Application of Ultrasound to Measure Coronal Curvature and Vertebral Rotation in Adolescent Idiopathic Scoliosis

by

Wei Chen

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biomedical Engineering
University of Alberta

Abstract

Adolescent idiopathic scoliosis (AIS) is a three-dimensional (3D) spinal deformity occurring during adolescence with no known causes. Both the coronal curvature and vertebral axial rotation (VAR) are important parameters to assess the severity, predict the progression and evaluate the outcomes of AIS. Currently, 2D radiography is the standard imaging method to diagnose AIS and the Cobb angle is the clinical practice to measure the severity of AIS. However, the radiographic method exposes the patients to harmful ionizing radiation and has limitation to reveal the true nature of scoliosis. Ultrasound as a non-ionizing radiation method has been proposed in this PhD study to measure the coronal curvature and VAR in children with AIS.

To optimize the ultrasound set up for spine imaging applications, experiments were performed to investigate the optimum configurations. Also, according to the ultrasound theory, the tissue bone interface can provide strong reflection signals when the surface is relatively flat. Therefore, after scanning a spinal phantom, the spinous processes (SP), laminae and transverse processes could be recognized on both the coronal and transverse views of the ultrasound images. Among these three landmarks, the center of lamina (COL) method was developed and reported to be the best estimation method for the coronal curvature measurement of scoliosis. The intra- and inter-observer reliabilities on both the *in-vitro* (a cadaver spine phantom) and *in-vivo* studies (26 Children with AIS) were found to be high (Intraclass correlation coefficient (ICC): >0.87, mean absolute

difference (MAD): 1.3°-4.1°). In addition, the coronal curvature measurements using the COL method showed high agreement and small deviation (ICC: 0.92-0.96, MAD: 1.7°-2.9°) compared with the clinical record of the Cobb angle from the local scoliosis clinic which was obtained on the same day.

To measure the VAR on the ultrasound transverse images, the COL method was used. *In-vitro* and *in-vivo* studies were then performed and the results demonstrated that the intra- and inter-observer reliabilities were high (ICC: 0.91-0.99, MAD: 0.3°-0.9°). The *in-vitro* study also showed that the VAR measurements from the ultrasound image using the COL method were more accurate than the VAR measurements from the radiographs using the Stokes method.

To improve the quality of the image and reduce the human measurement errors on both the coronal curvature and VAR measurements using ultrasound, a semi-automatic measurement program was developed. The program includes three parts: the pre-processing part processed the original ultrasound data to improve the quality of the ultrasound images by reducing the sparkle noise using the wavelet soft threshold method and improve the processing time by reducing the data size; the image reconstruction part generated the coronal and transverse images for measurements; and the semi-automatic measurement part required the operators to point out the laminae, and then the program segmented the laminae from the background using optimum global thresholding based on the Otsu's method, determined the centers of lamina more precisely, and automatically calculated the coronal curvature and the VAR measurements. The reliability and

validity of the measurements were investigated on the *in-vitro* and *in-vivo* data. The coronal curvature measurements had high intra-observer reliability (ICC: 0.85-0.98, MAD: 1.4°-2.4°). However, only fair inter-observer reliability was obtained for the *in-vivo* data (ICC: 0.76, MAD: 3.4°). The VAR measurements by the program showed high intra- and inter-observer reliabilities (ICC: >0.94, MAD: 0.2°-0.9°).

This thesis reported that 1) the ultrasound imaging method could be used to assess the coronal curvature and VAR of AIS; 2) a new center of lamina (COL) method was developed and validated to measure the coronal curvature and VAR on ultrasound images reliably; 3) a semi-automatic program was developed to improve image quality and reduce the human measurement errors for both of the coronal curvature and VAR measurements.

Preface

This thesis is an original work by Wei Chen. The research project, of which this thesis is a part, received research ethics approval from the Health Ethics Board of the University of Alberta, with the Project Name "Using ultrasound to assess spinal deformity for AIS", No. Pro00005707, starting January 22, 2010.

Part of the materials in Chapter 3 of this thesis was published in 'Chen W, Lou EHM, Le LH (2011) Using ultrasound imaging to identify landmarks in vertebra models to assess spinal deformity. Conf Proc IEEE Eng Med Biol Soc:8495-8498'. I was responsible for conducting the experiment, data collection, data analysis as well as the 4-page abstract preparation. Dr. Lou and Dr. Le gave the guidance of the experimental set up, methodology and edited the submitted abstract.

Part of the materials in Chapter 3 of this thesis was published in 'Chen W, Le L, Lou E (2012) Ultrasound imaging of spinal vertebrae to study scoliosis. Open J Acoustics 2:95-103'. I was responsible for both the *in-vitro* and *in-vivo* experiments, data collection, and the manuscript preparation. Dr. Lou and Dr. Le provided supervision on experimental setup and edited the submitted manuscript.

Part of the materials in Chapter 5 of this thesis was published in 'Chen W, Lou EM, Zhang P, Le L, Hill D (2013) Reliability of assessing the coronal curvature of children with scoliosis by using ultrasound images. Journal of Children's Orthopaedics 7:521-529'. I contributed to the experiment design, data collection, and performed the measurements, analyzed the results and wrote the manuscript. Both Dr. Lou and Dr. Le assisted the design of the experiment, and edited the manuscript. Dr. Lou also involved in coronal curvature measurements. Ms. Zhang assisted the design of the experiment and data acquisition, performed the measurements and wrote part of the manuscript. Mr. Hill edited the manuscript.

Part of the materials in Chapter 7 of this thesis is included in the abstract which has been accepted for the 5th International Conference on the Development of Biomedical Engineering, Vietnam, June 16-18th, 2014, as 'Chen W, Lou EM, Le L (2014) A reliable semi-automatic program to measure the vertebral rotation using the center of lamina for adolescent idiopathic scoliosis'. A 4-page abstract will be published by Springer in the IFMBE Proceedings Series. I developed the semi-automatic program, performed the measurements, analyzed the results, and wrote the abstract. Dr. Lou and Dr. Le assisted the abstract preparation and edited the abstract.

Dedication

To my beloved husband, Chao

Thank you for your love and support

And my little boy, Daryl

Although you delayed my graduation

You made me stronger

Acknowledgements

I would like to thank my supervisors Dr. Edmond Lou and Dr. Lawrence H. Le for their patience, guidance and support during my PhD research work and to prepare this thesis. I would like to thank Dr. Eric Parent and Dr. Samer Adeeb for providing their insightful advices during my research work.

I am grateful to Dr. Douglas Hedden, Dr. Marc Moreau and Dr. James Mahood for supporting patient's recruitment during their busy scoliosis clinics.

I would like to thank the China Scholarship Council for the generous financial support over the years. I would like to thank the Faculty of Graduate Studies and Research for awarding the Provost Doctoral Entrance Award. I would also like to thank the financial support from the Department of Radiology and Diagnostic Imaging, National Sciences and Engineering Research Council of Canada, Women and Children Health Research Institute, and Alberta Health Services. The financial support from these sources made me be able to concentrate my PhD work without worrying about funding.

I would like to thank my colleagues and friends: Rui, Chan, Hongjiang, Min, Tho, Cuong and many others for their support and help during my graduate work.

Last, but not least, I would like to thank my parents and my husband, Yu-Fu,Yi, and Chao, for their love and support during the course of my PhD work.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2D	Two dimensional
3D	Three dimensional
AIS	Adolescent idiopathic scoliosis
CI	Confidence interval
COL	Center of lamina
COP	Center of pedicle
CR	Computed radiography
CT	Computed tomography
DDF	Dynamic depth focusing
ICC	Intraclass correlation coefficient
IS	Idiopathic scoliosis
LS	Laser scanner
MAD	Mean absolute difference
MIP	Maximum intensity projection
O1	Observer 1
PA	Posteroanterior
PRF	Pulse repetition frequency
ROI	Region of interest
SAP	Superior articular process
SD	Standard deviation
SP	Spinous process

SPA	Spinous process angle
SPAA	Spinous process angle at apex
TP	Transverse process
US	Ultrasound
VAR	Vertebral axial rotation

Chapter 1.
Introduction

Adolescent idiopathic scoliosis (AIS) is a three-dimensional (3D) deformity of the spine occurring in adolescence with no known causes (Deacon *et al.*, 1984). The most commonly used imaging method for AIS is taking radiographs for assessing and monitoring the curvatures. However, this radiographic technique exposes patients to harmful ionizing radiation. There is a growing concern from families and health professionals that patients are exposed to too much ionizing radiation. Especially for the growing children who have AIS, the accumulated radiation from the follow-up clinics may increase the risk of breast cancer (Hoffman *et al.*, 1989; Doody *et al.*, 2000). Thus, alternative imaging methods without radiation are desirable to assess scoliosis. Ultrasound, which is a non-ionizing radiation and non-invasive method, has been applied to study scoliosis. More than two decades ago, ultrasound was attempted to image vertebrae and was able to identify the lamina and spinous process (Suzuki *et al.*, 1989). However, ultrasound has not been widely used because of challenges in data acquisition of the spine and the poor quality of ultrasound images.

Furthermore, the traditional radiograph only provides 2D images, which cannot reveal the true nature of scoliosis. The Cobb angle measured on the coronal view is the clinical standard to assess the severity of scoliosis. Vertebral axial rotation (VAR) is another important parameter to understand the spinal deformity. These two parameters play important roles on predicting curve progression, determining treatment options, and evaluating treatment outcome (Drerup, 1985; Kuklo *et al.*, 2005). Therefore, accurately measuring these two parameters from ultrasound images is important.

Although the technology and literature support ultrasound to image the human spine, there is a lack of clinical evidences that ultrasound can be used to measure and monitor the severity of scoliosis in children with AIS. Furthermore, the theory, the bony landmarks and the measurement method of using ultrasound to evaluate the spinal curvature have not been fully described.

1.1 Objectives of the thesis

The objectives of this PhD work are 1) to investigate and optimize the ultrasound scanning protocol for children with AIS, 2) to determine which anatomic features can be used to assess the severity of AIS, 3) to determine the reliability and validity of the coronal curvature and VAR measurements from ultrasound images, and 4) to develop a semi-automatic program calculating the coronal curvature and the VAR in AIS patients within 15 minutes during scoliosis clinic.

To address the research goals, four tasks have been performed 1) a feasibility study of using ultrasound to image vertebrae on spinal phantoms and healthy volunteers; 2) a study to determine the measurement methods for coronal curvature assessment using the recognized landmarks; 3) reliability studies of the ultrasound measurements on both the coronal curvature and VAR from the *in-vitro* and *in-vivo* images; 4) the development of a semi-automatic program to measure the coronal curvature and VAR of AIS, and the evaluation of the program performance on the *in-vitro* and *in-vivo* ultrasound data, the Cobb angle from scoliosis clinical record and experimental setups for VAR were used to validate the program .

1.2 Thesis organization

This thesis includes eight chapters. Chapter 1 gives a summary of the existing problems, introduces the objectives of this project and the thesis organization. Chapter 2 provides the background and the literature review on the terminologies of human spine anatomy, scoliosis and medical imaging methods for spinal deformity. According to the literature, ultrasound had the potential to assess scoliosis. Then the ultrasound equipment, the interface software applied in this study, and the transducer optimum element group for imaging is described.

In chapter 3, the *in-vitro* experiments are performed using a cadaver vertebra and a spinal phantom to determine which bony landmarks can be recognized from the ultrasound images. After that, a healthy volunteer is scanned by the ultrasound system to verify its feasibility to identify the landmarks on the *in-vivo* images. Then the bony landmarks are determined from the *in-vitro* and *in-vivo* results.

Chapter 4 presents a study using the recognized features described in Chapter 3 to assess the coronal curvature of AIS. In order to investigate which bony landmarks could provide reliable coronal curvature measurements, two methods, the center of pedicle (COP) and the apex of the spinous process angle (SPAA), are proposed and implemented on radiographs. The reliability of these two methods is investigated and the validity is determined by comparing with the Cobb angle measured from the corresponding radiographs. Then the bony landmarks providing better assessment of the coronal curvature are determined to be used on ultrasound images.

Chapter 5 describes a reliability study of the ultrasound measurements of coronal curvature using the center of lamina (COL) method. The intra- and inter-observer reliabilities of the COL method are investigated using both the *in-vitro* and *in-vivo* data. The measurements of the COL method on subjects' data are also compared with the Cobb angle from scoliosis clinical record to determine its validity.

Chapter 6 reports a reliability study of the vertebral axial rotation (VAR) measurements on the *in-vitro* and *in-vivo* data using the COL method on ultrasound transverse images. The VAR measurements using the COL method on ultrasound images are compared with the VAR measurements using the Stokes' method on radiographs to investigate if there is association between these two methods. The validity of the VAR measurements is investigated on the *in-vitro* images by comparing with the experiment setup.

Chapter 7 describes the details of the semi-automatic program that has been applied to measure the coronal curvature and VAR on both the *in-vitro* and *in-vivo* data. The purpose is to improve the image quality and reduce human errors while selecting the centers of laminae. The intra- and inter-observer reliabilities are studied on *in-vitro* and *in-vivo* data to investigate the performance of the program.

Chapter 8 provides a summary of this thesis and a future recommendation to improve this research.

Chapter 2.
Background

This chapter introduces the background on scoliosis and the ultrasound equipment with the interface software. The terminologies of the anatomical directional terms and body planes are first introduced, and then the structures of the vertebral column and individual vertebrae are presented. The definition, the classification, and the prevalence of scoliosis are described. Two important parameters, the Cobb angle and the vertebral axial rotation (VAR), for assessing the severity of the spinal deformity are introduced and their measurement methods and reliabilities are reported. The treatment options for scoliosis are also described. Different imaging methods including the X-ray based methods (radiograph, computed topography (CT), EOS system), magnetic resonance imaging (MRI), surface topography, and ultrasound are presented. At the end of the chapter, the ultrasound equipment used in this study including the main ultrasound unit, the transducer, the encoder, the interface software and the custom frame are described in details. The properties of the ultrasound equipment have been studied to determine the optimal setup for spine imaging.

2.1 Human anatomy

2.1.1 Terminology

The terminologies of the directions and the planes of the body used in this thesis are described in the following and illustrated in Figure 2.1.

The superior/cephalic indicates the upper part of a structure, while the inferior/caudal refers to the lower part of a structure. The anterior and posterior represent the front and the back of the body, respectively. The lateral means sideward which is away from the midline of the body.

The sagittal plane is a vertical plane that divides the body into right and left parts. The coronal plane (also known as the frontal plane) is a vertical plane, which divides the body into anterior and posterior parts. The transverse plane lies horizontally and divides the body into superior and inferior parts.

