



Comparison of ultrasound scanning for scoliosis assessment: Robotic versus manual

Maria Victorova¹ | Heidi Hin Ting Lau¹ | Timothy Tin-Yan Lee¹ |
David Navarro-Alarcon² | Yongping Zheng¹

¹Department of Biomedical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

²Department of Mechanical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

Correspondence

Maria Victorova, Department of Biomedical Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong.
Email: maria.victorova@connect.polyu.hk

Funding information

The Hong Kong Polytechnic University through the Intra-Faculty Interdisciplinary Project, Grant/Award Number: ZVVR

Abstract

Background: Ultrasound (US) imaging for scoliosis assessment is challenging for a non-experienced operator. The robotic scanning was developed to follow a spinal curvature with deep learning and apply consistent forces to the patient's back.

Methods: Twenty three scoliosis patients were scanned with US device both, robotically and manually. Two human raters measured each subject's spinous process angles on robotic and manual coronal images.

Results: The robotic method showed high intra- (ICC > 0.85) and inter-rater (ICC > 0.77) reliabilities. Compared with the manual method, the robotic approach showed no significant difference ($p < 0.05$) when measuring coronal deformity angles. The mean absolute deviation for intra-rater analysis lies within an acceptable range from 0 to 5° for the minimum of 86% and maximum 97% of a total number of the measured angles.

Conclusions: This study demonstrated that scoliosis deformity angles measured on ultrasound images obtained with robotic scanning are comparable to those obtained by manual scanning.

KEYWORDS

medical robotics, scoliosis, spine, ultrasound imaging

1 | INTRODUCTION

Quantitative assessment of curve severity for adolescent idiopathic scoliosis patients is vital for progression monitoring and treatment planning. The gold standard of scoliosis curve assessment is Cobb's angle¹ on coronal X-ray. Although it is well-known that X-ray radiation causes higher risks of breast cancer,² regular check-up of scoliosis is suggested for at least 6 months intervals.³ To reduce the radiation exposure caused by X-ray assessment, various ultrasound methods for scoliosis assessment have been developed.⁴⁻⁷ Table 1 shows a comparison of ultrasound and X-ray differences.

Various researchers have addressed this problem. For example, a portable 3D ultrasound system Scolioscan Air⁸ was designed for scoliosis curve assessment and progression monitoring for patients in

upright position, resting arms on supporting wall with the elbows at 90°. The system presented in Figure 1A uses a palm-sized ultrasound probe with an optical 3D tracking device and a tablet personal computer (PC) with reconstruction software. Using ultrasound B-mode spinal images with respective spatial information the Scolioscan software reconstructs the 3D spinal curve and generates coronal images for scoliosis angle measurements using the volume projection image (VPI) method.⁹ With a non-planar re-slicing approach, where the skin surface is the reference for a slicing plane, the VPI method obtains an averaged intensity of all voxels of the volumetric image at a specified depth to form an image in the coronal plane. Jiang et al.,¹⁰ developed a 9-times faster coronal image generation approach than original VPI method by skipping the unnecessary 3D reconstruction and directly mapping the B-mode

TABLE 1 Comparison of the characteristics of two scoliosis assessment methods, ultrasound and X-ray

Method	Hazard	System cost size and installation	Application	Angle measurement	Human efforts needed	Image quality	Reliability
Ultrasound	No radiation	Low cost (10k–30k), can be portable, and can be used anywhere	Scoliosis screening and progression monitoring. Can be used as progression monitoring in between X-rays	Coronal curvature, sagittal curvature, and axial rotation with a signal scan	Setup, apply gel, launch scan (apply force, rotate, track), clean gel	Image clear enough to show curvature of bony structures and good to view muscles and soft tissues	Reliable as demonstrated by previous studies ⁴
X-ray	Ionising radiation	High cost (40k–100k), big and shielding room needed	Diagnosis confirmation. Gold standard, not advised to do more than once a year	Coronal curvature with Cobb's angle	Setup, launch scan	Clearer image of bony structures	Widely adopted in clinical practice

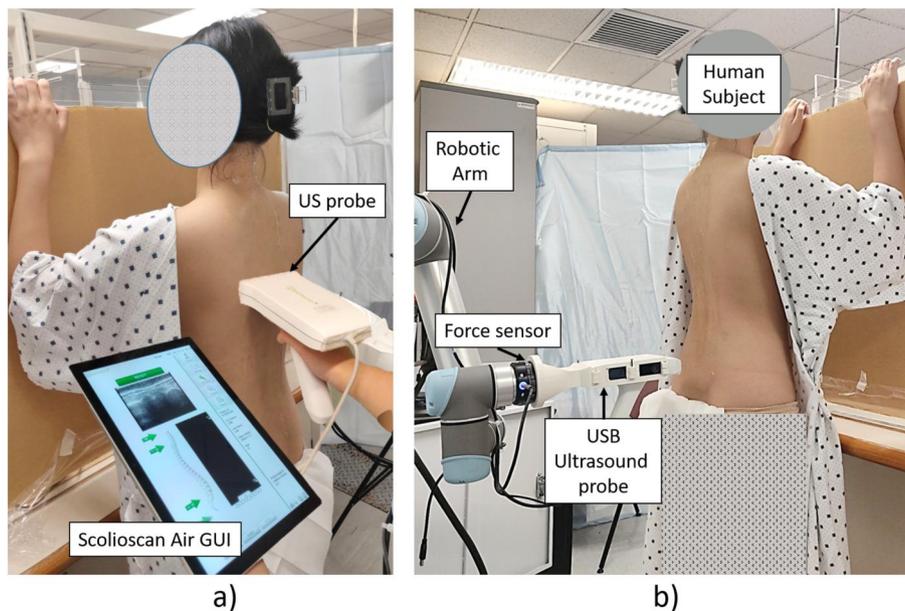


FIGURE 1 Setup for ultrasound scoliosis assessment: (A) manual approach using Scolioscan Air system comprising of tablet personal computer with reconstruction software and USB ultrasound probe. (B) Robotic approach using robotic arm UR5, force sensor FT300 and USB ultrasound probe similar to the Scolioscan Air system.

images and their corresponding positional data. Another method suggested by Vo et al.,¹¹ creates the 3D spine reconstruction compounded from segmented bony features, which allows them to choose the plane of maximum curvature for coronal image projection.

There are several methods to measure scoliosis angle on ultrasound coronal images, depending on which anatomical feature is used for the measurement, such as centre of lamina,⁵ spinous process⁴ or transverse process.⁶ For Scolioscan system deformity angle measurements were based on spinous process. Zheng et al.,⁴ showed that the spinous process angle (SPA) measured on US images had moderate to strong correlation with Cobb's angle measured on X-ray images. Many studies use reliability as measure of similarity of the scanning procedure on repeated trials. SPA had high intra-rater and inter-rater reliabilities with interclass correlation coefficient (ICC)

higher than 0.9 and 0.87, respectively. Another study supports this observation⁶ for Scolioscan system, excellent linear correlation ($R^2 = 0.97$) was observed between the SPA and Cobb's angle with high intra-rater and inter-rater reliability ICC was 0.97 and 0.95, respectively. The validation study for Scolioscan Air showed the angles measured on coronal images obtained with Scolioscan⁴ and portable version Scolioscan Air had a strong correlation, with R^2 higher than 0.7, and with intra-rater and inter-rater reliabilities with ICC larger than 0.94 and 0.88 respectively.

