2014;14(6):973-983. e2.
25. Patel M, Liu X-C, Yang K, Tassone C, Escott B, Thometz J. 3D back contour metrics in predicting idiopathic scoliosis progression: retrospective cohort analysis, case series report and proof of concept. Children. 2024;11(2):159.
26. Wilczyński J. Relationship between muscle tone of the erector spinae and the concave and convex sides of spinal curvature in low-grade scoliosis among children. Children. 2021;8(12):1168.
27. Yan B, Lu X, Qiu Q, Nie G, Huang Y. Predicting adolescent idiopathic scoliosis among Chinese children and adolescents. Biomed Res Int. 2020;2020:1784360. https://doi.org/10.1155/2020/1784360.
28. Negrini S, Donzelli S, Aulisa AG, Czaprowski D, Schreiber S, de Mauroy JC, Diers H, Grivas TB, Knott P, Kotwicki T, Lebel A, Marti C, Maruyama T, O'Brien J, Price N, Parent E, Rigo M, Romano M, Stikeleather L, Wynne J, Zaina F. 2016 SOSORT guidelines: orthopaedic and rehabilitation treatment of idiopathic scoliosis during growth. Scoliosis Spin Disord. 2018;13(1):3. https://doi.org/10.1186/s13013-017-0145-8.
29. Lonstein JE, Carlson JM. The prediction of curve progression in untreated idiopathic scoliosis during growth. JBJS. 1984;66(7):1061.
30. Tajdari M, Tajdari F, Shirzadian P, Pawar A, Wardak M, Saha S, Park C, Huysmans T, Song Y, Zhang YJ, Sarwark JF, Liu WK. Next-generation prognosis framework for pediatric spinal deformities using bio-informed deep learning networks. Eng Comput. 2022;38(5):4061–84. https://doi.org/10.1007/s00366-022-01742-2.
31. Deng L, Hu Y, Cheung JPY, Luk KDK. A data-driven decision support system for scoliosis prognosis. IEEE access. 2017;5:7874–84.

Publisher's Note