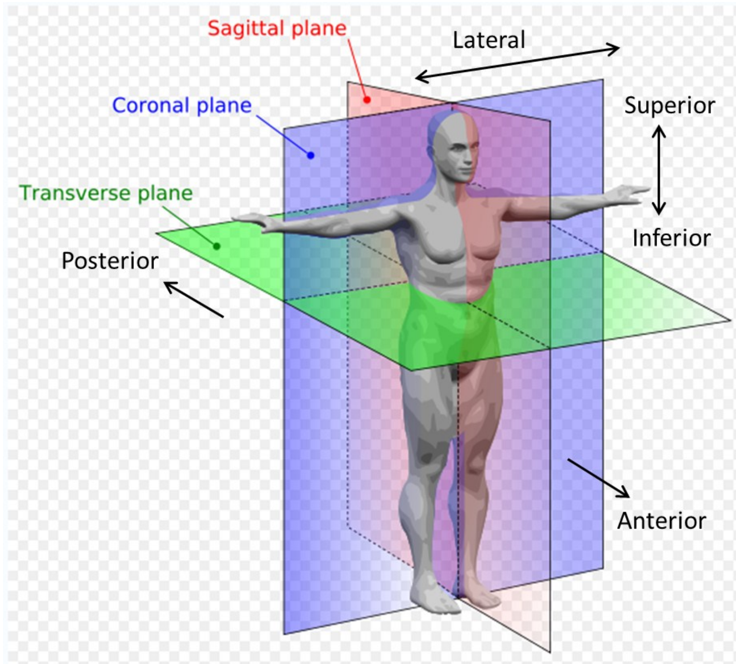
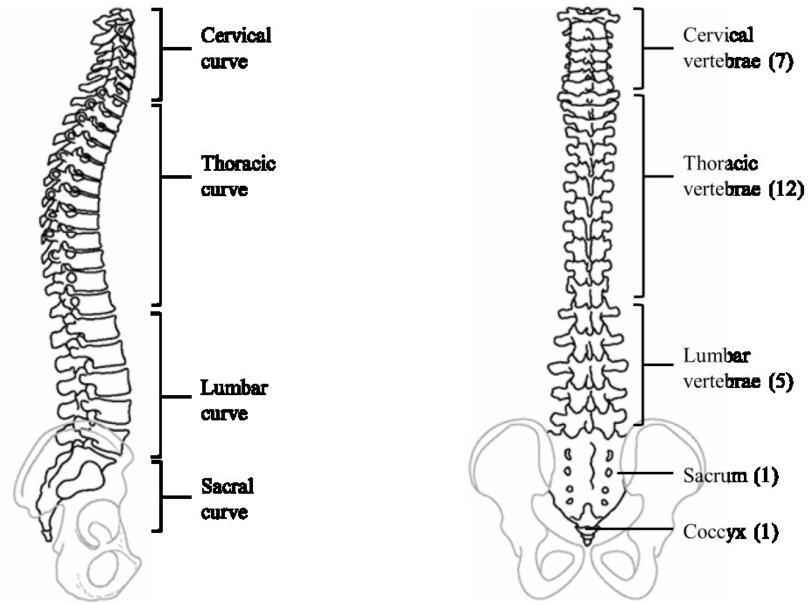


Figure 2.1 Anatomical planes and directions of axes.

2.1.2 Spine

The Spine (also known as the vertebral column) is a bony structure formed by vertebrae. The spine consists of 7 cervical vertebrae, 12 thoracic vertebrae, 5 lumbar vertebrae, and 9 fused vertebrae in the sacrum and the coccyx from the top to the bottom, as shown in Figure 2.2. There are 33 vertebrae in the human spine. The vertebrae are named with respect to their locations in the spine: C1-C7 in the cervical area around the neck, T1-T12 in the thoracic area linking with the ribs, L1-L5 in the lumbar area, S1–S5 in the sacrum area and the last 4 fused coccygeal vertebrae linking with the pelvis. Figure 2.2a and Figure 2.2b show the sagittal and coronal views of a normal spine, respectively.

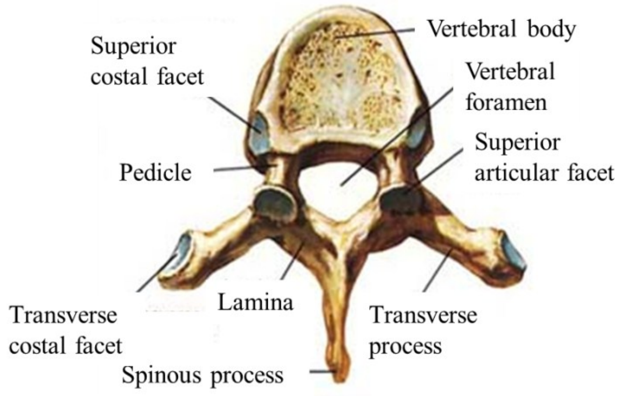


(a) (b)

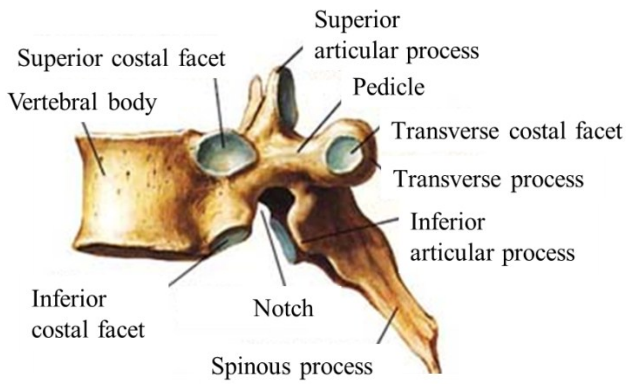
Figure 2.2 The vertebral column. (a) The sagittal view, and (b) the coronal view.

2.1.3 Vertebral structures

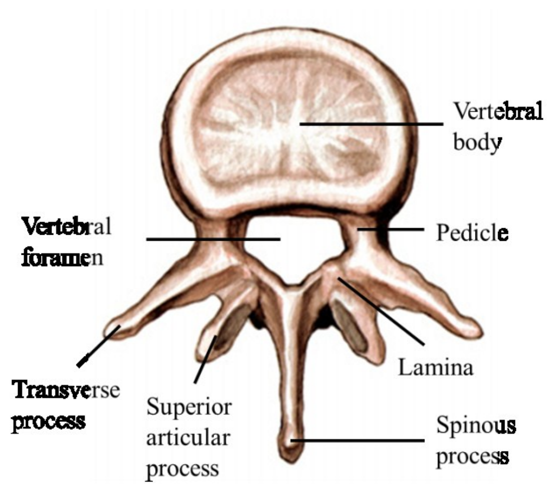
Each vertebra consists of a vertebral body, a vertebral arch, and processes. The vertebral body is the largest part of the vertebra with the shape of cylinder. The vertebral arch includes a pair of pedicles and a pair of laminae. It supports the processes such as spinous process, transverse processes, inferior articular processes, and superior articular processes. Although each vertebra has the similar structure components, the size and shape are different depending on its locations. The top and side views of the thoracic and lumbar vertebrae are shown in Figure 2.3. The dimensions and the shapes of the lumbar vertebrae are larger than and different from the thoracic vertebrae. Since thoracic vertebrae articulate with the ribs, they have the demifacets and the facets articulating with the tubercles of the ribs that the lumbar vertebrae do not have.


Vertebral body
Superior
costal facet
Vertebral
foramen
Superior
articular facet
Pedicle
Lamina
Transverse
costal facet
Transverse
process
Spinous process

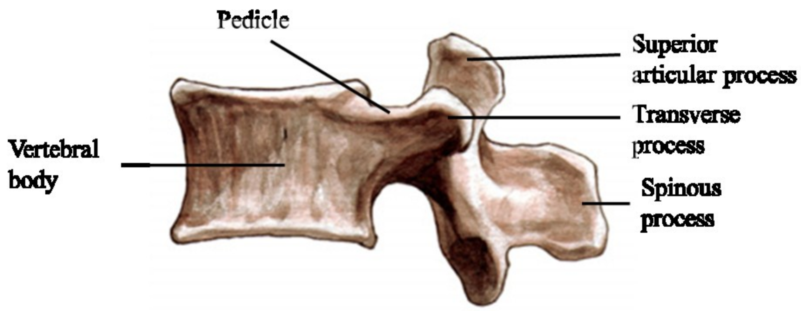
(a)


Superior
articular process
Superior costal facet
Pedicle
Vertebral body
Transverse costal facet
Transverse process
Inferior
articular process
Inferior
costal facet
Notch
Spinous process

(b)


Vertebral
body
Vertebral
foramen
Pedicle
Lamina
Transverse
process
Superior
articular
process
Spinous
process

(c)



(d)

Figure 2.3 Vertebral structures. (a) and (b) display the top and side views of a thoracic vertebra, and (c) and (d) show the top and side views of a lumbar vertebra, respectively.

2.2 Scoliosis

2.2.1 Definition

Scoliosis is a complex three dimensional deformity of the spine in the coronal, transverse, and sagittal planes, which associated with axial rotation (Deacon *et al.*, 1984). The lateral curvature affects the rib cage and presents as deformities of the trunk as seen in Figure 2.4. These deformities of the trunk may include asymmetrical elevations of the shoulders and/or hips, prominence of the scapula, waist asymmetries, and trunk rotation. Therefore, annual scoliosis screening is recommended for children aged 10-14 years by Scoliosis Research Society. The forward bend test requires a child to bend forward at the waist with the knees straight and palms together. The examiner looks for any unevenness from the behind, front and side of the trunk. Also, the examiner may measure the angle of trunk rotation with a scoliometer (Morrissy and Weinstein, 2006).

According to etiology, scoliosis can be classified to congenital scoliosis, idiopathic scoliosis (IS), scoliosis in generalized diseases and syndromes, traumatic scoliosis, and degenerative scoliosis (Van Goethem *et al.*, 2007). Approximately 80% of all scoliosis cases are IS for which there is no known cause. IS are usually divided into four classifications based on age of diagnosis: infantile (0-4 years), juvenile (>4 and ≤10), adolescent (>10 and ≤17) and adult (>17). Scoliosis can also be classified by the location of its apical vertebra, known as the apex, which is the most laterally deviated vertebra of the curve: cervical (apex between C2 and C6), cervicothoracic (apex between C7 and T1), thoracic (apex between T2 and T11), thoracolumbar (apex between T12 and L1), lumbar (apex between L2 and L4), or lumbosacral (apex at L5 or below).

The prevalence of idiopathic scoliosis in childhood and adolescence is ranging from 0.5% to 3% (Morrissy et al., 2006). The ratio of girls to boys is equal among patients with mild curves, but girls predominate as the curve magnitude increases, with the ratio reaching 8:1 among those requiring treatment (Bunnell, 1988).

(a) (b)

Figure 2.4 a) External asymmetries of a scoliotic patient, and b) internal alignment of the corresponding patient. (Morrissy *et al.*, 2005)

2.2.2 Cobb angle

The apex of a curve is the most laterally deviated disk or vertebra of the curve shown in Figure 2.5. The end vertebrae of a curve define the proximal and distal extent of a curve and are determined by locating the most tilted vertebrae from the horizontal.

The Cobb method (Cobb, 1948) is the clinical standard used in clinic to measure the coronal curvature from the standing radiograph. To measure the Cobb angle, one should outline the superior end plate of the top end vertebra and the inferior end plate of the bottom end vertebra, construct a perpendicular to each of these lines, and then measure the angle formed by these two lines as shown in Figure 2.5. According to the Scoliosis Research Society, scoliosis is defined with a Cobb angle greater than 10°. Besides assessing the magnitude of the curvature, the Cobb angle is also used to monitor the progression of the curve, decide the treatment options, and evaluate the treatment outcomes (Helenius *et al.*, 2003; Katz *et al.*, 1997; Tan *et al.*, 2009).

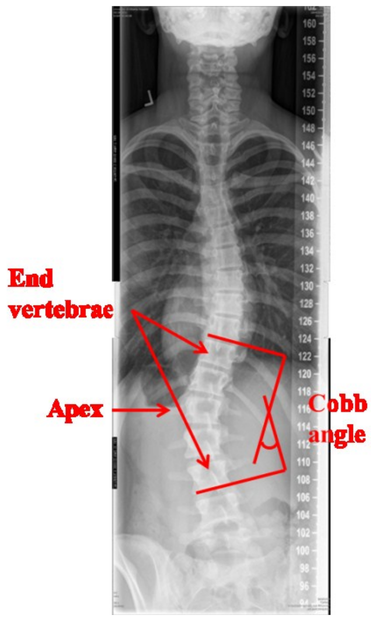


Figure 2.5 Posteroanterior radiograph showing the vertebrae to define the curvature.

The intra- and inter-observer measurement errors of the Cobb angle were reported to be in the range of 3°-8° (Carman *et al.*, 1992; Morrissy *et al.*, 1990; Shea *et al.*, 1998). Standing position and postural changes can have some effects on the Cobb angle, but most of the errors are still from human measurements. Studies comparing measurements from digital and traditional radiographs indicated that using digital radiographs had the potential to reduce errors by avoiding using different drawing tools (Mok *et al.*, 2008; Shea *et al.*, 1998; Srinivasalu *et al.*, 2008). Accompanied with the popularity of the digital radiographs, some computer-aided methods were developed to reduce human measurement error by minimizing the subjective factors such as selection of end vertebra and experience of the observer (Chockalingam *et al.*, 2002; Stokes and Aronsson, 2002; Zhang *et al.*, 2009). The intra- and inter-observer errors were reduced to 3° when using the computer-aided method to perform the Cobb angle measurements on 84 radiographs by three observers (Zhang *et al.*, 2010).

2.2.3 Vertebral axial rotation

Although the Cobb angle is the clinical standard to measure the severity of scoliosis, this measurement is limited to two-dimension (2D) which cannot reveal the full extent of the 3D spinal deformity. Vertebral axial rotation (VAR) is another important parameter that can be used to assess the severity of scoliosis, to predict the risk of curve progression, and evaluate the treatment outcome (Drerup, 1985; Kuklo *et al.*, 2005). Therefore, VAR becomes increasingly prominent in the study of scoliosis. From the standing PA radiograph, vertebrae having axial rotation show a pair of asymmetric pedicles with respect to the vertebral body (Figure 2.6).

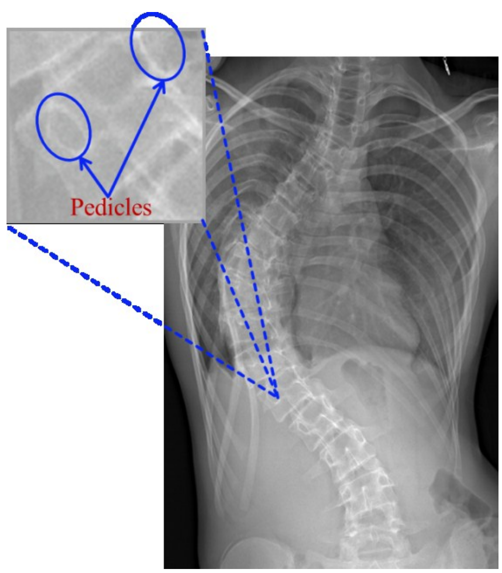


Figure 2.6 The vertebra with vertebral axial rotation.