Ultrasonography is a safe and reliable technology for scoliosis evaluation,^{4,6,8} however, due to the nature of ultrasound imaging, the inherent speckle (signal-dependent) noise affects the image quality, making it challenging for operators to distinguish anatomical characteristics. Normally the Scolioscan Air scanning protocol⁸ requires



an operator to adjust the time gain compensation setting, then move the probe upwards, follow the spinal curve, apply sufficient pressure and change the probe's orientation normal to the human's back, at the same time observing the captured B-mode image quality, which is a non-trivial task for a non-experienced user. Usually, it takes around 3–6 months for an operator to become experienced. Poor scanning performance, inconsistent contact with skin, and jerky movements directly reduce the quality of the coronal images. The problems encountered during manual scanning can be solved by using a robotic arm equipped with a force sensor, where robot can apply a constant force to the human back and adjust the probe's orientation in real-time to keep it normal to the skin.

To reduce the effect of the human factor on spinal ultrasound examinations, several research groups proposed using robotic ultrasound systems.^{12–15} While several robotic ultrasound systems are available for spinal applications (with very few for scoliosis assessment), they primarily demonstrate proof-of-concept results, for example, tests on phantom tissues or on a very small sample size of human subjects. Up to this day, no extensive study has been conducted to compare the performance of robotic ultrasound assessment for spinal assessment with standard manual ultrasound procedures. Li et al.,¹⁵ presented a robotic approach for ultrasound imaging-based navigation to locate spinal features using deep learning (DL) and reinforcement learning on dataset of 648 images of spinous process to imitate the human decision-making during US scanning. The design was evaluated in an ultrasound simulation environment, and the resulted average error on the test set was 9.5 ± 10.0 mm. Another study by Tirindelli et al.¹³ proposed a robotic-ultrasound system for spinal levels counting and injection site localisation based on US images DL processing. Tirindelli et al. used 3972 images for training, resulting in an average localisation accuracy of 2.1 ± 2.6 mm on the test set. Although these papers do not address scoliosis, the approaches employed can be used for various spinal applications. The following two studies do address scoliosis assessment: Yang et al.¹⁴ proposed to use a robotic arm to navigate through the spine based on the pre-calculated trajectory yield from the tracking camera image DL processing, using 2500 images for training. The scanning was done in a prone position, different from the Scolioscan Air protocol. The validation was performed on the phantom with different curvatures; the mean absolute error was 1.6 for moderate scoliosis. The concept was only demonstrated on two human subjects (this shows that a more extensive human study is needed to fill this gap in the literature). Human scanning was performed for 120 s capturing 150 ultrasound frames per each.

The robotic ultrasound system used in this study (Figure 1B) was designed for scoliosis assessment and was initially presented in Victorova et al.¹² and extended in Victorova et al.¹⁶ It uses a DL-based navigation control to detect the centre of the spine, which is defined as spinous process tip, in real-time at each B-mode ultrasound frame; thus, there is no need for additional external sensors, such as a camera. The proposed robotic system can also adjust the probe orientation to keep the ultrasound probe normal to the surface and

maintain good acoustic coupling with the back of the subject. Good acoustic coupling can be defined as consistent contact with the human skin, where there is no air gap between the probe and skin, which would affect the resulting image quality. The details for the proposed robotic ultrasound system are described in Section 2.3. The dataset consisted of 25 774 images of spinous processes, and the resulting mean localisation error for the human subjects test set was 1.0 ± 0.99 mm. The real-time experiments on human showed that robotic system was able to centre the spine in the field of view with error 1.0 ± 0.8 mm, compared to manual scanning 4.6 ± 4.6 mm for the subject with mild scoliosis; for the moderate scoliosis the errors were 2.8 ± 1.8 mm for robotic approach and 4.3 ± 3.9 mm for manual.

The contribution of this study is to demonstrate that robotic scanning for 3D ultrasound imaging can achieve the same assessment for scoliosis in comparison with manual scanning. Figure 2 shows the study flow chart. We first collected human subject profiles and explored whether robotic scanning is beneficial and can replace manual scanning. Then, statistical analysis was conducted for inter- and intra-rater reliability of scoliosis angle measurement, with ICC based on a single-rating ($k = 3$), absolute-agreement, 2-way mixed-effects model. The reliability shows an extent to which a scanning procedure yields the same results on repeated trials. The mean absolute deviation (MAD) of angles measured on images obtained by different methods would be compared with those reported previously. Two human raters measured spinal process angles on coronal images of six scans for each subject, three robotic and three manual scans.

To the best of our knowledge, this is the first study focussing on comparing robotic ultrasound scanning with human manual scanning for scoliosis assessment on a group of human subjects. In addition, it is the first work using a robotic ultrasound system which includes evaluation of robotic approach based on patient's comfortability survey.

2 | METHODS

2.1 | Subjects

Twenty-three adolescents (21 female, 2 male) with confirmed or suspected scoliosis were recruited for this study. The main inclusion criteria were: mild or moderate scoliosis, age 14–18 years; no previous interventions on spine, since it may result in unpredictable image artefacts; no overweight ($BMI < 25.0$ kg/m²) since a high BMI may result in poor ultrasound image quality with the 7.5 MHz probe utilised in this investigation. Ethical approval was granted for this study.

The subject leaned on the wall or wall-imitating stand during the scanning procedure. The subjects lifted their hands so that the angles between the elbow and the wall, elbow, and body are 90°, as in Figure 1. The subject was asked not to move but to breathe freely during the procedure. Total six scans were performed on each

subject, three times with manual approach and three times with robotic scanning. Before each scan, the subject was asked to relax and then stand in scanning position so that each scan would be independent of the previous; thus, the posture might vary slightly. Of all the subjects, 10 chose to participate in the survey on the comfortability of robotic and manual scanning. The summary of the differences of robotic and manual scanning presented at Table 2.

2.2 | Manual scanning procedure

The manual scoliosis assessment was performed with a Scolioscan Air system, consisting of an ultrasound transducer of 80 mm aperture and frequency of 7.5 MHz (UW-1C, Sonoptech), a Realsense depth camera (T265, Intel), and a tablet PC. The ultrasound probe communicates with PC at 60 fps. The software installed on the PC collects the ultrasound and spatial data, processes it to form 3D reconstruction and coronal spinal images. The software interface enables manual measurement of the scoliosis deformity angle using the feature profile in ultrasound image formed by spinous process,

previously denoted as SPA. The spatial data of the ultrasound device is obtained through a camera attached to the probe using a vision/odometry-based vSLAM algorithm.¹⁷ This algorithm tracks the changes in the environment (due to the probe motion during the scanning) from the initial position and outputs the position and orientation of the tracking camera.

For the assessment procedure, the operator of the ultrasound probe set the initial (at the level below L5) and final (around C5 of the cervical spine) points of the scanning and then started to move the probe upwards, applying enough pressure to maintain the stable probe-skin contact to see the bony features. The operator tried their best to keep the spinous process in the middle of the ultrasound frames available in real-time for observation. The average time to complete a scan was around 60 s.

The manual ultrasound scanning of the human back was performed three times by an experienced operator with Scolioscan Air.⁸ After each scan, nine coronal images at different slicing depths were generated for the operator to visually assess the quality of the scan and measure the spinal curvature to generate the on-site report for the subject's use.

