Journal of Digital Imaging

Validity and Reliability of Active Shape Models for the Estimation of Cobb Angle in Patients with Adolescent Idiopathic Scoliosis

Shannon Allen, BSc, Eric Parent, PhD, Maziyar Khorasani, BSc, Douglas L. Hill, MBA, Edmond Lou, PhD, and James V. Raso, MASc

Choosing the most suitable treatment for scoliosis relies heavily on accurate and reproducible Cobb angle measurement from successive radiographs. The objective is to reduce variability of Cobb angle measurement by reducing user intervention and bias. Custom software to increase automation of the Cobb angle measurement from posteroanterior radiographs was developed using active shape models. Validity and reliability of the automated system against a manual and semiautomated measurement method was conducted by two examiners each performing measurements on three occasions from a test set ($N=22$). A training set ($N=47$) of radiographs representative of curves seen in a scoliosis clinic was used to train the software to recognize vertebrae from T4 to L4. Images with a maximum Cobb angle between 20° and 50°, excluding surgical cases, were selected for training and test sets. Automated Cobb angles were calculated using best-fit slopes of the detected vertebrae endplates. Intraclass correlation coefficient (ICC) and standard error of measurement (SEM) showed high intraexaminer (ICC>0.90, SEM 2°–3°) and interexaminer (ICC>0.82, SEM 2°–4°), but poor intermethod reliability (ICC=0.30, SEM 8°–9°). The automated method underestimated large curves. The reliability improved (ICC=0.70, SEM 4°–5°) with exclusion of the four largest curves (>40°) in the test set. The automated method was reliable for moderate-sized curves, and did detect vertebrae in larger curves with a modified training set of larger curves.

KEY WORDS: Active shape models, Cobb angle scoliosis, automated measures, reliability

INTRODUCTION

Scoliosis is the most common cause of spinal deformity in adolescents. It involves lateral deviation of the spine in the frontal plane, often accompanied by rotation of individual vertebrae. This rotation causes noticeable asymmetries of the trunk such as uneven shoulders, a prominent shoulder blade, a waist crease, or side-to-side leaning. Approximately 2–4% of adolescents have some form of scoliosis, but only about 0.5% will require treatment.[1] In 80% of the cases, scoliosis develops for unknown reasons (idiopathic) in otherwise healthy adolescents.[2] Treatment consists of observation, bracing, or surgery for progressive curves in growing teens. Before initiating bracing or contemplating surgery, which are both major undertakings, the status of the scoliotic curve must be established as stable, slowly progressing, or requiring intervention. Assessment of the severity of the scoliosis involves measuring the largest curvature of the spine, known as the Cobb angle (see Fig. 1).[3] The Cobb angle is determined on the back-to-front radiograph by selecting the most tilted vertebra at the top and bottom of the curve. The summation of the angles of these two vertebrae relative to the horizontal is the Cobb angle. In a straight spine (no scoliosis), the angle will be near zero, and has to be more than 10° to be classified as scoliosis. The error in measuring the Cobb angle consists of selecting different vertebrae that are most tilted

From the Department of Rehabilitation Technology, Capital Health, Glenrose Rehabilitation Hospital Site, 10230 111 Ave., Edmonton, Alberta, T5G 0B7, Canada.

Correspondence to: Douglas L. Hill, MBA, Department of Rehabilitation Technology, Capital Health, Glenrose Rehabilitation Hospital Site, 10230 111 Ave., Edmonton, Alberta, T5G 0B7, Canada; tel: +1-780-7358289; fax: +1-780-7357972; e-mail: dhill@cha.ab.ca

Online publication 6 March 2007
doi: 10.1007/s10278-007-9026-7

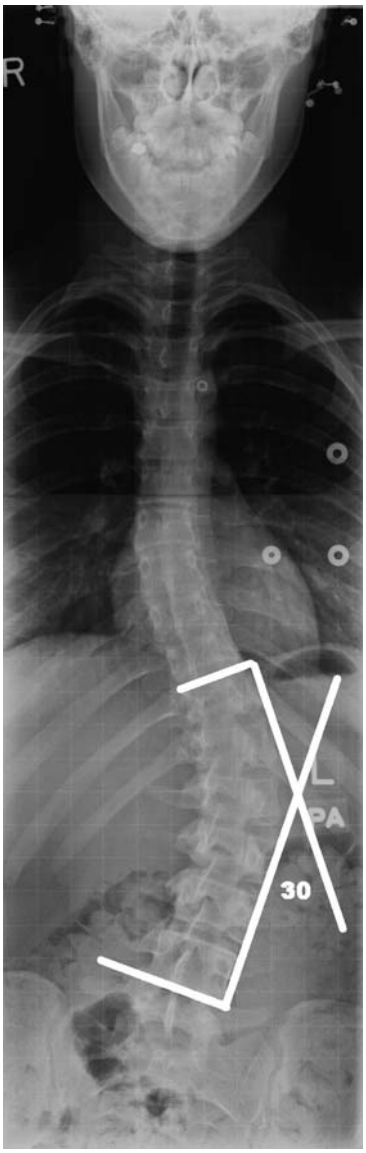

Fig 1. The Cobb angle is the measurement used for evaluation of curves in scoliosis on a PA radiograph. It is measured as the angle between the two perpendicular lines drawn from parallel lines running through the most tilted superior and inferior endplates.