Some methods have been developed to measure VAR on radiographic images using projected landmarks such as the spinous process and pedicles (Cobb, 1948; Nash and Moe, 1969; Perdriolle and Vidal, 1985; Stokes *et al.*, 1986). The Cobb method has simple procedure using SP to grade the rotation, but there is no means to quantify rotation from this gradation scheme; the method by Nash and Moe can quantify the rotation, but the angle may be over-estimated; the Perdriolle method is affordable and accurate, however, the accuracy is reduced when measuring large degrees (over 30°) of rotation; the Stokes method has similar results as the stereoradiograph measurements, while introducing random error when marking the vertebral edges (Lam *et al.*, 2008). However, the measurements on the radiographs are not directly performed on the transverse plane and the spinal deformity especially the axial rotation could change the size and shape of the projected landmarks, which limits the accuracy of the 2D measurements in the coronal plane.

Other 3D methods were proposed to measure the VAR on the transverse images such as computed topography (CT) (Aaro and Dahlborn, 1981; Ho *et al.*, 1993) and magnetic resonance imaging (MRI) (Birchall *et al.*, 1997). The CT-based measurement methods, which provide more details on the vertebra structures, are able to display the transverse plane images and more precise to assess the axial rotation. However, CT requires more processing time, is relatively more expensive, and exposes patients to more radiation than the standard radiograph. Not many clinics use the CT method to diagnose scoliosis regularly. MRI can also be used to measure the VAR on the transverse images and does not introduce ionizing radiations. However, MRI is also very costly and time consuming; it is not common to use MRI to evaluate scoliosis either. Also, the supine position applied by the CT and MRI methods significantly affects the lateral curvature and VAR measurements (Yazici *et al.*, 2001).

The quantitative evaluation of VAR measurements were reviewed and discussed on both the radiograph and CT methods (Vrtovec *et al.*, 2009). Among these methods, the Ho method measured from the CT images was the most reliable method with small standard deviation for both the intra- and inter-observer comparisons (<2°) (Ho *et al.*, 1992). However, the CT method is not common for assessing scoliosis. To improve the performance of the radiograph methods, computer-assisted methods were proposed to avoid the manual operation and exclude the impact of the experience from observers (Chi *et al.*, 2006; Zhang *et al.*, 2010). Up to now, there is no routine method available for the visualization and quantitative evaluation of the VAR.

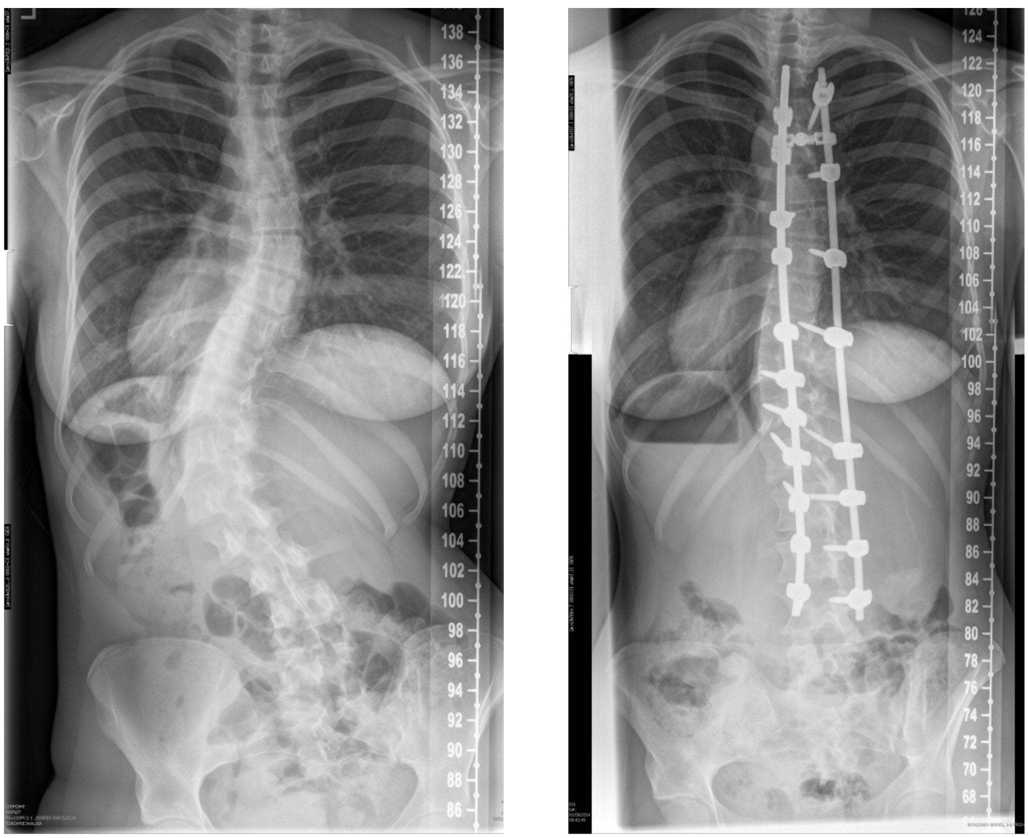
2.2.4 Treatment options

Treatment for scoliosis is based on consideration of the patient's physiologic maturity, curve severity and location, surface topography, and the projected potential for progression (Morrissy and Weinstein, 2006). In general, patients are treated differently according to the Cobb angle.

Most AIS patients, who have curves among 10°-25° with no significant progression, are monitored every 4 to 12 months. The orthotic (brace) treatment is considered for growing children and adolescents with a curve between 25° and 45° (Weinstein *et al.*, 2008). Treatment with bracing is to prevent curve progression during the high-risk period of the adolescent growth phase by applying the external forces. Figure 2.7 shows a patient wearing a custom brace. Studies indicated that bracing significantly decreased the progression of the high-risk curves to the threshold for surgery in AIS patients (success rate 72% and 75%), and brace wear for long hours (>12.9h) per day was associated with high success rate (90% to 93%) in avoiding the need for surgery (Weinstein *et al.*, 2013).

Figure 2.7 A patient with a custom brace.

Surgical correction is reserved for curves more than 45° in the mature patients by implanting instrumentation, such as screws, rods, and connectors, along the treated vertebrae for internal support and fusion. The purpose is to halt the progression of the curve and reduce the spinal deformity. Surgery significantly corrects the spinal deformity; however, complications may occur such as hook dislodgment, infections, and neurologic deficits (Helenius *et al.*, 2003). The radiographs of pre-operation and at two months after post-operation of an AIS patient are shown in Figure 2.8. The Cobb angles were reduced from 42°/65° to 19°/19°.



(a) (b)

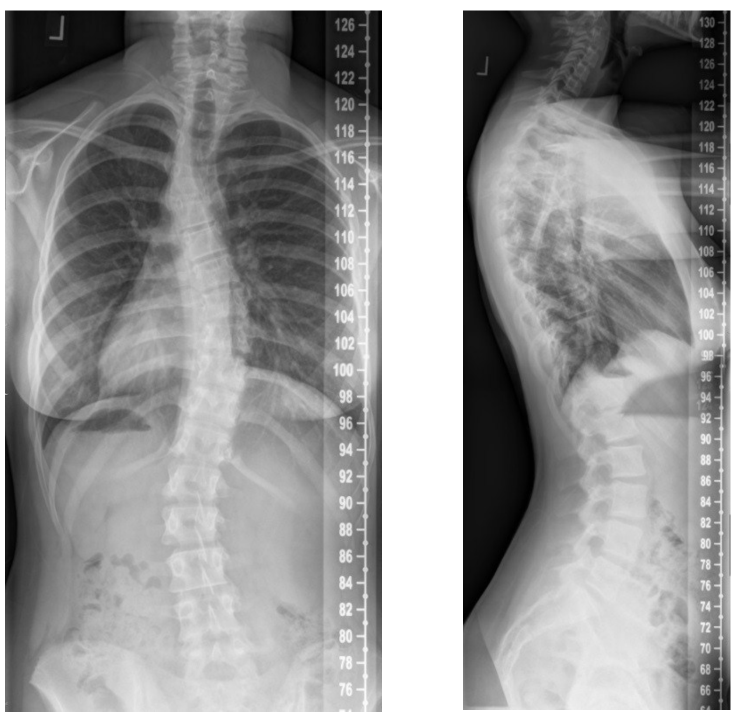
Figure 2.8 Radiographs of (a) pre-operation and (b) post-operation at two months of an AIS patient.

2.3 Imaging methods

2.3.1 X-ray based techniques

Radiography is the use of X-rays to view the objects for diagnosis. When a primary beam passes through the body, some of the radiation is absorbed known as attenuation. X-rays passing through the object are captured to represent the structures. High density parts have a higher rate of attenuation than that of the low density parts, so bones will absorb more x-rays than soft tissues. Thus, a high energy photon source is required for imaging the bone structures.

Generally, posteroanterior (PA) radiography is used to obtain images of spine for scoliosis in an upright standing posture. The Cobb angle can be measured on the PA radiograph (Figure 2.9a). Also, sagittal (side view) radiographs are usually taken to assess thoracic kyphosis and lumbar lordosis (Figure 2.9b).



(a) (b)

Figure 2.9 Radiographs of the scoliosis patient: (a) coronal view, and (b) sagittal view.

Computed tomography (CT) is a medical imaging method employing tomography principles. 3D image can be reconstructed from a series of CT images. CT scans can provide detailed morphological structures of the spine (Heo *et al.*, 2010), and is considered as the standard in assessing the 3D reconstruction of other methods such as the biplanar radiographs with EOS (Glaser et al., 2012). CT techniques have been applied to assist the spinal surgery. Some studies used CT to evaluate the size of pedicles for pedicle screw selection and trajectory (Gstoettner *et al.*, 2011; Kuraishi *et al.*, 2013). The pre- and postoperative parameters were measured from CT images to evaluate the corrections of scoliosis (Asghar *et al.*, 2009; Hong *et al.*, 2011). However, routine CT scans are not used in patients who have AIS because they expose patients to more harmful ionizing radiations (Brenner and Hall, 2007). Also, the supine position alters the magnitude of curves both in the coronal plane and transverse plane (Torell *et al.*, 1985; Yazici *et al.*, 2001).

The EOS 2D/3D radiography system (Biospace Med, Paris, France) is a biplane imaging device using slot-scanning technology (Figure 2.10). It includes two orthogonal x-ray sources, and is able to capture the coronal and sagittal images simultaneously (Deschenes *et al.*, 2010). This system reduces radiation dose 6 to 9 times compared to the computed radiography (CR) system, and provides images with comparable or better quality than CR (Deschenes *et al.*, 2010; McKenna *et al.*, 2012; Samei *et al.*, 2005). With the station sterEOS, the system can reconstruct 3D models of the spine based on statistical modeling and bone shape recognition from the coronal and lateral images. After a few descriptor parameters, such as the width and depth of the T1 and L5 endplate, and the length of the spinal curve, were digitized on both the coronal and lateral images, the vertebrae models were generated based on longitudinal and transversal inferences (Humbert *et al.*, 2009). The scoliotic parameters such as the Cobb angle, vertebral axial rotation, and sagittal balance parameters can be measured on the reconstructed models (Ilharreborde *et al.*, 2011; Illes *et al.*, 2011).

Figure 2.10 EOS slot-scanning radiograph system. (Deschenes *et al.*, 2010)

Plain radiography exposes patients to harmful ionizing radiation and CT scanners emit more radiation than conventional radiography (Brenner and Hall, 2007). Young scoliotic patients require years of follow-up to monitor the progression of spinal curvature with the possibility of accumulating higher radiation dosages due to increasing number of radiographic exposures. Some cohort studies involving female patients showed that frequent exposure to low dose radiation during childhood and adolescent may increase breast cancer risk (Hoffman *et al.*, 1989; Doody *et al.*, 2000). Although the EOS 2D/3D system can reduce the radiation dose, the patient health benefits by the reduction of radiation need more evidence (McKenna *et al.*, 2012).

2.3.2 Magnetic resonance imaging

Magnetic resonance imaging (MRI) makes use of the property of nuclear magnetic resonance to image nuclei of atoms inside the body. It is an ionizing-radiation free technique. MRI can visualize detailed internal structures, especially tissues with much hydrogen and little density contrast, such as brain, muscle, and most tumors. Thus, MRI screen is usually applied for scoliosis related to the spinal abnormality and tumors (Morrissy and Weinstein, 2006; Thomsen and Abel, 2006).

MRI is able to generate images of the spine in the coronal, sagittal, and transverse planes, and scoliotic parameters can be measured from the different planes (Birchall *et al.*, 2005; Schmitz *et al.*, 2001). Birchall *et al.* (2005) used 3D MRI method to measure mechanical torsion of the vertebral bodies and discs of AIS. However, MRI was not recommended as the routine method for the AIS patients since a low intraspinal anomalies rate was detected (<3%) (Do *et al.*, 2001; Nakahara *et al.*, 2011). Besides, the Cobb angle measurement from the traditional MRI, which scans a body on a supine position, underestimates the curvature (Torell *et al.*, 1985). Also, MRI is very costly and time consuming; it is not common to use MRI to diagnose or monitor AIS.

2.3.3 Surface topography

Back only surface topography is a three-dimensional measurement of back's surface using scanned light, photographic techniques or sensors (Oxborrow, 2000).

Several systems based on surface topography were suggested such as optical based Moire fringe topography, the integrated shape imaging system, ISIS (Oxford Metrics, Ltd., Oxford, UK), and the Quantec Spinal Imaging System, QSIS (Quantec Inc., Lancashire, England) (Daruwalla and Balasubramaniam, 1985; Goldberg *et al.*, 2001; Turner-Smith *et al.*, 1988; Weisz *et al.*, 1988). The Moire fringe topography used the fringe contour lines pattern on the back of the patient to interpret the internal spinal deformity (Daruwalla JS, Balasubramaniam P, 1985). The ISIS system was designed to quantify the distortion of horizontal scanned line over a subject (Turner-Smith AR, 1988). The QSIS was a raster stereography to provide 3D information (Quantec angle) of a trunk (Curran, P and Groves D, 1990).

For the past decade, researchers have focused on the full torso surface topography. A patented INSPECK system (Song *et al.*, 2002) and a commercial non-contact 3D digitizer (Vivid 910, Konica Minolta, Japan) have been used to capture the full torso and have tried to correlate the surface features with internal alignment. However, the correlation of the surface features and the Cobb angle was not very strong, (Parent *et al.*, 2010; Pazos *et al.*, 2005).

Another method, which used the sensor of the Ortelius 800^{Tm} system (Orthoscan Technologies Inc., MA USA) to touch spinous processes of scoliotic patients, was able to reconstruct the spinal alignment and to calculate the vertebral locations and the curvature angles. A study involving 52 AIS patients showed good agreement between the curvature angles from the Ortelius and the Cobb angle measurements (Parisini *et al.*, 2006). However, another study did not support the previous findings (Knott *et al.*, 2006). The controversial results caused the technology to not be commonly used.