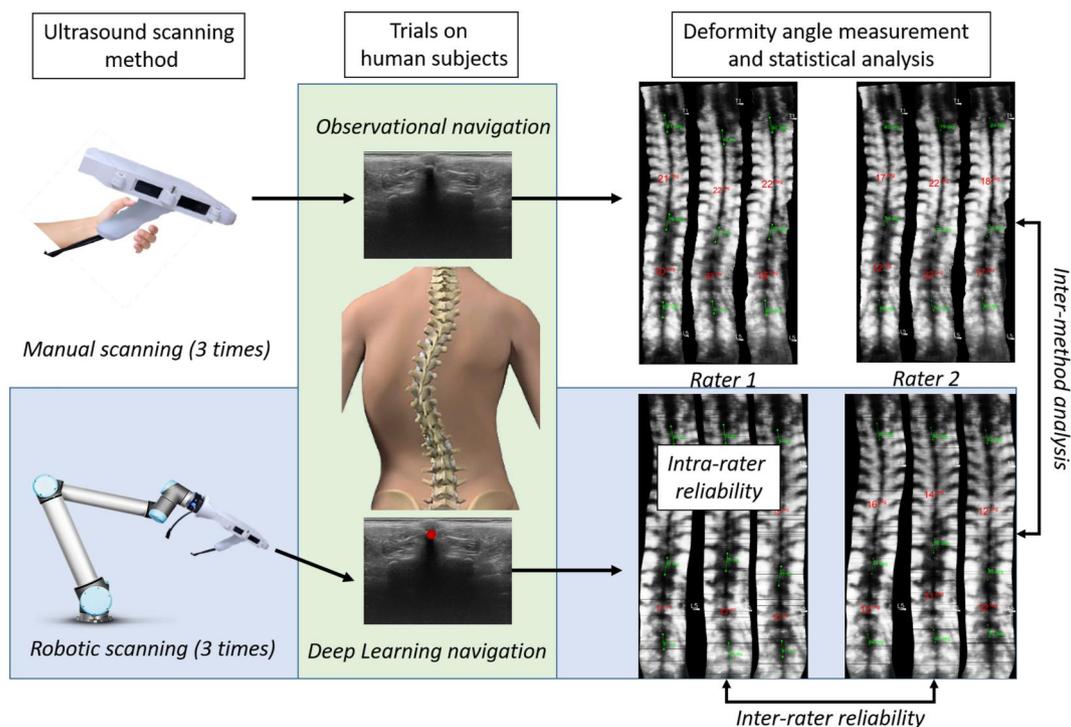


FIGURE 2 The study workflow diagram. Ultrasound scanning was performed with two methods: robotic and manual, three times each. During the robotic procedure the spinal navigation were performed with Deep Learning (DL) approach, whereas during manual procedure the navigation relies on human expertise in reading ultrasound images. Resulted coronal images of reconstructed spine were measured to assess the scoliosis severity by two human raters. The various analysis on reliability and validity were performed.

TABLE 2 Comparison of the characteristics of two scoliosis assessment scanning methods, ultrasound manual and ultrasound robotic

Method	Estimated system cost (USD)	Navigation system	Force control	Scanning time (s)
US manual	30k	Hand and visual	Variable force	60
US robotic	45k	Robot-based and US-based	Constant force 10, 15 N	120

2.3 | Robotic scanning procedure

The setup for robotic scoliosis assessment, shown in Figure 1B, employs a robotic arm (Universal robot UR5) with a force sensor affixed (FT300, Robotiq) and a palm-sized ultrasound probe (Sonoptek). The linear ultrasound probe with a width of 80 mm has a frequency of 7.5 MHz with an imaging depth of 6 cm and transfers raw data at 10 frames per second via a USB port to a PC, where the images are produced in 640×480 pixels. The robot was connected to a PC through TCP/IP protocol, and the control algorithm was launched at a 30 fps Hz rate.

The patient's posture is similar to the Scolioscan Air scanning procedure. The operator drags the robotic actuated probe using build-in admittance control to the base of the subject's spine, with the probe parallel to the ground. From this position, the robot starts to drive the ultrasound probe further until it reaches the human skin (force detected). Then the robot starts moving upwards, applying the preset force, adjusting rotations, and following the curve by detecting spinous processes in real-time. The preset force was chosen after taking into account results from other robotic studies, which suggested using 1–3 N for soft tissues¹⁸ and 4–8 N for the abdominal

region.^{19,20} We tested larger forces because the bony structures were more rigid. In Tirindelli et al.¹³ it was found that 10 N produces better quality during robotic scanning of the spine for subjects with BMI < 23 kg/m², and 15 N is more appropriate for subjects with BMI > 23 kg/m².

The Scolioscan Air system operates at frame rate of 60 fps, which results in average 3600 ultrasound frames per scan. In contrast due to current hardware limitations, the robotic system operates at 30 fps, but receives images at the frame rate of 30 fps. To get the same amount of ultrasound images as for manual scanning the speed of robotic scan should be twice slower. Since the usual scanning speed is around 8–10 mm/s, the robotic scan was performed with 4 mm/s. Figure 3 illustrates some of the timestamps of the robotic procedure, showing how the robotic arm follows the profile of the human back, adjusting the probe's orientation. The average scanning time for one scan was around 120 s.

Robotic control method. The schematic overview of the robotic scanning approach is presented in Figure 4A. The robotic arm is controlled to maintain the constant force applied to the subjects back with PID control,^{21,22} which takes the control error as a difference of measured force to reference force (10–15 N, depending on BMI¹³).



FIGURE 3 The images from the robotic scanning procedure video for one of the subjects. The robot adjusts the ultrasound probe rotation according the force sensing and smoothly scans through the patient's back, maintaining constant contact.

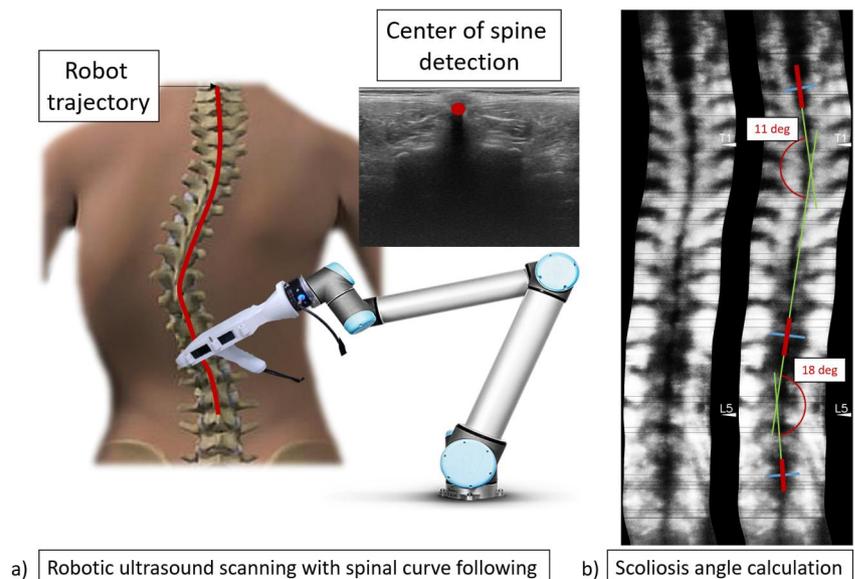


FIGURE 4 (A) Overview of the robotic scanning approach, where the robotic arm automatically follows the spinal curvature based on B-mode images; (b) Principles of spinous process angle measurement based on the indication of the most tilted vertebrae on the coronal image of spinal reconstruction.



Meanwhile, for automatic curvature following, the algorithm fuses the network located spinous processes into a continuous path based on the Kalman filter, compensating for the missing parts (intervertebral gaps). It also adjusts the probe's orientation along the normal vector to a subject's back, the speed of orientation change depends on the spinal region, which is predicted by the network described below.

The robot uses a machine-learning algorithm to follow the spinal curvature and detect the spinal region for rotation adjustments, described in earlier stage of the current work.¹⁶ The fully connected network model was designed to detect the spinous process location to indicate the spine centre at each ultrasound B-mode frame in real-time. The model presented in Figure 5 is built on ResNet²³ backbone, which is used to extract the feature vector from the input ultrasound image. The model has two heads, first serves as spinous location prediction, where the backbone is followed by three deconvolutional layers with batch normalisation²⁴ and ReLU activation,²⁵ which serve as a decoder to generate the image features heatmap.²⁶ The deconvolutional layers are followed by a single 1×1 convolutional layer, which converts the resulting features matrix to a final heatmap where each image pixel intensity represents the probability of being one of the classes; in our case, the maximum intensity represents the high probability of the pixel belonging to a spinous process class. The mean squared error determines the difference between the actual and forecasted heatmaps. The final location is calculated from the heatmap's greatest intensity pixels.