and in estimating the best-fit line for the slope of the vertebrae. In this study, the Cobb angle of adolescent idiopathic scoliosis (AIS) patients using an automated method to detect the vertebrae was assessed. Adolescent idiopathic scoliosis usually becomes apparent in adolescents aged 10–16, and managing its progression and/or course of treatment relies heavily on accurate Cobb angle measurement from successive radiographs, typically taken 3–6 months apart.

When manually determining the Cobb angle from an AIS radiograph, intra- and interexaminer measurements have been reported to vary between 3°–10°.[4–6] This range is critical to the assessment of curve progression because a true increase in curvature of ≥5° over a 6-month period usually warrants some course of treatment.[4,7] To confirm its validity, the automated method was compared to a semiautomated method and the standard manual method. Unlike other studies[8,9] which have attempted to improve Cobb angle measurement either through computer-assistance or the use of a new tool (Cobbometer), this study can be externalized to a clinical scenario. End vertebrae were not preselected for measurement and the examiners used their measurement tools of choice, whereas the before-mentioned studies eliminated such variables.

One similar study used a semiautomated computerized method developed within MATLAB to measure Cobb angle.[4] The statistical tool standard error of measurement (SEM), which is utilized in this study, is the same as the technical error of measurement (TEM) or the measurement error standard deviation used by Chockalingam.[4] TEM provides an estimate of measurement error in the units of measurement of the variable, which in this case are degrees. Mean intraobserver TEM was 0.7° and the mean coefficient of reliability was 0.98, while interobserver TEM was 1.2° with a mean coefficient of reliability of 0.99. Manual measurement of the same PA radiographs produced an interobserver TEM of 1.8° and a mean coefficient of reliability of 0.8. However, this study was based only on nine scoliotic radiographs. In addition, manual measurement was performed only once, and there was no indication of how the measurement was done (whether end vertebrae were pre-selected or same protractor used by all three examiners).

The overall objectives of this study were to develop a program for a more automated measurement of Cobb angle, and then to test its validity and reliability against both a semiautomated and manual Cobb angle measurement. It was predicted that the error would be reduced in measuring the severity of the scoliosis by automating the Cobb angle measurement with a program that minimized examiner input, thus limiting the amount of bias or error attributed within and between examiners. Subsequently, this

would allow individuals with varying levels of expertise (from novice to expert) to calculate the Cobb angle with approximately the same amount of validity and reliability, regardless of their skill level or previous experience. Having an automated evaluation of Cobb angle will ultimately allow for improved understanding of curve progression and better treatment decisions.

MATERIALS AND METHODS

Active Shape Models

Developing and testing a program that allowed for the automated measurement of Cobb angles from spine radiographs was achieved in two stages. The first stage involved using a training set of radiographs to "teach" the computer program to recognize vertebrae shape and subsequently perform calculations of superior and inferior endplate slopes. Training was achieved using a custom program developed in MATLAB and C^{++} and was based on the "active shape models" algorithm described by Cootes and Lindley.[10–13] The second stage consisted of estimating and comparing the intraexaminer, interexaminer, and intermethod reliability of the automated, manual, and semiautomated method.

Active shape models (ASM) were used to detect specific vertebra on PA radiographs. Active shape models are "trained" to recognize what a vertebra looks like in a radiograph through a series of sample images. The advantage of ASM over other forms of image recognition is that they are specific to the application, and are "trained" only in the detection and recognition of a specific target object and no other. Because they are only familiar with the variation provided in the training set, ASM cannot significantly deviate from that variation, and thus can only generate shapes similar to the training set.

Training creates two models, the object shape model and the object appearance model. After training, by deforming the shape models, and comparing the image with the appearance model, the ASM algorithm iteratively converges to the target object in the image. The object shape is described by the point distribution model (PDM). During training, the boundary of the object is identified by manually digitizing n landmark points around the perimeter of the object in the image. This is done for a series of training set images, consistently placing the points in the same positions throughout the images in the training set.