2.3.4 Ultrasound

Ultrasound is a high frequency sound wave, with the frequency above 20 kHz. The frequencies for medical uses of ultrasound are usually in the range of 2 MHz to 10 MHz (Bushberg *et al.*, 2002). Ultrasonography (diagnostic sonography) is an ultrasound-based diagnostic imaging technique and has been used in clinics for many years. For imaging, pulse-echo techniques are used to provide an acoustic map of tissues. The transducer transmits an ultrasonic pulse into the tissues under examination and receives reflection signals which occurred at the interface between different tissues. The echo data is then displayed for diagnostic purposes. Ultrasound helps visualize soft tissues like muscles, vessels and internal organs for possible pathology or lesions.

Using ultrasound to image bone is not common because ultrasound signals are reflected and attenuated when ultrasound travels through soft tissue. The tissue-bone interface is a strong reflector and reflects most of the US signals, which makes vertebral imaging possible. Using a medical ultrasound scanner (Shimazu SDL-300) with a 5.0 MHz transducer, researchers from Kyoto University (Suzuki *et al.*, 1989) successfully identified the spinous processes and laminae from 47 scoliotic patients in a prone posture. The VAR was measured and a linear correlation was found with the Cobb angle in untreated patients.

Another study showed the accuracy of using ultrasound to locate lumbar intervertebral level (Furness *et al.*, 2002). One group developed their own ultrasound system for the assessment of AIS (Cheung *et al.*, 2010). The spinous processes and transverse processes were manually marked from the stack of 2D ultrasound images. Then a 3D image was formed from these markers, and the transverse processes were exploited to measure the curvature. However, the experiments were only performed on Sawbones spinal column phantoms; no *in-vivo* data was included in the study yet. Also, a considerable time was needed to process the data, which was a critical issue for clinical application. Thus, further validation study is needed before applying their method to clinic. Another group applied a freehand 3D ultrasound system to image spine (Purnama *et al.*, 2009; 2010). The data was acquired from a volunteer without scoliotic curvature in a prone posture. Six scans with overlap areas were performed. A volume reconstruction was then done including four stages: bin-filling, hole-filling, volume segment alignment and volume segment compounding. This study showed the feasibility of using ultrasound to image spine. However, only one volunteer who had no scoliosis was recruited and the volume reconstruction procedure took a day for the processing, concluding that further development is required. Recently, ultrasound was applied to improve the orthotic (brace) treatment by tracing spinous process angle (SPA) during the brace fitting procedure (Li *et al.*, 2010; 2012). The SPA was the summation of the angle formed by two lines joining three neighbor spinous processes in the curve. First, the pre-brace SPA was measured without putting on the brace. Then the different in-brace SPAs were calculated when the major pressure pad inside the brace was adjusted to different positions. The optimal location of the pressure pad was determined at the spot where the largest curvature correction was achieved according to the differences between the pre-brace SPA and the in-brace SPA measurements. As a result, 62% of patients in this study benefited from the ultrasound assistance.

All the above ultrasound studies indicate that it is feasible to use ultrasound to image the spine and US should be able to measure coronal curvature and VAR. However, the reliability and validity of the ultrasound measurements have not been reported. This question becomes part of the focus of this research.

2.4 Ultrasound system

2.4.1 Equipment

In this study, an Olympus TomoScan Focus LT™ Phased Array Ultrasound system (Olympus NDT Inc., Canada) was used as shown in Figure 2.11. This non-medical ultrasound system was chosen because it provides data portability and flexibility in data acquisition. Data portability is important as further post-acquisition image processing can be applied if needed. Beam forming and transmit/receive focusing for different depth zone can be generated by applying time delays among transducer elements. The scanning pattern can be linear or sectorial (or angular). The ultrasound system is connected to a computer via Ethernet port for fast data transfer. A computer installed with the Tomoview™ software (Version 2.9 R12) was used to control the data acquisition process and to modify the parameters of the ultrasound beam such as scanning mode, beam angle, focal position, and active aperture. The acquired data could be exported for further post-acquisition analysis. Real-time data compression and signal averaging were also available. For a 3D US data set, different views could be generated by Tomoview software. Coronal, transverse and sagittal views could be displayed either in a single slice or projected image by multiple slices. The position of the single slice or the range for the projected image can be determined by the reference cursors on other views (Figure 2.12). For example, the depth range for the projected coronal view shown in Figure 2.12 was determined by the reference cursors from the sagittal and transverse views.

Figure 2.11 The TomoScan Focus LT^{TM} phased array ultrasound system.

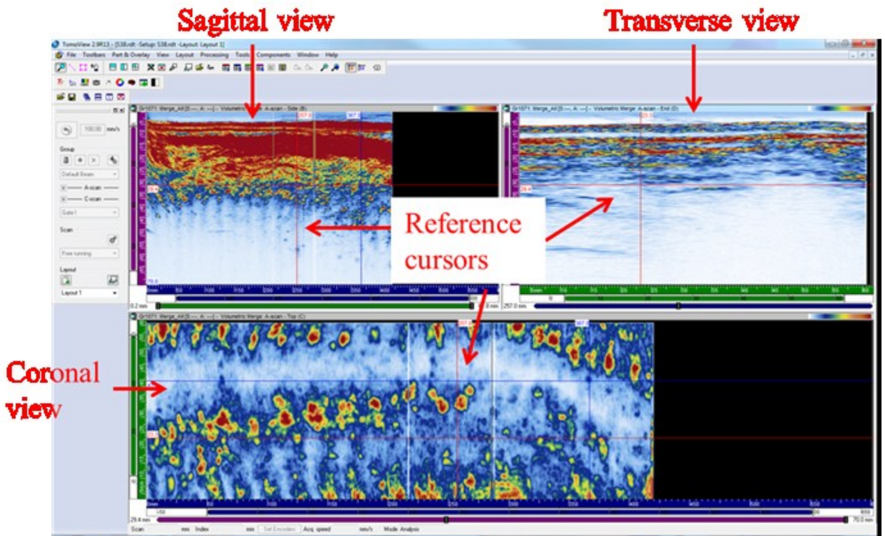


Figure 2.12 The Tomoview^{TM} software analysis interface.

The ultrasound transducer is a 5.0-MHz 64-element array probe (5L64-I1) shown in Figure 2.13. The probe has an active area of 38.4 mm (length) by 10 mm (elevation) with a pitch (p) of 0.6 mm. Different number of elements can be grouped together to form beam at different angle or depth by applying time delays (Figure 2.14a). For transmit focusing, the outer elements are fired before the inner elements to achieve focus at a certain depth. Increasing time delays between adjacent elements decreases focal depth while reducing the delays causes distal depth focusing. For receive focusing, the echoes coming back from a certain depth are stored, delayed, and then summed to produce an ultrasound signal. Generally, the echoes received by the outer elements travel longer distances than the echoes by the inner elements. In order to align the echoes for summing, the echoes received at the center element of the group are delayed the most while those received at the edge of the groups are delayed the least. The amount of delay is such that it ensures the phases of the echoes are aligned to be summed and form an A-line. The ultrasound beam converges at the focal depth where the beam width is the minimum, providing the best lateral resolution and beyond which the beam starts to diverge. Lateral resolution is depth dependent. The TomoScan™ system has the dynamic depth focusing (DDF) capability. DDF is a beam forming technique to extend the focal zone over a depth range by adjusting the time delays to focus on different depths (Figure 2.14b).

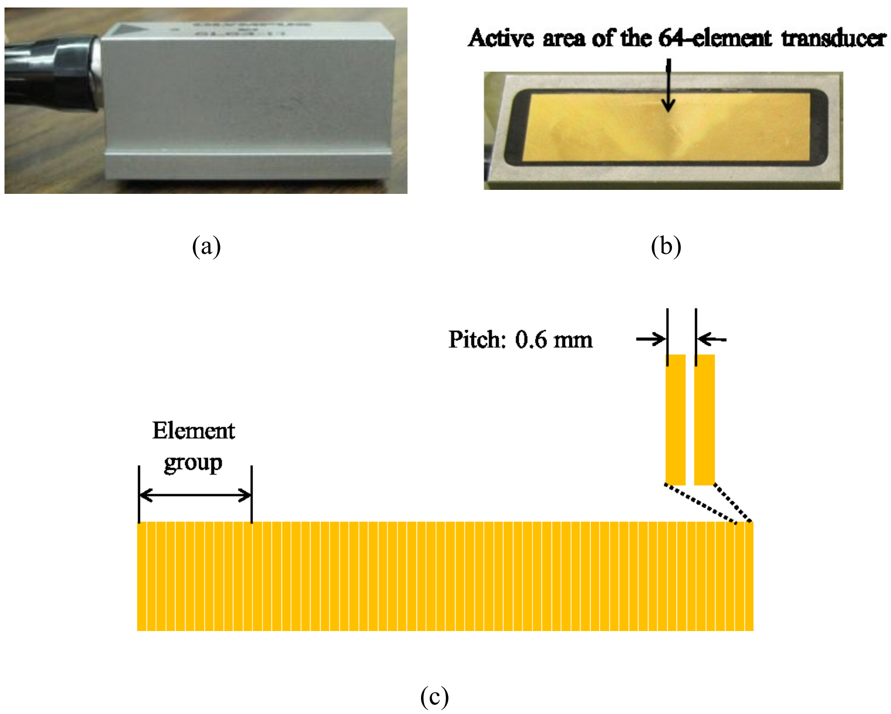
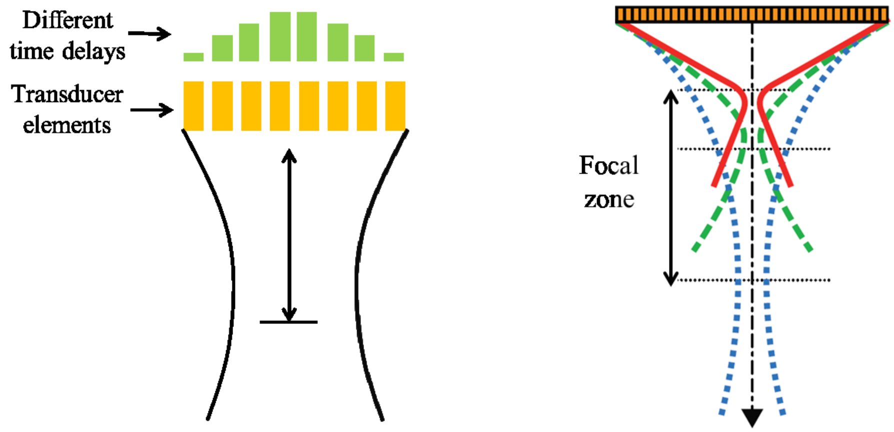


Figure 2.13 Transducer. (a) Side view of the transducer, (b) bottom view of the transducer, and (c) elements in the transducer.



(a) (b)

Figure 2.14 (a) Time delays for beam transmitting, and (b) dynamic depth focusing (DDF).

The encoder (shown in Figure 2.15) works in synchronization with the transducer movement during data acquisition. The position information of the transducer can be recorded by the system.

Figure 2.15 The mini-wheel encoder.

2.4.2 Frame

To acquire the ultrasound images of the whole spine, a custom frame was designed by a mechanical designer to assist the scanning procedure. Several requirements were used for the frame design: 1) vertical scanning procedure could be performed from the top to the bottom along the subject's back; 2) the transducer and encoder part attached on the frame should have the freedom to come in and out since there are spinal curves from the sagittal view; 3) the beginning of the scan position can be adjusted.

According to these requirements, the custom frame shown in Figure 2.16 was built. A XYZ-coordinate system was defined. XY-plane, YZ-plane, and XZ-plane indicate the coronal, transverse, and sagittal views, respectively. X-axis indicated the scan direction along which the transducer moved with the vertical bar (B). The horizontal arm (D) was able to come in and out along the Z-axis with the holder part attached to keep the transducer and encoder moving together. The vertical bar (B) could move to the left or right in the Y direction. The position of the vertical bar can be quantitatively adjusted by the scale engraved on the top and bottom crossbars (A, E). The metal bar (C) was applied for the posture of the subjects.

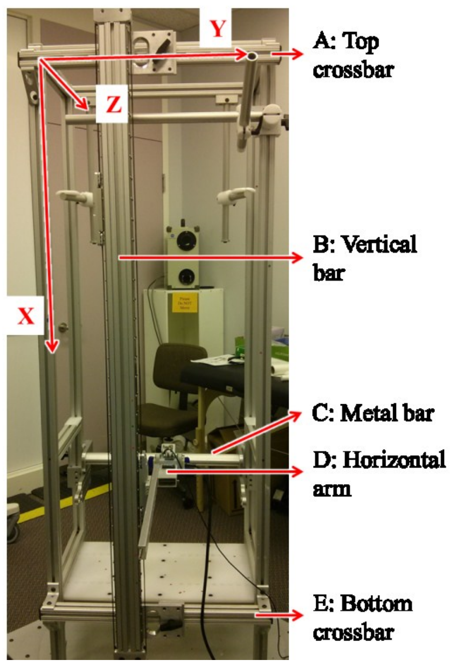


Figure 2.16 Custom frame.

2.4.3 Determining the number of elements for depth focusing

To better understand the performance and limitation of the ultrasound equipment, experiments were performed using a tissue-mimicking phantom (Model 549, ATS Labs Inc., Bridgeport, CT, USA) (Figure 2.17). According to information provided by the manufacturer, the material has a sound velocity of 1450 m/s at 0.5 dB/cm/MHz measured at room temperature. The phantom was designed to evaluate the axial/lateral resolution (the ability to distinguish two objectives along direction of US beam/perpendicular to the US beam), vertical/horizontal measurement calibration, dead zone (the distance from the front face of the transducer to the first identifiable echo), and focal zone (the region has the best lateral resolution).

To determine the optimum configuration for imaging application, tests were performed with different number of elements in one group from eight to forty. The transducer was put on the scan surface with a 200 N standard load on the top of the transducer to maintain a constant pressure. Water was laid on the scan surface to avoid air (Figure 2.17).

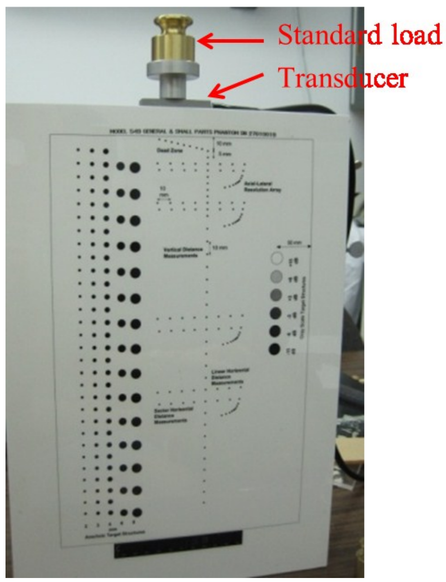


Figure 2.17 The tissue mimicking phantom Model 549.

The images with the different elements number on the same area are shown in Figure 2.18. According to the images, a group of twelve elements provided good lateral resolution and clear identification of the objectives in the image. Also, more elements in a group would decrease the imaging length, which changes from 34.2 mm to 15 mm with the elements number from eight to forty. A group of twelve elements was selected to be the optimum one with the imaging length of 31.2 mm.