The second head of the network has a fully connected layer after the backbone. The output is the probability of each ultrasound image belonging to a spinal region class: 'sacrum', 'lumbar,' 'thoracic.' According to the output class, the algorithm changes the control gain K_{pitch} which specifies how fast the rotation of the probe is changing, $r_x = -K_{pitch}m_x$, where the torque m_x is obtained from the force sensor. For example, the thoracic part of the spine generally has larger

curves in the sagittal view than the lumbar part; thus, a larger control gain is needed. 25 774 images were used for training and testing of the network resulting in a five-fold cross-validation accuracy of 0.939 ± 0.010 for the classification task and 0.973 ± 0.006 for the localisation task with mean distance error of 1.0 ± 0.99 mm. This distance error can be compared to anatomical references. According to Aylott,²⁷ the average spinous process width in the lumbar region ranges from 6 to 11 mm, whereas the thoracic one is narrower. The error is therefore contained within anatomical limits.

2.4 | Deformity angle measurement

According to the scoliosis application and the nature of the images, the observational criteria for image quality assessment can include a (a) clear spinous process shadow path, which allows unambiguous deformity angle measurement; (b) visibility of the spinal features, such as lamina, transverse process, etc.; (c) uninterrupted acoustic coupling, thus no missing spots due to the contact lost with the skin.

These criteria help to ensure that the image quality of the resulting coronal images is suitable for the subsequent scoliosis deformity angle measurement, which is the assessment's main goal. Because the image quality criteria is only observational and there is no unambiguous method of assessment, a quantitative method of image quality analysis is required. Since the main goal is to measure scoliosis deformity, the measured angles are used as an indirect quality marker. If the measurements of the robotic approach are the same or comparable with angles obtained by the manual method, the image quality is considered to be comparable.

In total, 23 subjects were examined with the proposed automatic scanning for scoliosis assessment; each had one or two deformity angles. Two raters, R1 and R2, conducted the measurement of spinal curve deformity angle (SPA) according to the method presented in

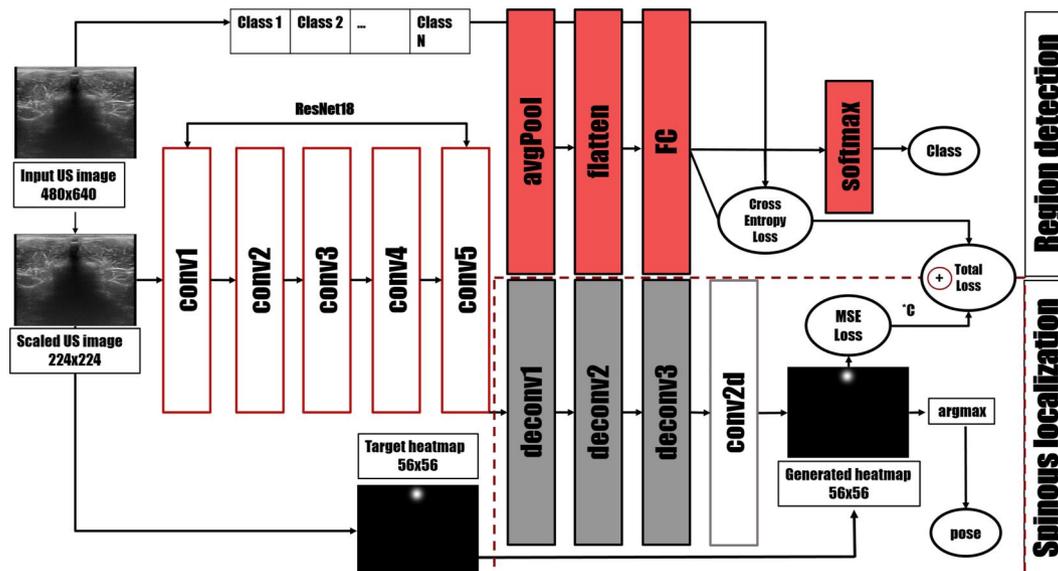


FIGURE 5 Network for spinous process localisation and spinal region classification.



Zheng et al.⁴ The rater R2 had more than 5 years of experience in measuring scoliosis, while the other one R1 was a newcomer who received basic instructions on the measurement procedure. The deformity angle measurement procedure is summarised in Figure 4B. The raters used software to load coronal projection images and draw the lines at the most tilted vertebrae levels of the spinal curve. Then, the software outputs the spinous process deformity angles, calculated as an angle between drawn lines. The raters measured the deformity angle on six coronal images from three robotic and three manual scans.

2.5 | Statistical analysis

Statistical analyses were performed using the IBM SPSS Statistics Version 25 (IBM). To assess the intra-rater reliability of spinal deformity angles measurement, the intra-class correlation coefficient ICC with 95% confidence interval was calculated. Koo et al.,²⁸ stated the guidelines for proper ICC test design selection. For intra-rater reliability, the usual model used was two-way mixed with absolute agreement definition; the reported value was single-rating since there was no averaging of the measurements. The absolute agreement was used to see how different the same angle measurements were among different scans by the same rater.

Similarly, the inter-rater analysis uses ICC to reflect the correlation between the measurements of two raters; it is performed on means of three scans of each method for each rater; thus, the reported value is average-rating. According to Portney et al.,²⁹ the ICC values between 0.5 and 0.75 indicate moderate reliability, values greater than 0.75 indicate good reliability and greater than 0.9 excellent reliability. Since 5° in angles measurement variability is acceptable and commonly reported,^{4,30} the percentage and number of curves, which MAD is less than 5° is reported.

The Wilcoxon signed-rank test was used to compare the angles measured on coronal images obtained from robotic and manual assessments to evaluate the validity of ultrasound assessment. The inter-method analysis employed a non-parametric Wilcoxon test instead of the standard t-test method. This test shows whether there is a statistical difference between the angles obtained by robotic and manual scanning. The analysis was performed on the mean of three scans for each scanning method. Another metric to assess the inter-method correlation is R^2 value computed with the linear regression method.

3 | RESULTS

The findings present an examination of the measured SPA scoliosis deformity angles obtained through manual and robotic scanning. Angles serve as an indirect measure of image quality, and by comparing angles, we can compare image quality. However, it serves a practical purpose in assessing scoliosis; we also investigate some observational criteria for robotic and manual scans.

The recruited subjects had an average body mass index (BMI) of 18.9 ± 2.2 kg/m² and age of 15 ± 1.0 years. Ten subjects had a single thoracic or lumbar curve, while 13 possessed a double curve, yielding a total of 36 curves for analysis in this study. The SPA measured angles of these curves ranged from 7 to 31°, and the average value was 15.7 ± 5.6 °. While all the subjects have mild or moderate scoliosis, the second curve of five subjects was slightly less than 10°. Since the purpose of the study is the curves angles comparison and not the scoliosis detection, these angles were also included in the study.

Figure 6 presents one robotic (left) and one manual (right) scan for each of six exemplary subjects. It can be noticed that the resulting curve angles are similar between the two scanning approaches. The survey result shows no significant difference between the compatibility of robotic (7.8 out of 10) and manual (8.7 out of 10) scanning.

Table 3 displays the ICC values for two-way mixed model with absolute agreement definition, single-rating type. The MAD, SD, and SEM values were calculated for each curve based on angles measured on three scans of one method (robotic or manual). The average values (with min and max) for all the curves for intra-rater analysis are presented in Table 3. The highest value of ICC = 0.868 was found for robotic method Rater 2, where MAD was below 5° except for one outlier. The lowest value of ICC = 0.753 was found for manual method Rater 1 with MAD below 5° for all angles.