The shape for each image can be then represented by a $2n$ element vector:

$$z = (x_1, \ldots, x_n, y_1, \ldots, y_n)$$

Where (x_i, y_i) are the x and y coordinates of the ith landmark point, respectively.

Using Procrustes analysis, the corresponding landmark points are matched along the training set images, and translated, rotated, and scaled for each training set shape. For this study, n=40 points per vertebra, for 13 vertebrae (from T4–L4) per radiograph were annotated in the training set, which consisted of 47 radiographs. The points were approximately evenly spaced around the four sides of the vertebra, with ten points per side. The number of pixels on each side of a vertebra examined during training is 10 and set to 30 during image search.

Principal component analysis (PCA) is then used to determine the axis along which the majority of the data, and hence the variation among the training set images, lies. Based on the predetermined level of accuracy, the most common modes of variation that describe the training set can be extracted.

The point distribution (i.e., the shape) model can then be represented by the following equation:

$$x = \overline{x} + Pb$$

Where:

$\overline{x}$	The mean of the aligned training set images
P	$2n \times t$ matrix representing the modes of variation
b	shape parameter represented by a vector of weights

The model can be deformed by varying the shape parameter b. The new shape ideally converges to the target object shape. The shape parameter b is constrained to ensure implausible shapes are not generated. For this study, the weightings have been set to 1.

The generation of shape variations is illustrated by the PDM of a partial spine in Figure 2 for the most common variations found in the training set. The ASM algorithm varies the shape within the three standard deviation limit for each mode of variation.

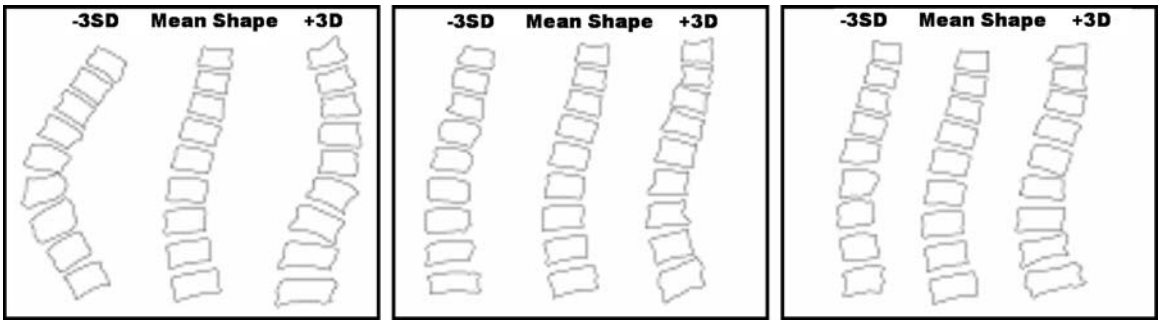

Fig 2. Three shape variations of a partial spine, altered between the three standard deviation limit for each image.

The second model generated during training is the image appearance, represented by analyzing the shade intensity around each landmark point. This process is illustrated in Figure 3. The program samples along a line normal to the intersection of two successive landmark points. The intensity profile is evaluated for all landmark points per vertebra and for all radiographs of the training set. The effects of intensity changes along the normal line are reduced by taking the derivative along the profile. Principal component analysis is then used once more to create the statistical model of the intensity profiles. In standard ASM, the profiles are normalized.

Once the object shape and appearance statistical models have been created, the ASM are used to locate the specific vertebra in the image. In this iterative scheme, an initial estimate of the vertebra shape is made by selecting initialization points to hasten convergence (see Fig. 4). Only the region around the initial estimate is examined using the appearance model. Scaling, translation, and rotation are applied to the point distribution model to improve the match. This procedure is repeated until either a maximum number of predetermined iterations occur, which in this study was 50, or convergence to a predetermined threshold (90% of the explained variance) is reached.

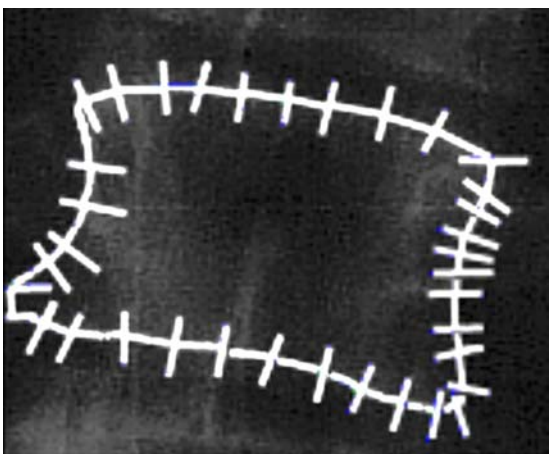

Fig 3. Generation of the image appearance by sampling the shade intensity along a line normal to the intersection of two landmark points.