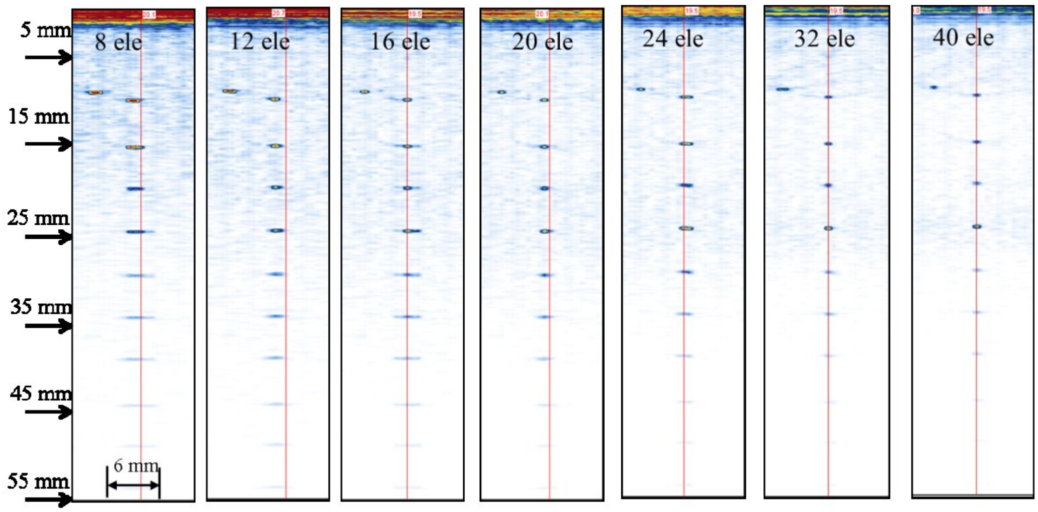


Figure 2.18 Images of groups with elements from 8 to 40 on the same area.

The maximum focal depth D_N for the 12-element group is 43.7 mm according to the following equation (Bushberg *et al.*, 2002):

$$D_N = \frac{d^2 f}{4c} \tag{2-1}$$

where the active group aperture d is 7.2 mm (12×p), the frequency f is 5 MHz, and the speed of sound c is 1480 m/s for water (Zhang *et al.*, 2011). Based on the resolution phantom study, the lateral resolution remains constant at 1.2 mm for depth between 5 mm and 43.7 mm using a group of 12-element with DDF.

2.5 Summary

In order to reduce the ionizing radiation by taking radiographs in children with AIS, non-ionizing radiation imaging methods were discussed. Among these, ultrasound has the potential to assess scoliosis. A non-medical ultrasound system was described for ultrasound data acquisition and the optimum element group was determined. Then the ultrasound system was applied for the studies in this project.

In order to achieve the objectives, several studies were performed and the flow chart is shown in Figure 2.19. A feasibility study was first performed to investigate the bony landmarks that could be recognized from ultrasound images. Then a comparison study was conducted on radiographs to determine which landmarks from the feasibility study could provide reliable measurements of the coronal curvature. After these two studies, the bony landmarks to be used on US measurements were determined. The reliability and validity of the coronal curvature and vertebral axial rotation measurements were investigated to find out if the proposed method was a promising tool. Then a semi-automatic program was developed based on the proposed measurement method to improve the image quality and reduce human measurements errors.

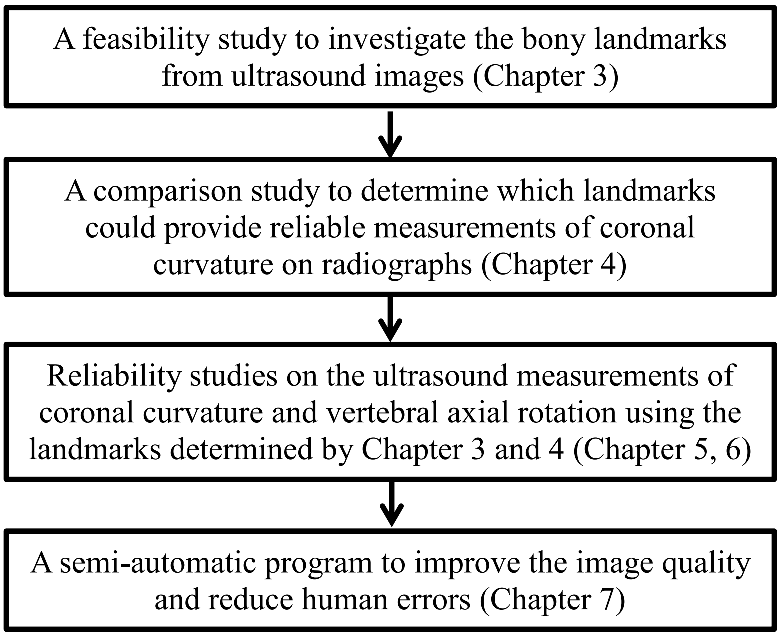


Figure 2.19 The flow chart of the studies.

Chapter 3.

Feasibility study on vertebrae phantoms and volunteer*

From the literature, it was possible to image human vertebrae using ultrasound (US) (Suzuki *et al.*, 1989; Li *et al.*, 2012; Purnama *et al.*, 2010). However, there is a lack of description about the ultrasound reflectors or markers that can be used to assess the spinal deformity. In this chapter, both *in-vitro* and *in-vivo* experiments using a non-medical US scanner were performed to obtain US images of the spinal vertebrae. In the *in-vitro* experiments, a cadaveric thoracic vertebra phantom and a Sawbones spinal column were scanned to investigate which anatomical markers could be reliably identified. Then an *in-vivo* study was performed on a healthy volunteer to investigate the feasibility of using US to image the human spine and study the effect of soft tissues.

3.1 *In-vitro* experiments

3.1.1 Vertebrae phantoms

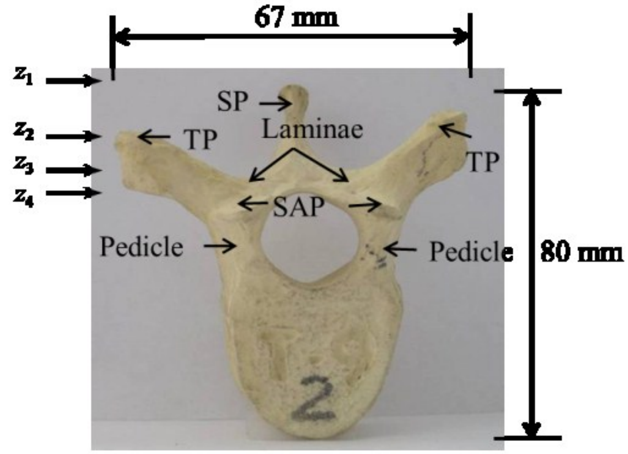
A cadaveric thoracic vertebra (T9) and a Sawbones spinal column phantom were used in this study.

Figure 3.1 shows the front, top, and side views of the cadaveric thoracic vertebra. With soft tissues removed, the specimen, which had the same physical properties as the human spine, was cleaned, dried, and treated for preservation

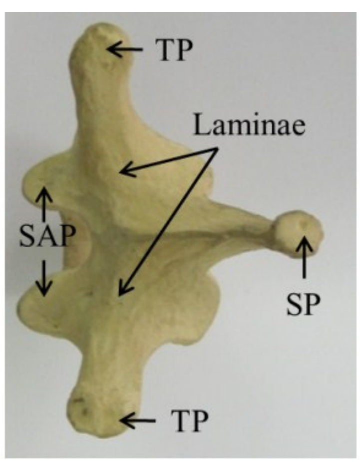
---

* Part of the materials in this chapter has been published in:1) Chen W, Lou EHM, Le LH (2011) Using ultrasound imaging to identify landmarks in vertebra models to assess spinal deformity. Conf Proc IEEE Eng Med Biol Soc:8495-8498. 2) Chen W, Le L, Lou E (2012) Ultrasound imaging of spinal vertebrae to study scoliosis. Open J Acoustics 2:95-103.

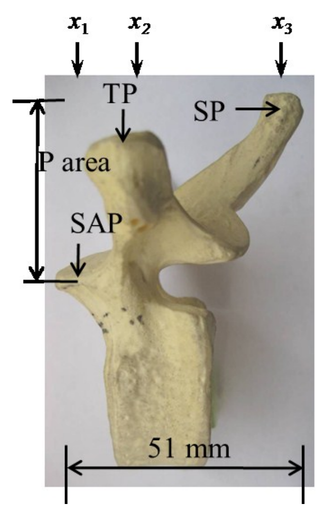
and frequent handling. The major posterior arch structures of the vertebra, which include the spinous process (SP), transverse process (TP), lamina, and superior articular process (SAP), are identifiable from the images (Figure 3.1). The maximum length, height, and thickness of the vertebra were 67 mm, 80 mm, and 51 mm respectively.



(a)



(b)



(c)

Figure 3.1 The vertebra phantom: (a) front view, (b) top view, and (c) side view. P area indicates the posterior area of the vertebra. The x-arrows (x_1 - x_3) show the positions of the displayed images presented later in Figure 3.7, and the z-arrows (z_1 - z_4) show the positions of the displayed images presented later in Figure 3.9.

The Sawbones spinal column phantom including vertebrae from T2 to T12 is fixed on a wooden platform (Figure 3.2). The maximum length, height, and thickness of the spinal column were approximately 65 mm, 240 mm, and 60 mm respectively. A 9-mm flexible metal wire threaded through the center of the vertebrae, which allowed free bending of the phantom and also was able to keep the shape of the phantom during movement. The TP, SP, and lamina can be recognized from the phantom.

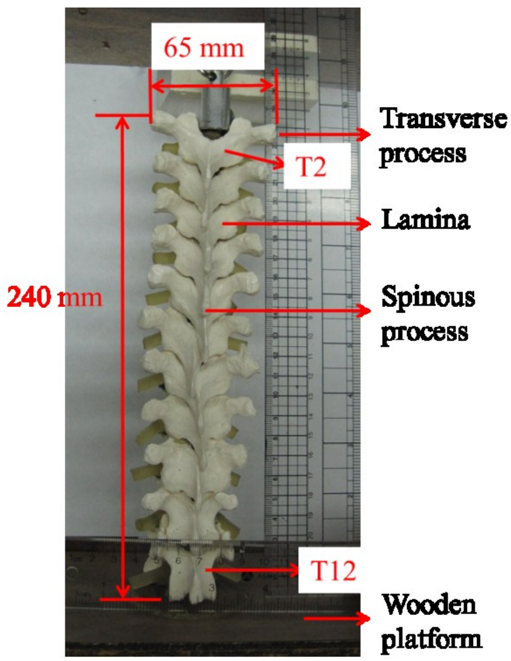


Figure 3.2 The Sawbones spinal column phantom including vertebrae T2 to T12.

3.1.2 Experimental setup

The phased array ultrasound system described in Chapter 2.4.1 with a 5MHz 64-element transducer and the mini-wheel encoder was used in this study (Figure 2.9-2.11). Twelve elements were set in a group and the dynamic depth focusing (DDF) was used for the data acquisition.

The cadaveric vertebra was secured at the bottom of a small water tank with the left and right transverse processes set at the same horizontal level to minimize vertebral axial rotation (VAR) (see Figure 3.3). A 2 mm thick polypropylene sheet with an US speed of 1628 m/s, calculated by the pulse-echo method (Bushberg *et al.*, 2002), was used to mimic the skin. The sheet was supported by screws and placed about 7 mm above the spinous process of the phantom. The tank was filled with water until it just covered the sheet. The human skin thickness varies between 1.55 mm and 2.54 mm (Laurent et al., 2007) with an average US speed 1645 m/s for epidermal layers and 1595 m/s for dermal layers respectively (Moran et al., 1995). Therefore, the polypropylene sheet was considered as an appropriate skin mimics. Water has an US speed of 1480 m/s similar to 1540 m/s for soft tissue and was used to simulate the soft tissue. The whole experimental setup was designed to mimic the human back including skin, soft tissues, and vertebra. The schematic diagram describes the experimental setup shown in Figure 3.3 and the experimental setup is shown in Figure 3.4a.

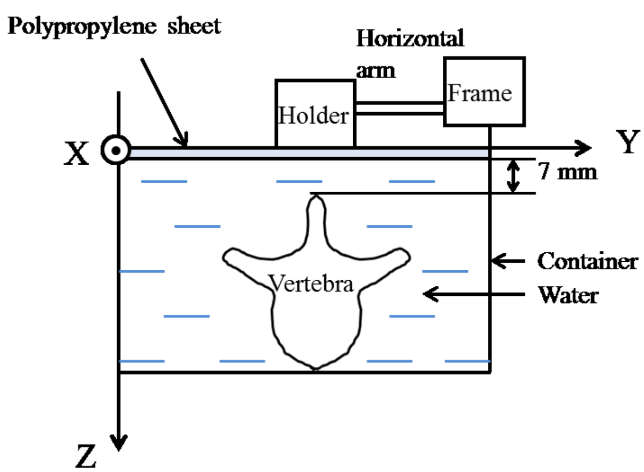
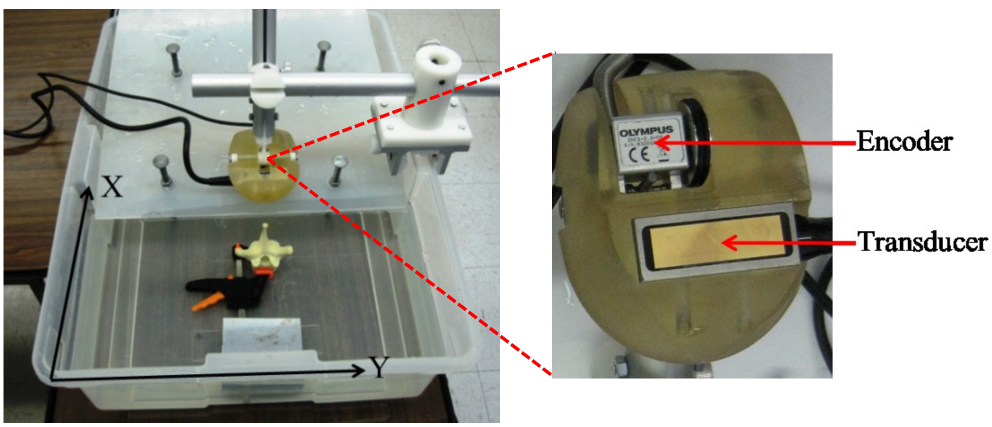


Figure 3.3 The schematic diagram of the experimental setup.

The transducer and the encoder were placed in a custom holder (Figure 3.4b). The function of the encoder was to record the movement of the transducer. The holder was mounted on an aluminum vertical arm and guided by a horizontal arm attached to the edge of the water tank (Figure 3.4a). The horizontal arm was engraved in metrics to indicate the *y*-position of the holder. During scanning, one operator held the holder arm and moved the glider forward along the *x*-direction with a speed of no more than 5 mm/s. The other operator controlled the software to record the data. At the moving speed faster than 5 mm/s, the transducer would not fire ultrasound signals since it exceeded the ability of the equipment to catch the signals, causing null or dead US image, which could be observed from the Tomoview software during scanning. When the null image was observed, the operator then needed to slow down the scanning speed. Among many other factors such as sampling frequency, pulse repetition frequency (PRF), group element number, and number of channels, the scanning speed depended on the depth of interest because a longer listening time was required for the transducer to record echoes from a deeper depth. A bigger depth would require a slower speed.