The results of the inter-method comparison are presented in Table 4. The Wilcoxon test with the value of 0.166 and 0.06 (with $p = 0.05$) showed no significant difference between angles measured on images obtained with a robotic system (average of three scans) and with manual scanning. The R^2 values of 0.73 and 0.72 demonstrated a moderate correlation between the SPA results by the two scanning approaches, robotic and manual. The correlation plot between the two methods is displayed in Figure 7. Table 5 presents the ICC, MAD, SEM values for inter-rater analysis for each scanning method. According to Table 5 inter-rater reliability had ICC = 0.772 for robotic and ICC = 0.843 for manual methods. The MAD was below 5° for all cases except four angles for manual scan and eight for robotic.

Observational criteria, as mentioned in Section 2.4, were also examined during the experiments, and there were several differences between the manual and robotic approaches. Section 4 Discussion and Conclusions presents the cases in which the observational criteria were not met. Figure 8 depicts manual scanning problems, while Figure 9 depicts robotic problems. The Table 6 depicts a summary of the observations made during scanning, emphasising the advantages and disadvantages of the robotic and manual approaches.

4 | DISCUSSION AND CONCLUSIONS

The current study shows, using various statistical methods, that the scoliosis angles measured by robotic scanning are comparable to those measured by manual scanning. The statistical analysis shows that both robotic and manual approaches have good reliability, with

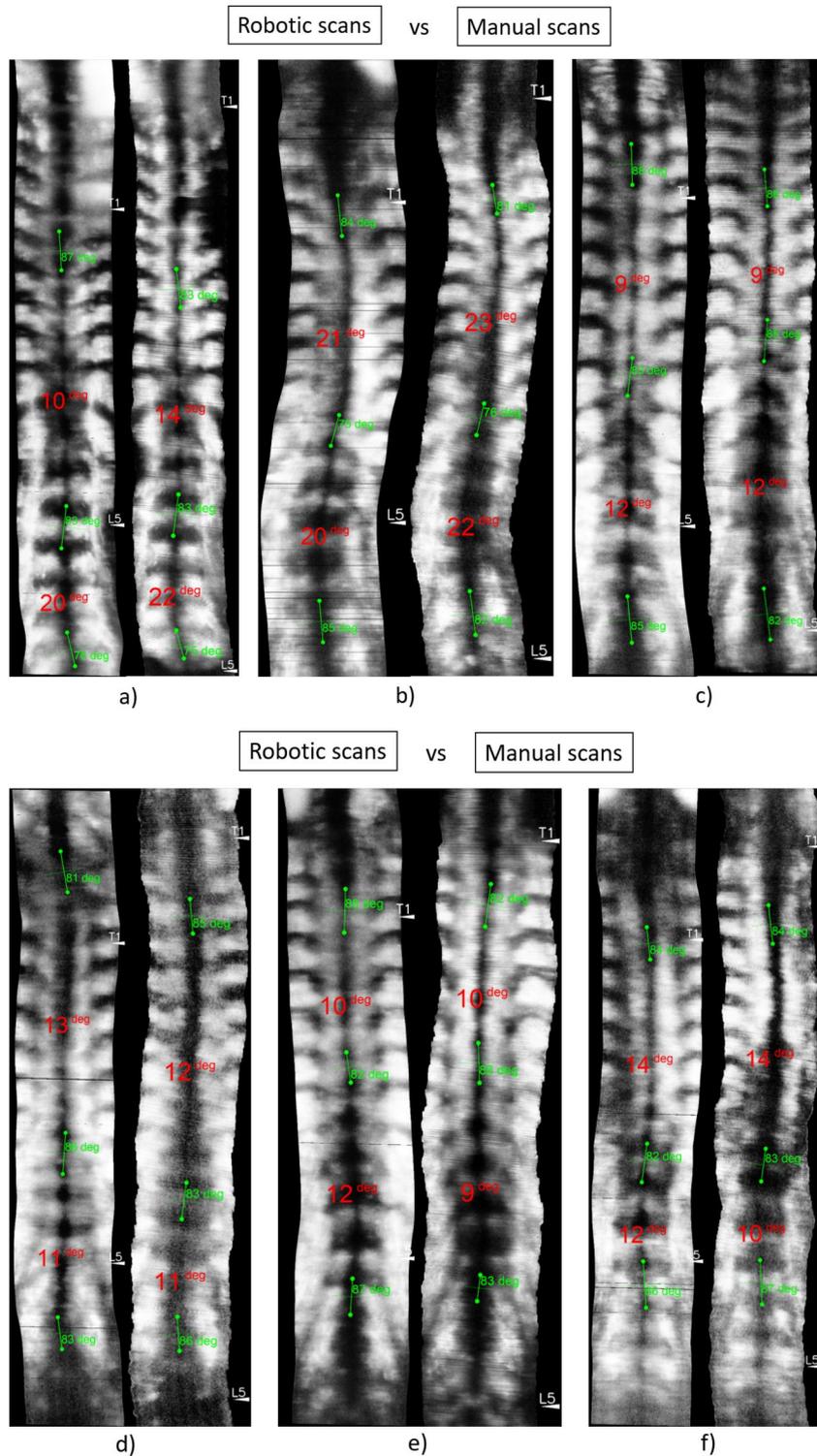


FIGURE 6 Comparison of coronal images resulted from robotic (LEFT) and manual (RIGHT) scanning with measured spinous process angles for six subjects (a–f) with different scoliosis severity. The robotic scanning yield smoother coronal reconstructions and more prominent spinal features.

values slightly higher for the robotic approach. According to the Wilcoxon test, both raters agreed that there was no significant difference between robotic and manual angles. A moderate linear correlation was found between robotic and manual approaches; however, the variation in the correlation numbers for the Wilcoxon test can be explained by the considerable difference in the rater's experience.

Kottner et al.³¹ summarised the interpretation of the ICC values; he showed that the values of 0.6–0.8 are often used as a minimum standard for research purposes, where the sample size is relatively smaller, however for clinical practice, the value should be at least 0.9. In the current study the ICC values were between 0.75 and 0.87. Previous studies have revealed that Cobb angle measurements of moderate scoliosis on X-ray images have a 3 to 5° intra- and inter-



TABLE 3 Intra-rater variation and reliability of coronal curvature assessments using robotic ultrasound scanning compared with manual ultrasound scanning

Methods	Raters	ICC (95% CI)	MAD (max), °	% below 5°	SD (max),°	SEM (max), °
Robotic	R1	0.804 (0.690–0.887)	2.3 ± 1.1 (4.0)	100 (36)	1.8 ± 1.0 (4.0)	1.1 ± 0.6 (2.3)
	R2	0.868 (0.785–0.925)	2.5 ± 1.4 (5.3)	97 (35)	1.9 ± 1.1 (4.0)	1.1 ± 0.6 (2.3)
Manual	R1	0.753 (0.619–0.855)	2.6 ± 1.0 (4.7)	100 (36)	2.1 ± 0.8 (3.6)	1.2 ± 0.5 (2.0)
	R2	0.829 (0.725–0.902)	2.5 ± 1.7 (6)	88.8 (32)	2.0 ± 1.3 (4.7)	1.1 ± 0.7 (2.7)

Abbreviations: % below 5° (num of angles), clinically not significant difference of angles below 5°; CI, confidence intervals; MAD, mean absolute deviation; SD, standard deviation; SEM, standard error of the mean.