Automated Cobb Angles

The radiographs of the training and test sets were analogue PA radiographs. The images chosen for the training and test set were selected from the scoliosis database whose maximum Cobb angle varied between 20° and 50°. These radiographs can be classified into five categories, which represent the majority of scoliotic curves seen at the clinic (see Table 1). Scoliosis presents most commonly with a curve bending to the right in the thoracic region (right thoracic curve). Frequently, there is a compensating left lumbar curve. Less common is a single curve either in the lumbar region that generally bends to the left or in the middle of the spine denoted as thoracolumbar.

These were digitized by taking pictures of the radiographs with a consumer 5 megapixel digital camera. The radiograph was placed on a viewbox, with the film occupying the maximal field of view. Some quality was lost in the digitizing process, making detection of the vertebrae more difficult than directly using digital radiographs. However, analogue radiographs are still more common than digital images in the clinic. This process results in a worst case scenario. Using a high-quality scanner or digital radiographs will improve the detection of features in the radio-

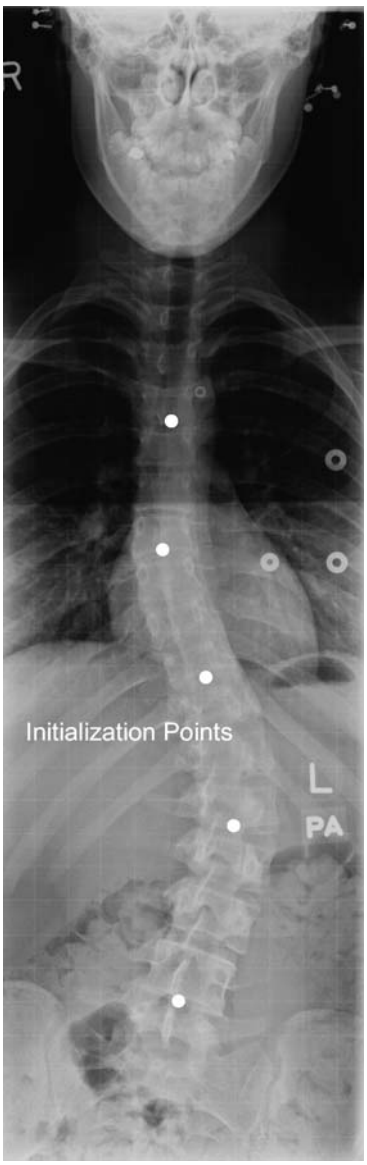

Fig 4. An illustration of the five initialization points.

graphs. In selecting training and test radiographs, the exclusion criteria were those that included or showed: congenital deformities, poor imaging clarity, very large curves, major left thoracic curves, or post-surgical radiographs. These radiographs would be less compatible with the ASM, which are trained to recognize relatively rectangular vertebrae, while the excluded radiographs listed above have a higher likelihood of irregularly shaped vertebrae. A training set of 47 images was created, while the test set had 22 images. An additional training set comprising only of the 10 largest curves in the original training set was used for large curves.

The training set was used with test radiographs to measure the automated Cobb angle of those images. These measurements were obtained using a custom software program that can read images in .TIF, .JPEG, or dicom format. After loading the training set and the test image, the program prompts the user to select five initialization points on the test image. Initialization points on every third vertebra were selected, beginning with T4 and ending with L4, and placing them in the middle of the superior endplate, with both examiners choosing which vertebrae was T4 without any guidance (see Fig. 4). These points allowed the program to converge to an initial estimate of the shape of the test image (see Fig. 5). After the program went through a maximum of 50 iterations of shape estimation, the angles (in degrees) of the superior and inferior endplates relative to the horizontal of each vertebra were displayed and stored. The Cobb angle was calculated by adding the absolute value of the local maximum and minimum angle values. Selection of initialization points for the test images demonstrated that the "automated" method was not fully automated and still required examiner input. However, the program required less intervention than previous studies have reported, thus the program was classified as "automated."

Table 1. Classification of Training and Test Set Radiographs

	Right (R) Thoracic	Left (L) Lumbar	R. Thoracic with L. Lumbar	Thoracolumbar	Small Curves	Total
Number of training images	12 (26%)	6 (13%)	15 (31%)	5 (11%)	9 (19%)	47
Number of test images	7 (32%)	3(13.6%)	7 (32%)	2 (9%)	3 (13.6%)	22

Scoliosis presents most commonly with a curve bending to the right in the thoracic region (right thoracic curve). Frequently is a compensating left lumbar curve. Less common is a single curve either in the lumbar region that generally bends to the left or in the middle of the spine denoted as thoracolumbar.