For the spinal column phantom, the experimental setup was similar to the single vertebra phantom. The spinal column phantom was immersed into water with the polypropylene sheet covered to stimulate a human spine. During scanning, one operator moved the glider and holder to scan along the surface of the sheet, the other operator controlled the software interface to record data.



(a) (b)

Figure 3.4 (a) Experimental setup, and (b) the holder with the mini-wheel encoder and transducer.

3.1.3 Data acquisition and processing

For ease of discussion, an XYZ-coordinate system is defined as shown in Figure 3.5. The X and Y axes are similar to the axes shown in Figure 3.4a. The X-axis denotes the scan direction. The Y-axis defines the element index or index of the transducer. The Z-axis is the depth axis. The probe moved along the X-axis. A 12-element group defined an A-line ultrasound signal, which is a distance-amplitude signal. Each A-line had 7418 points separated by 0.01 μsec interval.

Since each group was shifted by one element interval, the spacing between A-lines was 0.6 mm. A maximum of fifty-three A-lines formed an image or frame with an effective aperture 31.2 mm. Each frame generated a 2D transverse view of the vertebra in the Y-Z plane (perpendicular to the X-axis). To convert the temporal interval Δt to depth interval Δz, one would divide Δt by 2 to account for the two-way travel time of the US and multiply it by 1480 m/s. The resultant depth interval was 0.0074 mm. At a 5 mm/s encoder speed, the spacing interval between frames was 1 mm. Signal averaging and data compression were not applied.

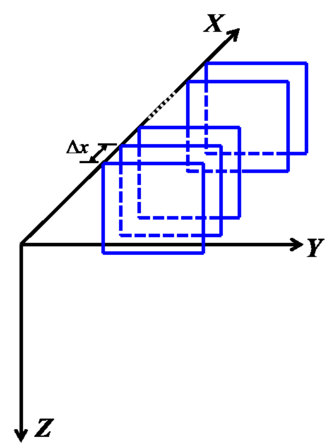


Figure 3.5 The coordinate system: X is the scan axis, Y is the axis for element index, and Z is the depth axis. The squares represent acquired frame.

Since the distances between the transverse processes (horizontal width) of each vertebra phantoms were wider than the effective aperture of the transducer array, each frame acquired by one scan did not cover the horizontal width of the vertebrae. Thus, three scans were required to cover both of the phantoms. A 70 mm × 81.2 mm area and a 240 mm × 81.2 mm area on the XY-plane defined the scanning area for the vertebra phantom and the spinal column phantom, respectively. After finishing one line scan, the holder was translated back to the beginning of the scan and moved along the Y-axis to a new index-position for the new scan. The interval between two scans along Y-axis was 25 mm. The starting (x, y) positions of each scan were recorded; the multiple scan files were merged to form a single data file of the whole phantom using the Tomoview™ software. When the overlapped data was merged together, the intensity of the resultant voxel was assigned using maximum intensity projection (MIP) method. The MIP method uses the maximum value along the projected direction to represent the intensity as shown in Figure 3.6. The merged data had a depth resolution of 0.074 mm after 10-fold decimation. Hereafter, a frame was referred to a merged frame formed from the three scans at the same *x* position.

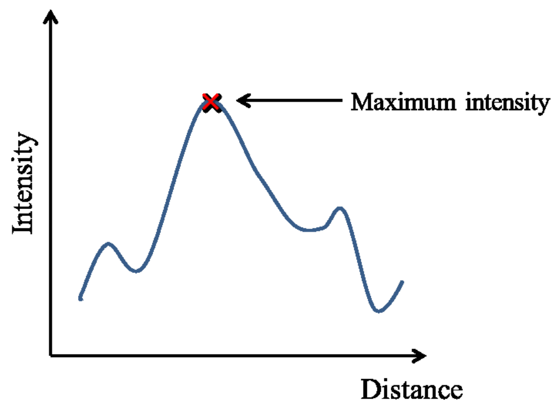


Figure 3.6 The maximum intensity projection method.

The data was further reconstructed to form the sagittal and coronal views on the XZ-plane and XY-plane respectively. The images could be displayed either by single slice or by merging the contiguous slices with varying slice thickness. The images were displayed with 16 colors, which linearly interpolated the intensity. The red color was for hyper-echoic area (>86.7%), blue was hypo-echoic (<33.3%), and white was anechoic (0%).

3.1.4 Vertebra phantom results

The transverse, sagittal, and coronal views of the vertebra phantom are shown in Figure 3.7, Figure 3.8, and Figure 3.9 respectively. The images were compared with the phantom in Figure 3.1a-c.

Three frames obtained at three scanning positions x_1, x_2, and x_3 of the vertebra phantom (indicated in Figure 3.1c) are shown in Figure 3.7a-c. The transverse view images showed the SAP, TP, laminae, and SP in their respective frames. There were 51 frames traversing the whole vertebra. Their projection on the YZ-plane provided an overall comparison of the echo strength from the vertebra structures (Figure 3.7d). The image showed the shape of a “W” as compared with the front view of the vertebra including the TP, laminae, and SP (Figure 3.1a). The echoes from the SP and laminae were relatively strong (in red color) compared to those from the TP and SAP.

The projection of all the sagittal images on the XZ-plane also displays similar observations about their relative reflections from the SAP, TP, laminae and SP (Figure 3.8). The shape of the sagittal view was relatively similar to the shape of the posterior area of the vertebra shown in Figure 3.1c.

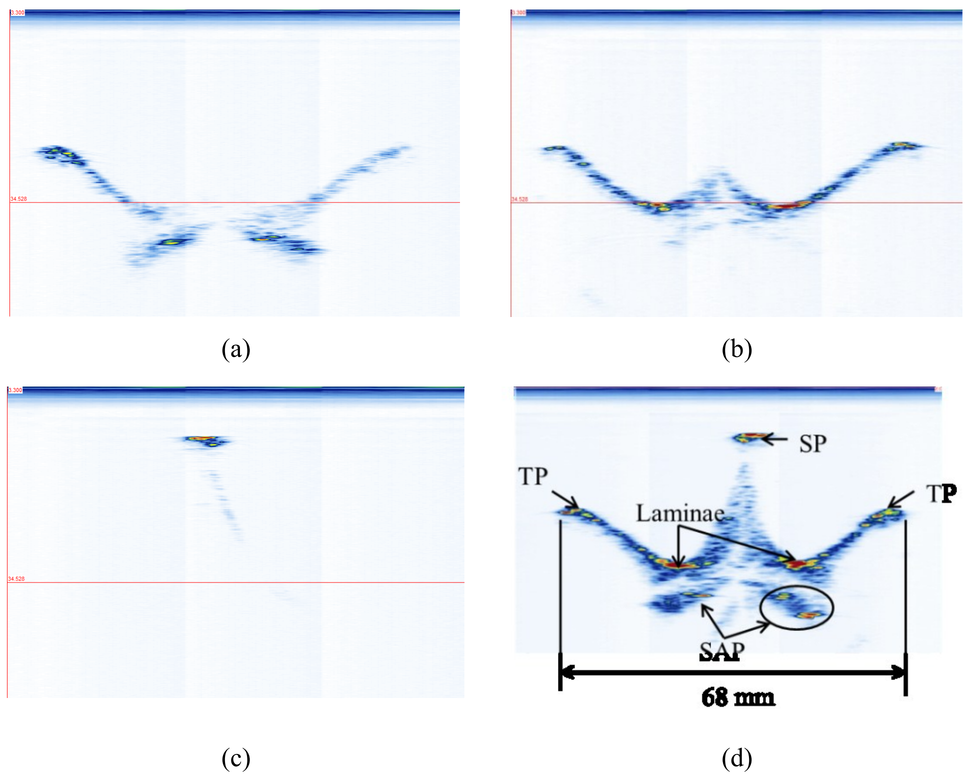


Figure 3.7 Transverse views of the vertebra phantom at different scanning positions as indicated in Figure 3.1c: (a) at x_1, (b) at x_2, (c) at x_3, and (d) stacking of 51 frames.

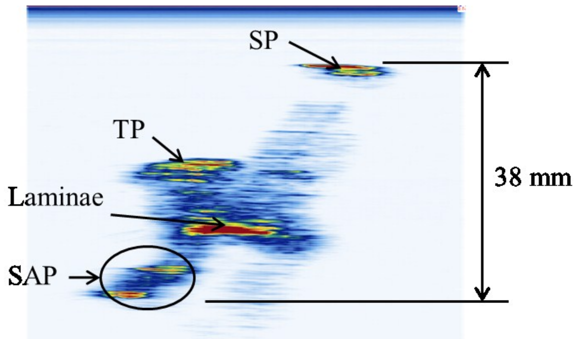
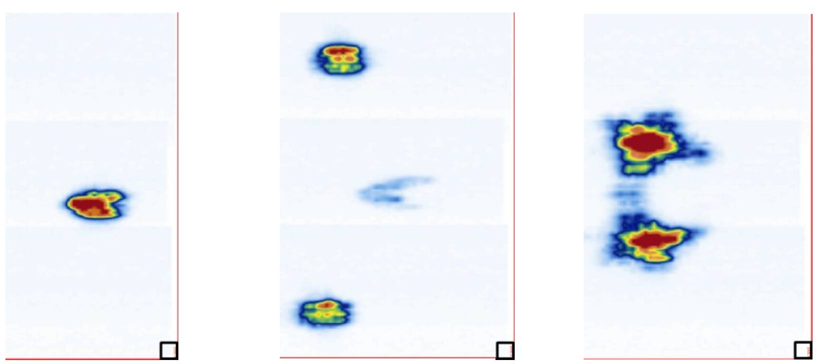
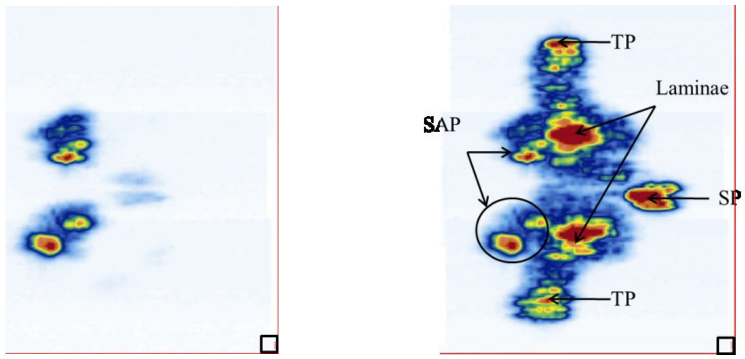


Figure 3.8 Projection of all sagittal images on the XZ-plane.

Coronal views are shown in Figure 3.9 at four different depths: z_1, z_2, z_3, and z_4 as shown in Figure 3.1a, indicating the structures of SP, TP, laminae, and SAP respectively. The 2D images were displayed with different slice thickness. Due to the uneven surfaces and curvature of the structures, thicker slices were necessary to delineate the structures and enhance the echo strength. The number of slices (equivalent thickness) used were 28 (2.07 mm), 24 (1.776 mm), 20 (1.48 mm), and 74 (5.476 mm) for the SP, TP, laminae, and SAP respectively. Figure 3.9e displays a coronal image formed by stacking all XY-plane images from the top of the SP to the bottom of the SAP, covering about 38 mm thick of the vertebra.

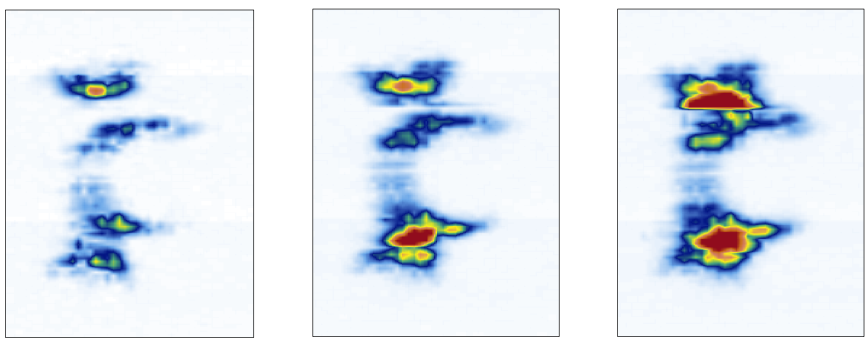


(a) (b) (c)

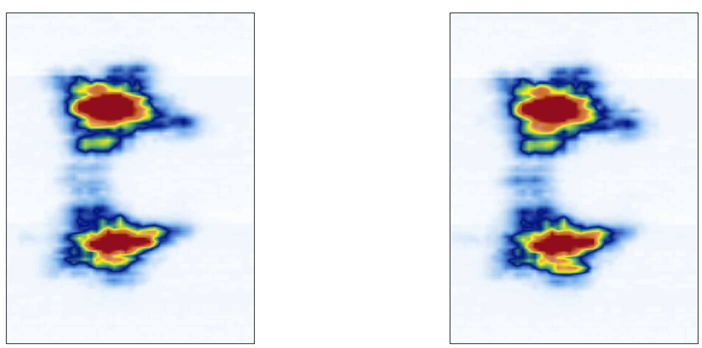


(d) (e)

Figure 3.9 Reconstructed coronal views at different imaging depths showing different structures of the vertebra: (a) SP (at z_1), (b) TP (at z_2), (c) laminae (at z_3), (d) SAP (at z_4), and (e) image stacking from the top of the SP to the bottom of the SAP. The square at the southeast corner of each image is used as a point of reference.



(a) (b) (c)



(d) (e)

Figure 3.10 The reconstructed coronal views of the laminae of different thickness: (a) 0.074 mm (1 slice thick), (b) 0.37 mm (5 slices thick), (c) 0.74 mm (10 slices thick), (d) 1.11 mm (15 slices thick), and (e) 1.48 mm (20 slices thick).

To verify the positions of the vertebrae structures that were recognized from the ultrasound images, the dimensions of the structures were further compared using measurements from both the phantom and its corresponding US images. Six measurements were performed on the phantom, which are shown in Figure 3.11. Three measurements, the distance between the centers of laminae (L1), the distance between the centers of the TP (L2), and the length of the vertebra (L3), were obtained from the transverse and coronal images. In addition, other dimensions, the height between SP and laminae (H1), the height between SP and TP (H2), and the height between SP and SAP (H3), were taken from the transverse and sagittal images using the top edge of each structure. A digital caliper was used to measure the dimensions on the vertebra phantom, and the cursors in Tomoview software were used to measure the dimensions on the US images. Each measurement was repeated three times by the same operator to perform the calibration. The means and the standard deviations of the measurements are listed in Table 3.1. Differences between the measurements from the phantom and the US images were less than 1 mm (<4%). To define the centers of laminae on the phantom was more difficult than the others, therefore L1 had bigger error.