TABLE 4 Comparison of coronal curvatures assessments between robotic ultrasound scanning and manual ultrasound scanning (on mean of three scans for each method)

Raters	Methods	MAD (max), °	% below 5°	Wilcoxon (p = 0.05)	R ²
R1	Robotic versus manual	1.8 ± 1.4 (5.7)	97 (35)	0.166	0.73
R2	Robotic versus manual	2.5 ± 1.6 (6.0)	92 (33)	0.06	0.72

Abbreviations: % below 5° (num of angles), clinically not significant difference of angles below 5°; MAD, mean absolute deviation; SD, standard deviation; Wilcoxon, a non-parametric signed-rank test.

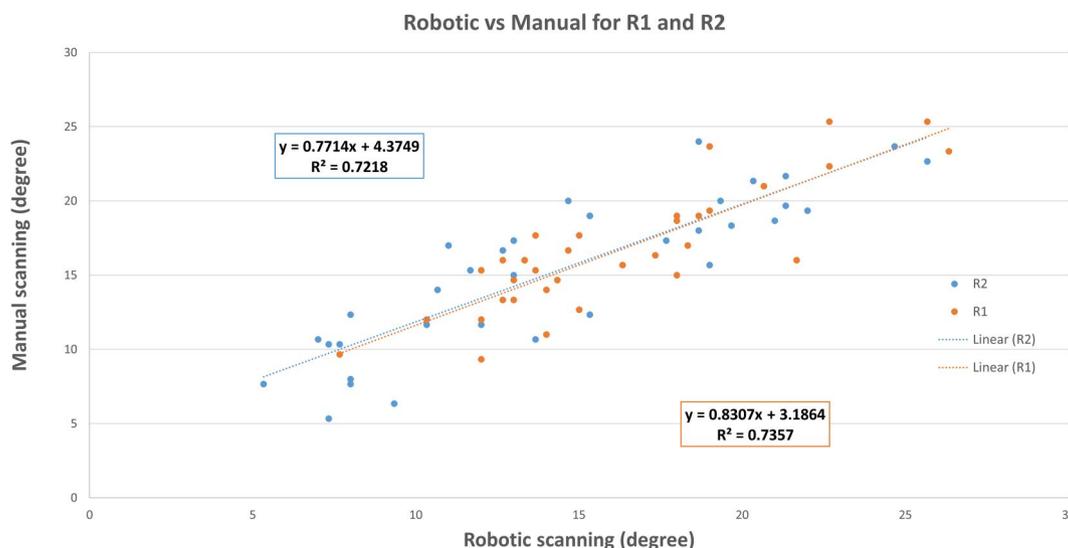


FIGURE 7 Correlation and equation between the spinous process angle measurement results by robotic scanning and manual scanning for mean of rater 1 and rater 2.

TABLE 5 Inter-rater correlation for robotic ultrasound and manual ultrasound approaches

Methods	Raters	ICC (95% CI)	MAD (max), °	% below 5°	SD (max),°	SEM (max), °
Robotic	R1 versus R2	0.772 (0.557–0.883)	3.6 ± 2.7 (10.6)	77.8 (28)	2.5 ± 1.9 (7.5)	1.8 ± 1.3 (5.3)
Manual	R1 versus R2	0.843 (0.694–0.919)	2.8 ± 2.2 (8)	89 (32)	1.9 ± 1.6 (5.6)	1.4 ± 1.1 (4.0)

Abbreviations: % below 5°(num of angles), clinically not significant difference of angles below 5°; CI, confidence intervals; MAD, mean absolute deviation; SD, standard deviation; SEM, standard error of the mean.

observer variability.^{1,32,33} The intra-class correlation coefficient ICC for Cobb's angle was from 0.83 to 0.99 according to a systematic review.³⁰ Angle measurements on ultrasound images had a variability of 2–5°.4–6 The current study shows that MAD for inter-rater and intra-rater mostly lays in a range of 0 to 5°. However, mostly the angles variability lays within acceptable range, there were a few

outliers for inter-rater analysis (Table 3). Because one of the raters was inexperienced, the inter-rater MAD could be influenced. MAD is expected to be lower with both experienced raters.

According to a review of medical image quality assessment,³⁴ currently, there is no standardised way to assess ultrasound image quality. Specifically to an application, US image quality can be defined

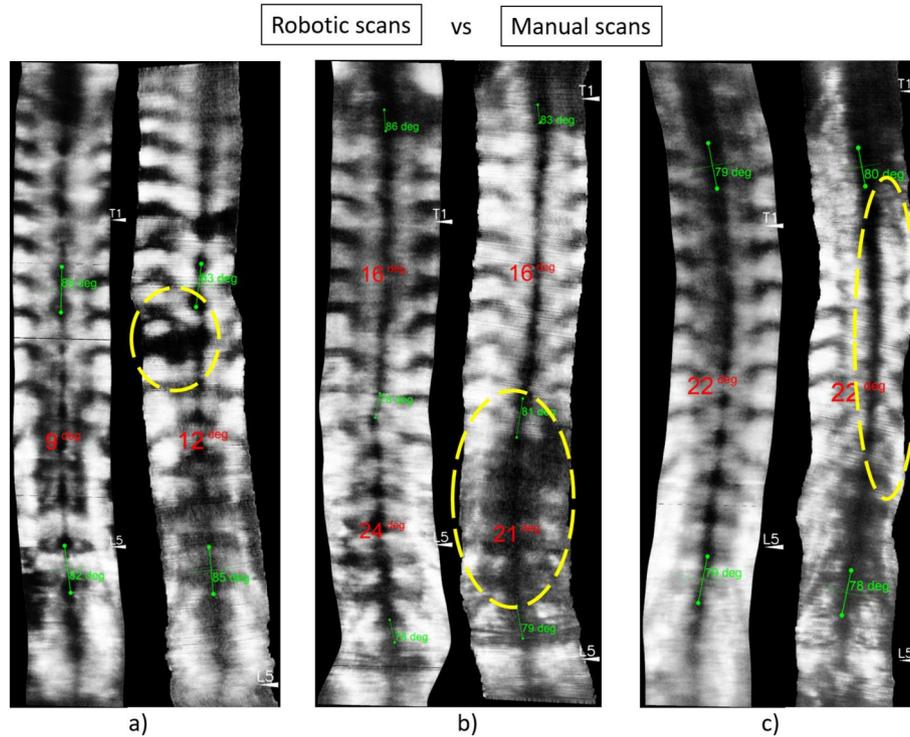


FIGURE 8 Comparison of coronal images obtained from robotic (LEFT) and manual (RIGHT) scanning shows challenges of the manual approach. (A) Skin contact lost (black spots); (B) spinal lumbar features are less prominent for manual scanning; (C) the challenges of following the curve leave the spinal aside of field of view.