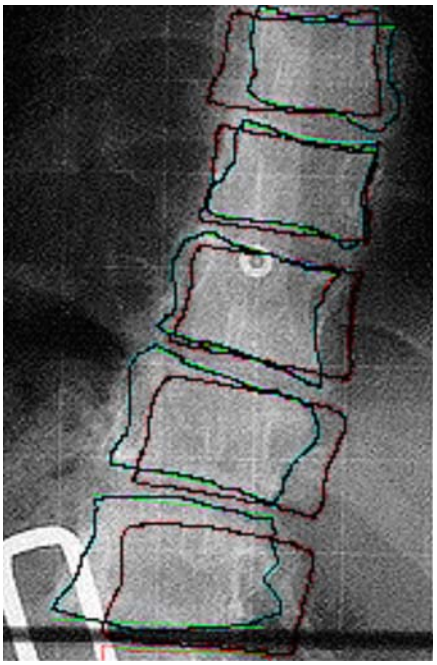

Fig 5. Initial estimate of the shape of the vertebra.

Clinical Test to Semiautomated and Manual Cobb Angle Methods

The automated method was validated by comparing it to a semiautomated and manual method using the same test images. The semiautomated method required the selection of four landmark points by the examiner; two points each for the most superior and inferior vertebrae that showed the greatest tilt towards the concavity of the spine (see Fig. 6). The landmark points are selected by marking positions that best describe the endplate orientation relative to the horizontal. The examiner could zoom in on the digitized radiograph to better place the points using a computer mouse. Based on these points, an upper and lower angle between the endplate and the horizontal was calculated, and the absolute value of their sum used as the semiautomated Cobb angle. For all three methods, two examiners (trained novice and intermediate skill levels) performed the Cobb angle calculation three times for the 22 test images, with trials spaced 2 to 3 days apart for a given method. The radiographs were presented in random order, with no record showing previous measurements.

For the manual method, each examiner was given computer print outs of the radiographs. They measured the Cobb angles using their choice of the most tilted end vertebrae and preference for measuring the angle between these vertebrae. The print outs were a scaled version of the analogue radiographs, with an approximate 3:1 reduction.

Analyses

Statistical analysis using Statistical Package for Social Sciences (SPSS) was used to calculate $ICC_{(2,1)}$[14] values with 95% confidence intervals (95% CI) to estimate the reliability of the intra-

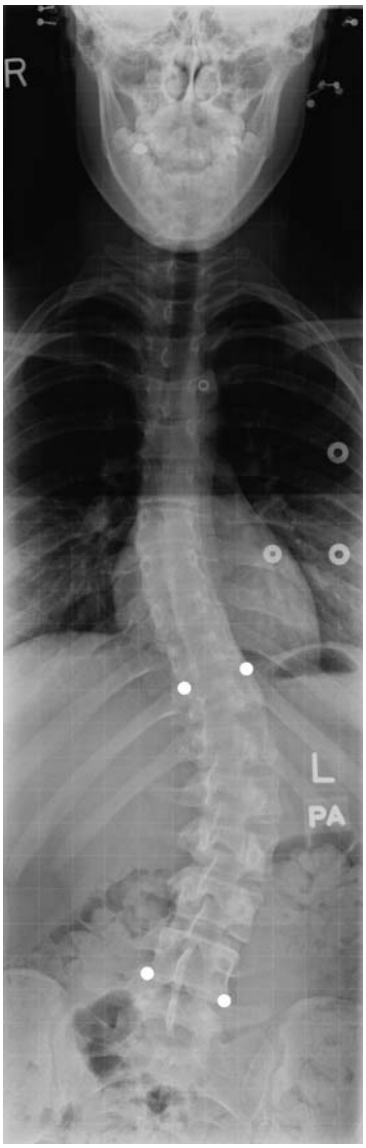

Fig 6. The semiautomated method involves the selection of four landmark points. Two each for the most superior and inferior endplates that tilt towards the concavity of the curve.

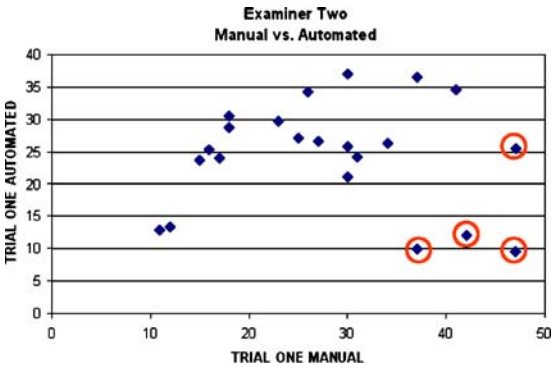

Fig 7. Scatter plot of intermethod results for examiner two's trial one. Points surrounded by circles indicate outliers.

examiner, interexaminer, and intermethod measurements. The value of ICC can vary between 0 and 1, with 1 indicating perfect reliability, 0.70 being recommended for inferences about groups, and values between 0.90 and 0.95 being recommended to make inferences for an individual.[15] The mean, standard deviation, and standard error of measurement $(\text{SEM} = \text{SD}^*\sqrt{(1 - \text{ICC})})$[16] were also calculated for each method and trial. Scatter plots were examined for outliers (see Fig. 7).