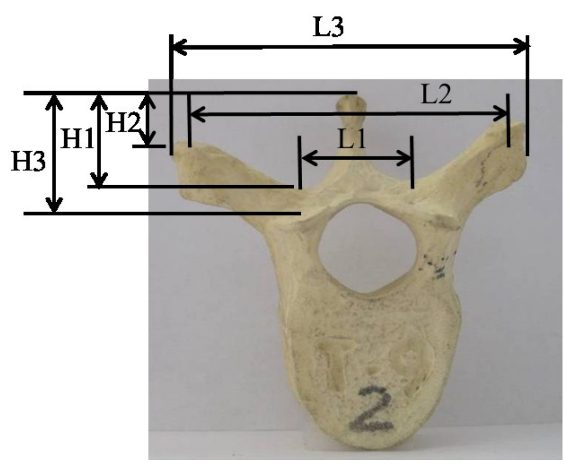


Figure 3.11 The six measurements labeled on the phantom. L1, L2 and l3 indicate the distance between the centers of laminae, the distance between the centers of the TP, and the length of the vertebra, respectively. H1, H2, and H3 indicate the height between SP and laminae, the height between SP and TP, and the height between SP and SAP, respectively.

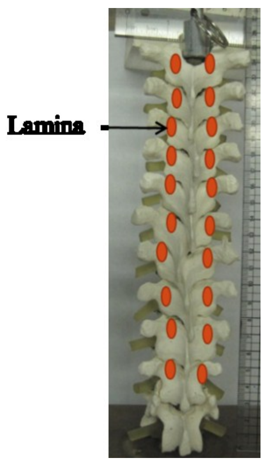
Table 3.1 Comparison between the measurements from vertebra phantom and US images.

	Measurements from phantom (mm)	Measurements from images (mm)	Difference (mm)
Length between centers of lamina (L1)	22.2±0.5	23.0±0.35	0.8 (3.6%)
Length between centers of TP (L2)	60.6±0.5	61.2±0	0.6 (1.0%)
Length of the vertebra (L3)	67.3±0.3	66.6±0	0.7 (1.0%)
Height between SP and lamina (H1)	26.2±0.6	25.9±0.04	0.3 (1.1%)
Height between SP and TP (H2)	15.9±0.3	15.3±0.04	0.6 (3.8%)
Height between SP and SAP (H3)	32.7±0.3	32.3±0.09	0.4 (1.2%)

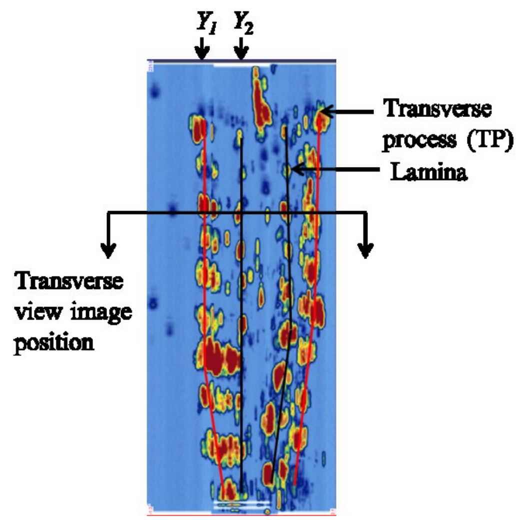
3.1.5 Spinal column phantom results

The figure of the spinal column phantom and the corresponding coronal view image are shown in Figure 3.12. The orange ellipses drawn on the phantom figure indicate the laminae areas based on the vertebral structures described in Chapter 2.2.3 (Figure 3.12a). The projection coronal view (to 38 cm depth) is shown in Figure 3.12b. The transverse process (TP) and lamina could be recognized where the outside red areas indicated the TP (linked with the red lines) and the inner red areas indicated laminae (linked with the black lines).

Sagittal view images of the phantom are shown in Figure 3.13. The TP and laminae could be distinguished by their different depths from the projected sagittal image (Figure 3.13a). On the coronal view (Figure 3.12b), the TP and lamina are located at different position along the Y-direction. Two sagittal images with single slice at Y_1 and Y_2 were exported to separate the TP and laminae (Figure 3.13b and c). The sagittal image at Y_1 included the TP on one side and the sagittal image at Y_2 showed the laminae on the same side of the phantom, in which the TP and laminae were not at the same depth level.

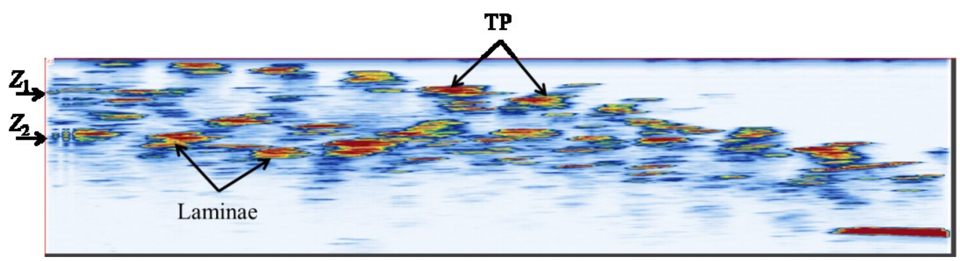


(a)

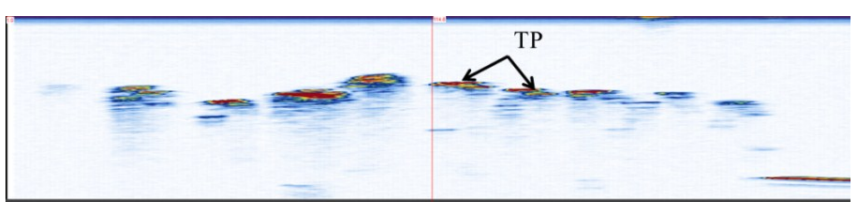


(b)

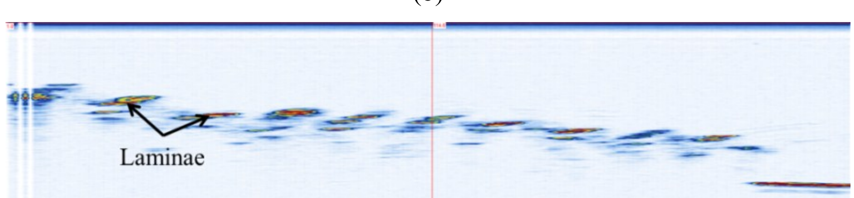
Figure 3.12 Spinal column phantom images: (a) phantom with orange ellipses indicating laminae, (b) coronal view of the phantom with the transverse process and laminae recognized, Y_1 and Y_2 pointed the positions of sagittal images in Figure 3.13.



(a)



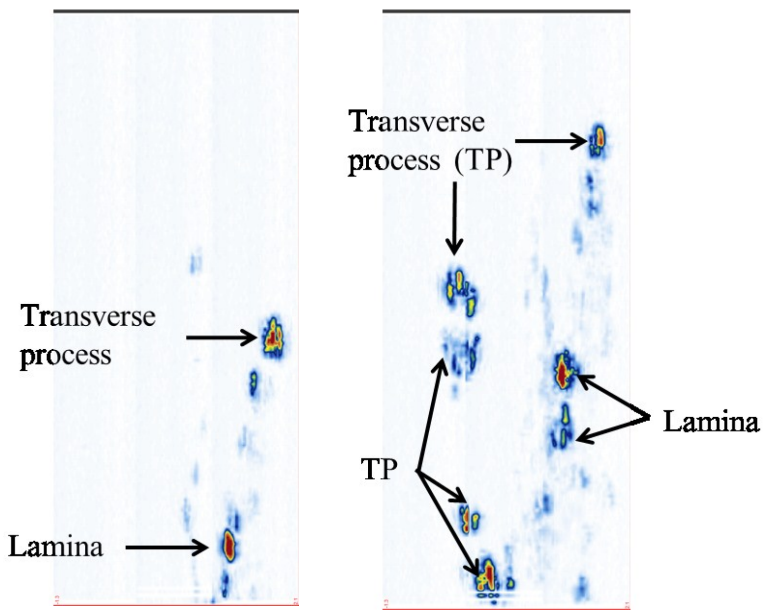
(b)



(c)

Figure 3.13 Sagittal images of the spinal column phantom: (a) projected sagittal image, Z_1 and Z_2 pointed the positions of coronal images in Figure 3.14, (b) single slice at Y_1 position, and (c) single slice at Y_2 position. The positions were indicated in Figure 3.12b.

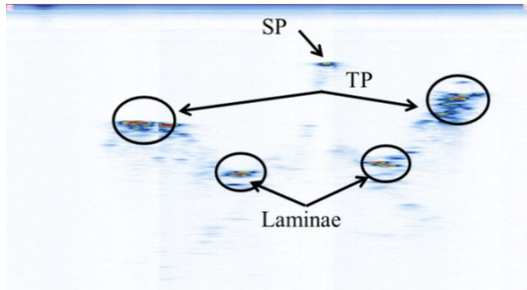
Two coronal images of the phantom at the depth Z_1 and Z_2 (Figure 3.13a) are shown in Figure 3.14, from which only several reflectors could be recognized because the structures from different vertebrae were not at the same depth. The reflectors of the same structures from the whole phantom, such as TP and lamina, could not be obtained in one single coronal slice. When vertebral axial rotation occurred, a single slice coronal view was not able to show both side of the structures from the same vertebra. Thus, projected coronal image was required to present the reflectors of the spinal column as shown in Figure 3.12b.



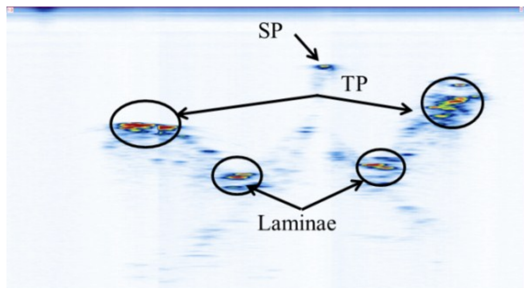
(a) (b)

Figure 3.14 The single slice coronal images of the spinal column: (a) at Z_1, and (b) at Z_2.

The transverse images from the vertebra at the selected position (Figure 3.12b) are shown in Figure 3.15. From the single slice transverse view image shown in Figure 3.15a, the TP, SP and lamina could be identified. Seven slices were applied to form the projected transverse image (Figure 3.15b) to enhance the reflected signals in which the TP, SP and lamina were easier to be recognized. However, compared with the transverse image of the single vertebra (Figure 3.7d), the SAP could not be recognized because the SAP was underneath the inferior articular process.



(a)



(b)

Figure 3.15 The transverse images: (a) the single slice transverse image from the location selected at Figure 3.12b, (b) the projected transverse image from the same vertebra including 7 slices.

3.2 *In-vivo* experiment

3.2.1 Volunteer

A healthy female volunteer (26-year, BMI: 19.4) with no scoliosis consented to participate in this pilot study.

3.2.2 Experimental setup

The volunteer wore a gown with the back opened during the scanning process. She sat straight inside the customized frame (Figure 2.16) and looked forward with arm resting on a metal bar in front of her chest position. This set up was used to minimize the body's sway and the gravity effect still applied to the spine. Ultrasound gel was applied to her back to ensure good coupling between the transducer and the skin. One operator controlled the laptop to acquire the data under the same setup as the *in-vitro* experiment. The other operator moved the holder vertically along the back scanning from T4 to L4. Based on the *in-vitro* experiment results, it was realized that a single vertical scan could not cover the full structure of each vertebra. Therefore, prior to the full spinal column scan, an inspection view was performed to investigate how many scans were required to capture the full structure. As a result, the operator found two scans were required to cover the entire spine column.

3.2.3 Results

The projected coronal image of the volunteer is shown in Figure 3.16. More noise was observed in the image when compared with the coronal view from the phantom (Figure 3.12b). The laminae were more difficult to distinguish, especially when there were strong reflectors from other bone structures in the lumbar part.

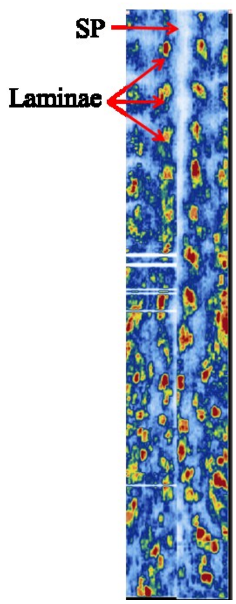
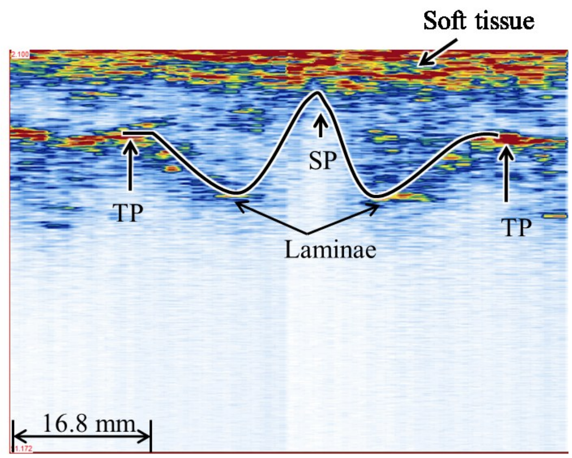
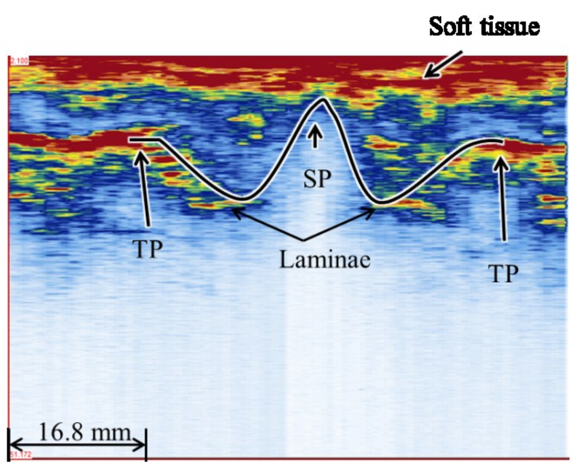


Figure 3.16 Projected coronal view image from the healthy participant.

A transverse image of the thoracic vertebra T4 of the volunteer is shown in Figure 3.17a. A stacked image of 11 slices from the same vertebra is also presented in Figure 3.17b. A 'W' shape consisting of SP, TP and lamina could be recognized from the transverse images. The reflected signals from the heterogeneous soft tissues were strongly represented by the top layer with red color; it was difficult to separate the SP from the soft tissue since they were linked together and both reflected strong signals. A shadow zone, where few US could go through, existed under the bone surface. Thus the SP could be indirectly recognized by identifying the white area in the middle of the 'W' shape. The contrast between the shadow zone and the other areas became stronger on the projected transverse image (Figure 3.17b), which made it more confident to localize the SP. The laminae were recognized as the red areas just besides the SP shadow zone at the lowest positions in the 'W' shape. However, it was difficult to define the locations of the TP since the US reflections from the ribs linked with the TP signals. Therefore, the SP and lamina were considered as the bony landmarks to be used in future investigations.