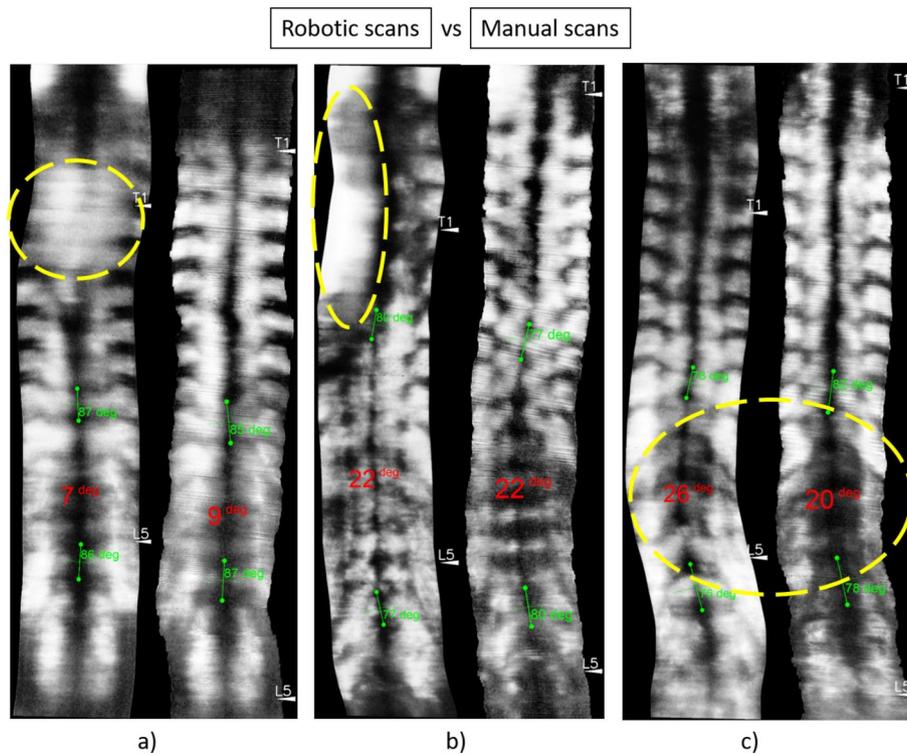


FIGURE 9 Comparison of coronal images obtained from robotic (LEFT) and manual (RIGHT) scanning show challenges of the robotic approach. (A) A subject with narrow space between scapula bones; (B) a subject with uneven back in coronal plane; (C) a subject bent his/her back under the scanning pressure due to the weak muscular system (a tall subject with low body mass index).



TABLE 6 Observations made during scanning experiments that highlight the benefits and drawbacks of two scanning methods, robotic and manual

Figure	Robotic	Manual	Num. of observed subjects
Figure 8 Robotic scanning advantages			
A	+ More regular contact with the skin	– Black area, contact lost	4
B	+ Lumbar spine is more prominent	– Lumbar spine is less distinguishable	10
C	+ Spine centring and surrounding features visible	– Spinous process surrounding features were out of the field of view	5
Figure 9 Robotic scanning disadvantages			
A	– Grey area, probe detached from the skin due to the narrow scapula space	+ Spine visible, more flexibility in rotation to overcome narrow scapula space	2
B	– Grey area, high concavity at one side of the back	+ Visible features, more flexibility for yaw rotation	1
C	– The measured angle appeared to be larger, the subject bended under robotic pressure	+ Subject bended less in lumbar region under the scanning pressure	3

by comparison to other modalities, such as X-ray, CT, MRI. Suppose there are no reference images of different modalities. In that case, the image quality is indirectly defined as an ability to see the features of interest (target anatomical structures) within the image.³⁵

Figure 8 illustrates the noticeable benefits of the robotic approach. Figure 8A shows the more regular contact with the skin surface during robotic scanning compared to manual, where the points of contact lost appeared in black for four scanned subjects. The other common observation is that the features of the lumbar spine were less prominent and poorly distinguishable for 10 subjects scanned manually compared to robotic scanning, as shown in Figure 8B. These advantages of the robotic approach can be explained by the stable force control, which includes the constant pressure onto the skin and probe's orientation adjustments. The other example (Figure 8C) shows the benefits of having a spinal curvature tracking algorithm to centre the spine in the field of view, which ensures that the spine is always in the field of view (which is especially useful for patients with severe scoliosis) and also gives a better view on the surrounding features such as laminae and transverse processes to utilise different approaches for angle measurements. For five subjects scanned manually, the spinous process surrounding features were out of the field of view compared to robotic scanning. In general, the generated coronal image of the robotic scan looks smoother, the spine is lying in the middle of the image, and bony features are more distinguishable than in the manual scan.

The number of US frames for reconstruction is similar in both robotic and manual approaches; the scanning speed in the robotic approach is twice slower than for manual. However, the speed difference may give an additional smoothness contour to the reconstructed image in the robotic scan; the concern is that a longer scanning time may cause instability in the subject's posture since the subject might make small motions during that time.

Other important benefits of the robotic system are constant force control; more precise navigation, centring the spine with error

1.0 ± 0.8 mm as opposed to manual scanning 4.6 ± 4.6 mm; liberation of the ultrasonographer from repetitive tasks and more concentration on diagnostics and treatment.

Several limitations for using the robotic approach were encountered during the study. Figure 9 illustrates less successful robotic scanning results. Coronal image of robotic scan for the first case (Figure 9A) represents the subject with low BMI and tight prominent scapula bones, which leaves too narrow space between the right and left scapula for a robotic actuated probe to pass through. Thus, the probe detached from the skin at the scapula level during the scanning, and image quality suffered in that region. In comparison, the manual scanning handled the tight scapula situation well since the operator could roll the probe to fit the narrow space and pass through. The issue with tight scapula was encountered in two subjects out of a total number of examined subjects. The robotic system cannot recognise the tight scapula cases and perform the necessary roll rotation.

The second challenging case for robotic scanning is displayed in Figure 9B. The subject had an uneven back, with high concavity at the left side of the back at the level of the thoracic region; thus, the left side of the probe was not compliant with the subject's back. There was only one subject with this peculiarity during the experimental trials. For such cases, two solutions can be used. The first one is to apply a greater amount of gel to fill in the gap; the second one is the greater robotic row rotation, which was implemented by a force sensing approach; however, this type of rotation was not fully studied since there were no subjects with severe scoliosis.

The third case shows a tall subject with low BMI (Figure 9C), where the subject could not resist the force applied by the robot (even the lowest settings of 10 N) and bend in the lumbar region as the probe was applying pressure on the back. Thus, the lumbar curve from a robotic scan is much larger than that from a human scan. This situation was observed for three subjects out of total. It was noticed that tall subjects with low BMI tended to have this problem. This might be because they had a weaker muscular system.



One of the study's limitations was that there were no X-ray reference images to measure the actual scoliosis angles. However, the goal was to compare ultrasound scanning approaches rather than to provide the final medical record. In clinical practise, the manual ultrasound system has already been validated and proven to be reliable for scoliosis assessment.⁴ Thus, if the robotic approach yields angles comparable to manual practise, it is a good sign to begin a more extensive study, where X-ray images will also be required for the experiment purity.

However, the study already shows a good correlation among the angles measured from robotic and manual scanning; it is suggested to do a more extensive case study on more subjects. Manual scans were performed by an expert who has done this procedure for more than 5 years. To train an operator, in general, it takes about 6 months to learn the scanning procedure on different types of subjects, to perform multiple scans for each to learn to correct the scanning path. It would also be meaningful to compare the robotic approach with a non-experienced user, which might show that the robotic approach performs better. For comparison purposes, it is better to obtain the ground truth spinal image from the patient before scanning, such as X-ray or Scolioscan image, which proved to be reliable by previous study.⁴

The comparison of robotic and manual ultrasound scanning for scoliosis evaluation was given in this study. The results of this study show that scoliosis deformity angles measured on images produced using a robotic approach are comparable to those obtained with manual ultrasound scanning. The benefits of employing a robotic approach were explored. This study paves the way for further research into the robotic approach on a larger group of participants with varying degrees of scoliosis severity, as well as eventual commercialisation.