RESULTS

Selection of the most tilted vertebrae for the determination of the Cobb angle had 70% agreement for the manual and semiautomated methods and 89% agreement for the automated method. Initial scatter plots of the three trials of each method against one another revealed a strong linear correlation between the trials (intraexaminer). Statistical Package for Social Sciences produced ICC values ≥0.9, and a 95% CI with lower limits ≥0.83 and upper limits ≤0.98 for the intraexaminer reliability using any of the three methods (see Table 2). All of the 95% CI of the intraexaminer reliability estimates overlapped.

Similar results were found for the interexaminer reliability using any of the three methods (ICC values ≥0.82 and a 95% CI with lower limits ≥0.62 and upper limits ≤0.98, see Table 3). The 95% CI of the interexaminer reliability estimates also overlapped.

However, plotting the three methods against one another showed near linear correlation between the examiners' results for the manual method versus semiautomated method, but noticeably poorer correlation between the manual versus automated, and semiautomated versus automated method (ICC values of approximately 0.30 and a 95% CI with lower limits ≥0.0 and upper limits ≤0.6, see Fig. 7, Table 4). Scatter plots of examiner one's intermethod trials, and examiner two's trials two and three are not shown because they produced similar results. In addition, the 95% CI of the intermethod reliability estimates overlapped.

From the intermethod plots, four outliers with large Cobb angles as measured by the manual and

Table 2. Intraexaminer Reliability (ICC), 95% Confidence Interval (CI), Standard Error of Measurement (SEM), Mean, and Standard Deviation (SD) for Each Method and Both Examiners

Method	Examiner	ICC	CI	SEM	Mean	SD
Manual	1	0.95	(0.91, 0.98)	2.09	26.74	9.53
	2	0.95	(0.90, 0.98)	2.28	29.02	10.31
Semiautomated	1	0.94	(0.88, 0.97)	2.97	33.45	12.19
	2	0.92	(0.85, 0.96)	3.39	34.41	12.33
Automated	1	0.91	(0.84, 0.96)	2.30	24.67	7.86
	2	0.96	(0.92, 0.98)	1.71	24.11	8.39

Table 3. Interexaminer Reliability (ICC), 95% Confidence Interval (CI), Standard Error of Measurement (SEM), Mean, and Standard Deviation (SD) for Each Method and Each Trial

Method	Trial	ICC	CI	SEM	Mean	SD
	1	0.95	(0.88, 0.98)	2.31	27.00	10.04
	2	0.93	(0.83, 0.97)	3.10	28.32	10.24
Manual	3	0.94	(0.86, 0.97)	4.09	28.32	9.74
	1	0.89	(0.76, 0.96)	3.95	34.57	12.19
	2	0.93	(0.85, 0.97)	3.10	34.02	11.96
Semiautomated	3	0.90	(0.77, 0.96)	4.09	33.20	12.76
	1	0.96	(0.90, 0.98)	1.63	24.46	8.05
	2	0.83	(0.63, 0.92)	3.37	24.74	8.07
Automated	3	0.94	(0.85, 0.97)	2.11	23.97	8.35

semiautomated methods were underestimated by the automated method. They were eliminated and the intermethod ICCs recalculated with the remaining 18 test cases. As expected, there was a sizable improvement in the ICC values when comparing the intermethod coefficients for each examiner (ICC values of approx. 0.70 and a 95% CI with lower limits ≥0.4 and upper limits ≤0.9, see Table 5).

The standard error of measurement (SEM) was lower compared to the standard manual technique. Including the outliers, the SEM for intraexaminer measurements for any method was between 2°–3°, and the SEM interexaminer measurements for each method was between 2°–4° (see Tables 2 and 3). The SEM for intermethod measurements for each examiner was approximately 8°–9°, improving to 4°–5° with the outliers eliminated (see Tables 4 and 5).