ACKNOWLEDGEMENTS

The authors thank Stella Chui Yi Chan and Kelly Lai for their contribution in managing the recruitment of the subjects.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Maria Victorova  <https://orcid.org/0000-0002-1655-1953>

David Navarro-Alarcon  <https://orcid.org/0000-0002-3426-6638>

REFERENCES

- Cobb J. Outline for the study of scoliosis. *Instr Course Lect AAOS*. 1948;5:261-275.
- Hoffman DA, Lonstein JE, Morin MM, Visscher W, Harris BS, III, Boice JD, Jr. Breast cancer in women with scoliosis exposed to multiple diagnostic x rays. *JNCI (J Natl Cancer Inst)*. 1989;81(17):1307-1312. <https://doi.org/10.1093/jnci/81.17.1307>
- Levy AR, Goldberg MS, Hanley JA, Mayo NE, Poitras B. Projecting the lifetime risk of cancer from exposure to diagnostic ionizing radiation for adolescent idiopathic scoliosis. *Health Phys*. 1994;66(6):621-633. <https://doi.org/10.1097/00004032-199406000-00002>
- Zheng Y-P, Lee TT-Y, Lai KK-L, et al. A reliability and validity study for Scolioscan: a radiation-free scoliosis assessment system using 3D ultrasound imaging. *Scoliosis Spinal Disord*. 2016;11:1-15. <https://doi.org/10.1186/s13013-016-0074-y>
- Wang Q, Li M, Lou EH, Wong MS. Reliability and validity study of clinical ultrasound imaging on lateral curvature of adolescent idiopathic scoliosis. *PLoS One*. 2015;10(8):0135264. <https://doi.org/10.1371/journal.pone.0135264>
- Brink RC, Wijdicks SP, Tromp IN, et al. A reliability and validity study for different coronal angles using ultrasound imaging in adolescent idiopathic scoliosis. *Spine J*. 2018;18(6):979-985. <https://doi.org/10.1016/j.spinee.2017.10.012>
- Chen W, Le LH, Lou EH. Ultrasound imaging of spinal vertebrae to study scoliosis. *Open J Acoust*. 2012;2(3):95-103.
- Lai KK-L, Lee TT-Y, Lee MK-S, Hui JC-H, Zheng Y-P. Validation of Scolioscan air-portable radiation-free three-dimensional ultrasound imaging assessment system for scoliosis. *Sensors*. 2021;21(8):2858. <https://doi.org/10.3390/s21082858>
- Cheung C-WJ, Zhou G-Q, Law S-Y, Mak T-M, Lai K-L, Zheng YP. Ultrasound volume projection imaging for assessment of scoliosis. *IEEE Trans Med Imaging*. 2015;34(8):1760-1768. <https://doi.org/10.1109/TMI.2015.2390233>
- Jiang WW, Zhou GQ, Lai KL, et al. A fast 3-D ultrasound projection imaging method for scoliosis assessment. *Math Biosci Eng*. 2019;16(3):1067-1081. <https://doi.org/10.3934/mbe.2019051>
- Vo QN, Le LH, Lou E. A semi-automatic 3D ultrasound reconstruction method to assess the true severity of adolescent idiopathic scoliosis. *Med Biol Eng Comput*. 2019;57(10):2115-2128. <https://doi.org/10.1007/s11517-019-02015-9>
- Victorova M, Navarro-Alarcon D, Zheng Y. 3D ultrasound imaging of scoliosis with force-sensitive robotic scanning. In: *2019 Third IEEE International Conference on Robotic Computing (IRC)*; 2019:262-265. <https://doi.org/10.1109/IRC.2019.00049>
- Tirindelli M, Victorova M, Esteban J, et al. Force-ultrasound fusion: bringing spine robotic-us to the next "level". *IEEE Rob Autom Lett*. 2020;5(4):5661-5668. <https://doi.org/10.1109/lra.2020.3009069>
- Yang C, Jiang M, Chen M, Fu M, Li J, Huang Q. Automatic 3-D imaging and measurement of human spines with a robotic ultrasound system. *IEEE Trans Instrum Meas*. 2021;70:1-13. <https://doi.org/10.1109/TIM.2021.3085110>
- Li K, Xu Y, Wang J, Ni D, Liu L, Meng MQ-H. Image-guided navigation of a robotic ultrasound probe for autonomous spinal sonography using a shadow-aware dual-agent framework. arXiv preprint arXiv:2111.02167. 2021.
- Victorova M, Lee MK-S, Navarro-Alarcon D, Zheng Y. Follow the curve: robotic ultrasound navigation with learning-based localization of spinous processes for scoliosis assessment. *IEEE Access*. 2022;10:40216-40229. <https://doi.org/10.1109/access.2022.3165936>
- Karlsson N, Bernardo E, Ostrowski J, Goncalves L, Pirjanian P, Munich M. The vSLAM algorithm for robust localization and mapping. In: *IEEE International Conference on Robotics and Automation (ICRA)*; 2005:24-29. <https://doi.org/10.1109/ROBOT.2005.1570091>
- Mathiassen K, Fjellin JE, Glette K, Hol PK, Elle OJ. An ultrasound robotic system using the commercial robot UR5. *Front Robot AI*. 2016;3:1. <https://doi.org/10.3389/frobt.2016.00001>
- Boman K, Olofsson M, Forsberg J, Boström S-Å. Remote-controlled robotic arm for real-time echocardiography: the diagnostic future for patients in rural areas? *Telemed e-Health*. 2009;15(2):142-147. <https://doi.org/10.1089/tmj.2008.0079>



20. Virga S, Zetting O, Esposito M, et al. Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms. In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE; 2016:508-513.
21. Navarro-Alarcon D, Li P, Yip HM. Energy shaping control for robot manipulators in explicit force regulation tasks with elastic environments. In: *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*; 2011:4222-4228. <https://doi.org/10.1109/IROS.2011.6094641>
22. Navarro-Alarcon D, Liu Y-H, Romero JG, Li P. Energy shaping methods for asymptotic force regulation of compliant mechanical systems. *IEEE Trans Control Syst Technol*. 2014;22(6):2376-2383. <https://doi.org/10.1109/TCST.2014.2309659>
23. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016:770-778.
24. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *International Conference on Machine Learning*. PMLR; 2015:448-456.
25. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1097-1105.
26. Xiao B, Wu H, Wei Y. Simple baselines for human pose estimation and tracking. In: *Proceedings of the European Conference on Computer Vision (ECCV)*; 2018:466-481.
27. Aylott CEW, Puna R, Robertson PA, Walker C. Spinous process morphology: the effect of ageing through adulthood on spinous process size and relationship to sagittal alignment. *Eur Spine J*. 2012;21(5):1007-1012. <https://doi.org/10.1007/s00586-011-2029-6>
28. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
29. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Upper Saddle River, NJ : Pearson/Prentice Hall, 2009;892:11-15.
30. Langensiepen S, Semler O, Sobottke R, et al. Measuring procedures to determine the Cobb angle in idiopathic scoliosis: a systematic review. *Eur Spine J*. 2013;22(11):2360-2371. <https://doi.org/10.1007/s00586-013-2693-9>
31. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int J Nurs Stud*. 2011;48(6):661-671. <https://doi.org/10.1016/j.ijnurstu.2011.01.016>
32. Scholten P, Veldhuizen A. Analysis of Cobb angle measurements in scoliosis. *Clin BioMech*. 1987;2(1):7-13. [https://doi.org/10.1016/0268-0033\(87\)90039-8](https://doi.org/10.1016/0268-0033(87)90039-8)
33. Tanure MC, Pinheiro AP, Oliveira AS. Reliability assessment of Cobb angle measurements using manual and digital methods. *Spine J*. 2010;10(9):769-774. <https://doi.org/10.1016/j.spinee.2010.02.020>
34. Chow LS, Paramesran R. Review of medical image quality assessment. *Biomed Signal Process Control*. 2016;27:145-154. <https://doi.org/10.1016/j.bspc.2016.02.006>
35. Li K, Xu Y, Meng MQ-H. An overview of systems and techniques for autonomous robotic ultrasound acquisitions. *IEEE Trans Med Robot Bionics*. 2021;3(2):510-524. <https://doi.org/10.1109/tmr.2021.3072190>

How to cite this article: Victorova M, Lau HHT, Lee TT-Y, Navarro-Alarcon D, Zheng Y. Comparison of ultrasound scanning for scoliosis assessment: robotic versus manual. *Int J Med Robot*. 2022;e2468. <https://doi.org/10.1002/rcs.2468>