DISCUSSION

Intraexaminer and interexaminer ICCs showed high reliability compared to intermethod ICC values, which were quite low, but improved upon the removal of the four outliers representing large curves. Part of the automated method's limitation may be because the apparent shape of vertebrae can change with increasing severity of scoliosis. Various conditions including osteoporosis, obesity, complex deformity, and previous surgery with implants may adversely affect the shape and hence the recognition of the vertebrae as well. With mild or moderate scoliosis, the vertebrae generally appear rectangular. As the scoliosis becomes more severe, the vertebrae may become more irregularly shaped as seen in Figure 2. Because the training set was taught to recognize rectangular vertebrae (representing mild to moderate scoliosis), vertebrae that are deformed (representing more severe scoliosis) may not be recognized, and therefore the automated program may not detect these vertebrae. The detection process of vertebrae by ASM are limited to the content of the training set, and because the training set did not include many radiographs with large curves, recognition of large curves in the test set was more difficult for the ASM. To demonstrate this, a second training set of ten radiographs representing the largest curves from

Table 4. Intermethod Reliability (ICC), 95% Confidence Interval (CI), Standard Error of Measurement (SEM), Mean, and Standard Deviation (SD) for Both Examiners and Each Trial

Trial	Examiner	ICC	CI	SEM	Mean	SD
	1	0.32	(0.06, 0.60)	8.64	28.11	10.51
T1	2	0.35	(0.09, 0.62)	9.34	29.24	11.58
	1	0.33	(0.07, 0.60)	8.93	29.25	10.91
T2	2	0.31	(0.05, 0.58)	9.02	28.80	10.84
	1	0.25	(−0.01, 0.53)	9.22	27.51	10.62
T3	2	0.30	(0.04, 0.58)	9.54	29.49	11.43

Table 5. Automated Method Revised for Intermethod Reliability (ICC), 95% Confidence Interval (CI), Standard Error of Measurement (SEM), Mean, and Standard Deviation (SD)

Trial	Examiner	ICC	CI	SEM	Mean	SD
T1	1	0.71	(0.48, 0.87)	4.28	26.53	7.91
	2	0.69	(0.46, 0.86)	5.08	27.59	9.13
T2	1	0.71	(0.48, 0.87)	4.38	27.39	8.08
	2	0.74	(0.53, 0.88)	4.13	27.18	8.14
T3	1	0.68	(0.44, 0.85)	4.34	25.84	7.70
	2	0.70	(0.47, 0.86)	4.50	27.62	8.21

the original training set was created. The four test radiographs with large Cobb angles were then tested against this new training set. Some improvement was observed in the automated Cobb angle measurement for two of the test radiographs, but showed less improvement for the other two. Therefore, creating multiple training sets for the various curve types may improve Cobb angle measurement, but it is likely that adjustment of the computer program's threshold limits will also be needed. In addition, although classified as "automated", our automated method does require user input, and therefore could be considered not truly automated. However, there are no known studies that have advanced as far in automating Cobb angle measurement as this study attempted to do, therefore calling the program "automated" seemed appropriate.

The high reliability demonstrated by the intraexaminer ICC values suggests that the automated method is effective regardless of the experience of the examiner. Because one of the examiners was a trained novice, the high reliability obtained suggests that the automated method could produce similar results in a clinical environment where those using the automated measurement program may have little prior experience. A period of 2–3 days between repeated measurements ensured that the examiners were not simply recalling their previous measurement. Examiners may have performed more than one measurement method on any certain day, but did not perform the same measurement method twice in 1 day.

Interexaminer ICCs were similar to the intraexaminer ICCs, which suggests that the semiautomated and automated methods are user friendly because the trained novice's results compared quite well to the intermediate examiner. The examiners differed in their level of experience of manual measurement, but were relatively similar in using the computer programs for conducting the semiautomated and automated measurements. Both had prior practice with the computer programs (MATLAB and the custom program) using sample radiographs, practicing on approximately five radiographs each.

Intermethod reliability was quite poor compared to intraexaminer and interexaminer reliability. In plotting the three methods against one another, it was obvious that the automated method produced different results compared to the other methods. Plots of manual versus automated, or semiautomated versus automated revealed four main outliers. These outliers were the four largest curves (>40° Cobb angle) and were single curves. The training set comprised of a subset of the 10 largest curves in the original training set allowed detection of these vertebrae. The 18 remaining test radiographs were represented by more moderate curves.

In this study, possible sources of intrinsic error were not controlled to increase externalization of our results such as selection of vertebrae used in calculating the Cobb angle, which other studies have eliminated. Both Rosenfeldt and Shea[8,9] preselected the end vertebrae before all measurements (both manual and computer or Cobbometer measurements), whereas this study allowed the examiners to choose their own. In addition, the examiners used their own protractor and radiograph markers unlike the prior studies where this was standardized. While this does add to the variability in Cobb angle measurement, it is a more realistic assessment of a clinical scenario in

which different examiners evaluate a patient's successive radiographs with various instruments. Although it would be better to have one examiner with the same measurement tools follow a certain patient, it is not the common practice.

Although Chockalingam's[4] mean TEM for both intra- and interobserver were noticeably better than this study's intra- and interexaminer, and intermethod SEM, there are various differences that may account for this. Firstly, the selection of two points per line, with a minimum of eight lines (16 points total) means that there could be less variability for the computer program to recognize, compared to this study which only required the selection of five points. However, the selection of five points requires less user intervention, and therefore a more automated measurement of the Cobb angle. Chockalingam[4] estimated angles on only nine scoliotic radiographs. The skill level of the observers and methodology in this study is unclear.

Future studies can be improved in several ways. Considering the range of error for manual measurement is between 3°–10°, these results suggest the automated method has potential, and may be of greater value with further refinement. Modification of the shape parameter and adjusting the scaling, translation, and rotation constraints may improve matching of larger curves. Using separate training sets for different curve types may improve detection of vertebrae. The use of higher quality digital radiographs would allow the individual vertebrae to be more distinct, and therefore produce even better results for all methods.

The Cobb angle can only represent the scoliotic deformity in a single plane, whereas scoliosis is truly three-dimensional (3D), but 3D parameters have not yet been developed that are widely accepted. Cobb angle measurement remains the most commonly used method for monitoring the progression of scoliotic curves. Considering that a variety of individuals (physicians, clinicians, medical students) with different skills and experience measure the Cobb angle in a scoliosis clinic, providing an automated method for angle measurement will help reduce intra- and interexaminer error. Reduced error ultimately leads to better assessment of the patient's need for treatment and allows for improved determination of whether the curve has truly changed over time.

CONCLUSION

Adolescent idiopathic scoliosis is both manageable and treatable, with early detection allowing for the most flexibility in making treatment decisions. However, choosing the most suitable treatment relies partly on the accuracy of Cobb angle measurement. Because treatment decisions are heavily based on how much the Cobb angle has changed, reliable Cobb angle measurement is necessary. This study attempted to develop an automated Cobb angle measurement program, requiring only initiation points, allowing various health professionals to assess the angle regardless of their prior experience and training. The automated method worked best for moderate curves.

ACKNOWLEDGEMENTS

This work was supported by the University Hospital Foundation Medical Research Competition.

REFERENCES

1. Lonstein JG: Adolescent idiopathic scoliosis. Lancet 344:1407–1412, 1994

2. Scoliosis Research Society. In Depth Review of Scoliosis. Available at: http://www.srs.org/patients/

3. Cobb JR: Outline for the study of scoliosis. Am Acad Orthop Surg Inst Course Lect 5:261–275, 1948

4. Chockalingam N, Dangerfield PH, Giakas G, Cochrane T, Dorgan JC: Computer-assisted Cobb measurement of scoliosis. Eur Spine J 11:353–357, 2002

5. Loder RT, Spiegel D, Gutknecht S, Kleist K, Ly T, Mehbod A: The assessment of intraobserver and interobserver error in the measurement of noncongenital scoliosis in children ≤10 years of age. Spine 29:2548–2553, 2004

6. Mior SA, Kopansky-Giles DR, Crowther ER, Wright JG: A comparison of radiographic and electrogoniometric angles in adolescent idiopathic scoliosis. Spine 21:1549–1555, 1996

7. Cobb JR: Outlines for the study of scoliosis measurements from spinal roentgenograms. Phys Ther 59:764–765, 1948

8. Rosenfeldt MP, Harding IJ, Hauptfleisch JT, Fairbank JT: A comparison of traditional protractor versus Oxford Cobbometer radiographic measurement—intraobserver measurement variability for Cobb angles. Spine 30:440–443, 2005

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

10. Cootes TF, Hill A, Taylor CJ, Haslam J: The use of active shape models for locating structures in medical images. Image Vis Comput 12:355–366, 1994

11. Cootes TF, Taylor CJ, Lanitis A: Active shape models: Evaluation of a multi-resolution method for improving image search. Proceedings of the British Machine Vision Conference, pp 327–336, 1994

12. Cootes TF, Taylor CJ, Cooper DH, Graham J: Active shape models—their training and application. Comput Vis Image Underst 61:38–59, 1995

13. Lindley K: Model based interpretation of lumbar spine radiographs. MSc Thesis, University of Manchester, 1992

14. Krebs DE: Declare your ICC type [letter]. Phys Ther 66:1431, 1986

15. Scientific Advisory Committee of the Medical Outcomes Trust: Assessing health status and quality-of-life instruments: Attributes and review criteria. Qual Life Res 11:193–205, 2002

16. Streiner DL, Norman GR: Health Measurement Scales: A Practical Guide to their Development and Use, 2nd edition. New York, New York: Oxford Medical Publications:106–119, 1995