(a)



(b)

Figure 3.17 Ultrasound transverse images from the healthy participant: (a) single frame of transverse view, (b) transverse view with 11 stacked frames.

3.3 Discussions

Ultrasonography is a successful diagnostic imaging technique to image soft tissues. However, using US to image bone is uncommon because of the strong reflections and high attenuations when ultrasound goes through soft tissues. The previous studies indicated that ultrasound could image bone (Le *et al.*, 2010; Zheng *et al.*, 2007). When US encounters an interface separating two media with different acoustic impedance $\hat{Z}$, a portion of energy is reflected at the interface. The amount of reflected energy R depends on the impedance contrast (normal incidence):

$$R = \left[\frac{\hat{Z}_{bone} - \hat{Z}_{tissue}}{\hat{Z}_{bone} + \hat{Z}_{tissue}}\right]^2 \tag{3-1}$$

The soft tissue-bone interface is a strong reflector. The impedances of bone and soft tissue are around 7.8×10^{-5} rayls and 1.63×10^{-5} rayls, respectively (Bushberg *et al.*, 2002), 43% of the incident energy is reflected. Based on this reflection theory, studies had demonstrated that echograms could provide valuable information on the surface structures of the spine (Furness *et al.*, 2002; Li *et al.*, 2010, 2012; Suzuki *et al.*, 1989).

In this study, a normal incident beam was chosen and the time delay function on the transducer was used for depth focusing. Since the surface of the vertebra is uneven and curved but the lamina, TP and SP areas are relatively flat; the reflection signals from these areas are relatively stronger. However, soft tissues attenuate the US signals significantly. Considering a 1-cm thick soft tissue, the amplitude of the echo is reduced by around 50% with the average attenuation coefficient of soft tissue 0.55 dB/MHz·cm, and US signals are attenuated as they travel to and from the tissue-bone interface. Thus, projection US images with multiple slices were required to enhance the reflection signals. Also, through the experiment results, one single slice could not include all bony markers from the image. Multiple slices were needed to display the information of the required bony markers.

Comparing the *in-vitro* and *in-vivo* studies described in this chapter, the participant's data presented further challenges to image a human spine. The scattering from the inhomogeneous soft tissues showed up as strong reflection signals at the top of the transverse images, which was mixed with the reflected signals from the SP. Despite of this, the SP was still recognizable as one could use the shadow zone under SP and the trace of W-shape to identify the area. However, the TP structure overlapped with ribs on the thoracic region which made it difficult to be distinguished from the ribs (Figure 3.17). Hence, the laminae and the SP were considered as the more consistently detectable markers in this study as compared to TP and SAP.

Furthermore, both the *in-vitro* and *in-vivo* experiments required multiple scans on the phantoms to cover the entire structure. For the *in-vitro* experiment, since the scanning surface was flat (z is constant), the stitching process was manageable and accurate. In this case, the reference point was the starting point (x, y) of the second scan, which provided information to merge the data from two scans. However, the merging issue became a challenge for *in-vivo* study. The 2D encoder could not provide the 3D coordinates of the reference point. Also, the unequal thickness of the soft tissue added more variation as it also depended on the pressure applied by the operator during the ultrasound scan. In addition, the scoliotic spine from AIS patients could cause back asymmetry, which introduce a more uneven surface of the back. To minimize merging problem, a longer transducer array with 128 elements was used after this point in all the further studies. A single scan from the longer transducer could cover the SP and laminae areas, which was able to eliminate the merging problem.

3.4 Summary

According to ultrasound theory, it was possible to image certain bony structures of a vertebra. Both the *in-vitro* and *in-vivo* experiments indicated that the spinous process and laminae could be recognized as the ultrasound markers. Also, from the *in-vitro* and *in-vivo* studies, projection images were required to enhance the reflection signals and display the reflectors more clearly. To avoid the stitching problem, a 128 element transducer would be used in all the studies after this chapter. Before performing the ultrasound measurements, which landmarks, SP or laminae, could provide more reliable measurements needed to be investigated.

Chapter 4.

A comparison of the coronal curvature measurements on radiographs by the center of pedicle method and the spinous process angle at apex

Chapter 3 indicated that the laminae and spinous processes were the bony landmarks, which could be identified on ultrasound spine images (Chen *et al.*, 2012). Further study was required to determine which landmark could provide better assessment of the spinal deformity. As the laminae and pedicles are attached to each other and are considered a rigid body, they shift and tilt together. Since the radiograph is the standard imaging method and Cobb angle is the clinical standard to assess the severity of spinal deformity, this chapter proposed two measurement methods: the center of pedicle (COP) and spinous process angle at apex (SPAA) on radiographs to measure the coronal curvature and compared to the standard Cobb angle method. The intra- and inter-observer reliabilities of these two methods were analyzed. Also, the comparisons of the COP vs. Cobb, and the SPAA vs. Cobb were investigated to find out the differences between the proposed methods and the standard method.

4.1 Assessment of spinal deformity

In a standard scoliosis clinic, a standing posteroanterior (PA) radiograph is used to measure the Cobb angle to assess the severity of scoliosis. On the PA radiographs, laminae cannot be seen but pedicles can be identified. The projected pedicles overlap with the positions of laminae on radiographs. Mehta *et al*. (2009) reported a strong correlation between the Cobb angle and the angle obtained using the pedicle method. Another study also reported that the orientation of lamina formed a useful reference plane for pedicle screw insertion, which meant the orientation of lamina was correlated to the pedicle orientation (Bayley *et al.*, 2010). Since the lamina is identifiable on the ultrasound image, it is hypothesized that the orientation of the lamina can be used to measure the coronal curvature of adolescent idiopathic scoliosis (AIS). Besides the lamina method, the spinous process angle (SPA) has been used to evaluate the Cobb angle during orthotic treatment, and a high correlation was found between the SPA and the Cobb angle (Li *et al.*, 2010, 2012). However, Herzenberg *et al.* (1990) reported that the SPA underestimated the severity of the coronal curvature, especially when axial rotation existed. To investigate which landmarks, spinous process (SP) or lamina, could provide a reliable method for the coronal curvature of AIS, the center of pedicle (COP) method and the spinous process angle at apex (SPAA) were proposed.

To develop a new method to measure the coronal curvature of AIS, a comparable reliability to the Cobb angle method is required. The hypothesis of this study is that the Cobb, COP and SPAA methods have the similar reliability on the coronal curvature measurements. Thus, the intra- and inter-observer reliability of the Cobb method and these two proposed methods needed to be investigated.

4.2 Materials and methods

4.2.1 Study population

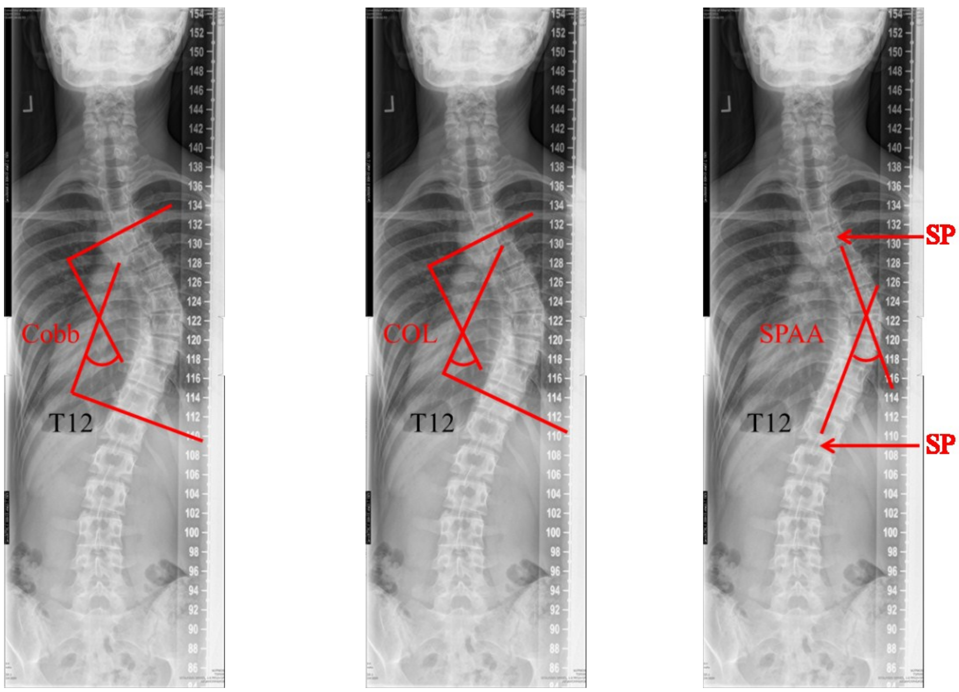
A retrospective study of 53 posteroanterior (PA) radiographs from AIS subjects (aged 9-18 years, 47 F and 6 M, Cobb: 36°±19°) were randomly selected from local scoliosis clinical records from 2009 to 2011. These subjects had no surgery prior to the study and were not wearing a brace when the radiograph was taken. All radiographs were obtained on a digital radiography system (Digital Diagnost, Philips, Canada). Patient identifiable information was removed from the images and the files were given a coded name. Three observers used their own judgment to record the curve measurements including the end vertebra selection. Thus, not all the curves were reported by every observer at every repetition. A total of 81 common curves, measured by each observer at each repetition, were found and assessed. Among these, one subject was measured with triple curves and 26 subjects were measured with double curves. The curves were divided into three groups for analysis to determine if the curve severity had an effect on the measurements: mild (<25° Cobb as reported in the clinical record, 31 curves), moderate (25°- 45°, 21 curves), and severe (>45°, 29 curves).

4.2.2 Observers

Three observers who had varying levels of experience carried out the assessment twice, independently, for each measurement: Cobb angle measurement, COP angle measurement and the SPAA measurement. The Cobb angle, as measured in the clinic by health professionals, was blinded to the observers. Observer 1 (O1) had experience reviewing and measuring radiographs for 15 months, observer 2 (O2) had experience reviewing and measuring radiographs for 6 months, and observer 3 (O3) had no experience. Prior to the study, O3 reviewed 10 radiographs for practice and was taught by an expert who had 15 years' experience to measure the Cobb angle. There was no training for the other two observers.

4.2.3 Measurement methods

The three observers measured the coronal angle by the Cobb method, the COP method, and the SPAA twice (2 sessions) with one week interval to blind to the earlier assessment. The location of T12 for each radiograph was mutually agreed upon by the three observers together prior to measurements and labeled on the image to provide consistency in labeling vertebrae. The ImageJ (NIH, USA) program was used to measure the angles. Measurements of the Cobb method and COP method (Figure 4.1) were performed with a three day interval to reduce bias. The measurements of the COP and SPAA methods were performed on the same day since different vertebrae were used as landmarks. For the Cobb angle measurement (Figure 4.1a), the end vertebrae of each curve (the most tilted vertebrae at the cephalic and caudal ends of a curve) were determined first and recorded. Then, lines were drawn along the cephalic and caudal endplates using the angle tool of ImageJ. The COP method used the centers of the pedicles on the most titled vertebrae to determine the tilt angle (Figure 4.1b). Each observer used their own recorded end vertebrae from the Cobb angle measurement to draw the lines through the center of pedicles. To measure the SPAA (Figure 4.1c), the apex of a curve was first determined. The SPAA was defined as the angle between two lines drawn through the spinous process tips of the upper and lower two vertebrae of the apex. There was a one week interval between sessions and all prior measurements were blinded on the second sessions. The measured angles as well as the end vertebrae and apex of each curve were recorded on both sessions.



(a) (b) (c)

Figure 4.1 Three measurement methods to measure the coronal curvature on radiographs: (a) Cobb, (b) COP, and (c) SPAA.

4.2.4 Statistical analysis

The intra- and inter-observer reliabilities were analyzed for each curve severity group and also the entire group.

The intra-observer differences of each method were calculated by using the mean absolute difference (MAD) and the standard deviations (SD) between the two sessions of each observer on each curve severity group. The intra-class correlation coefficient (ICC) (2-way random and absolute agreement) was used to analyze the measurements between the two sessions of each observer on each method (McGraw and Wong, 1996; Shrout and Fleiss, 1979). The ninety-five percent confidence interval (CI) was also calculated. The Currier criteria for ICC values were adopted: 0.90-0.99 = high reliability, 0.80-0.89 = good reliability, 0.70-0.79 = fair reliability, <0.69 = poor reliability (Currier 1990).

To evaluate the inter-observer reliability, the ICC and MAD±SD were applied for the comparisons between any two observers for all three curve severity sub-groups of each method.

The ICC (2-way random and absolute agreement) was also used to compare the different methods using the average measurements of 2 sessions between Cobb vs. COP and Cobb vs. SPAA on each curve severity group. The differences between the measurements for the agreement comparisons were also analyzed by the MAD±SD.

The standard error of measurement (SEM) is not applied because the MAD and SD are used and the SEM is related to the SD of difference ($SEM = SD/\sqrt{2}$) (Stratford 2004).

The Bland-Altman plot (Bland and Altman, 1986; 1999; 2010) was used to show the differences and check if the differences related to the magnitude of the measurements by Pearson's correlation coefficient: 1) intra-observer comparison of the COP method from O1, 2) the inter-observer comparison of the COP method between O1 and O2, and 3) the differences between the Cobb method against the COP from O1. The significant level was set at $\alpha = 0.05$, that the differences was significantly correlated with magnitude when the P-value was smaller than α.

4.3 Results

4.3.1 Intra-observer reliability

Table 4.1 summarizes the intra-observer reliability of each curve severity group on each method for three observers. For the Cobb method measurement, the ICC values indicated the results were good to highly reliable in the moderate (0.88, 0.91, 0.84), severe (0.96, 0.92, 0.90) and overall (0.99, 0.98, 0.98) curve groups for all 3 observers. Only O1 showed good reliability (0.82) in the mild curve group while the O2 and O3 showed fair reliability (0.77, 0.71). For the COP method, the measurements of the severe curve group for each of the three observers (0.95, 0.86, 0.91) showed high reliability, but for the mild and moderate groups, respectively, only O1 showed a good to high reliability (0.84, 0.90), while O2 (0.65, 0.79) and O3 (0.77, 0.66) had poor to fair reliability. For the SPAA method, all the ICC values were not as good as the other two methods, which were in the range from 0.27 to 0.77 in the three curve severity groups indicating poor to fair reliability. For the overall measurement, the ICC values showed high reliability on the Cobb (0.99 vs. 0.98 vs. 0.98) and COP (0.99 vs. 0.98 vs. 0.98) methods for all three observers. On the SPAA overall measurement, only O1 showed high reliability (0.94), but O2 (0.75) and O3 (0.74) showed fair reliability.

The MADs between two sessions were small in all curve groups for the Cobb method (<3.7°) and COP method (<4.1°), while much larger in the SPAA method (<11.8°). The standard deviations (SD) were small for the Cobb and COP methods (<3.4°), especially in the mild and moderate curve groups (<2.3°). However, the SD values of the SPAA method had larger values which increased from the mild to severe curve groups especially in O3 (from 4.0° to 10.4°). The larger value of the MAD±SD related to the lower ICC values when compared the same curve severity sub-group using the same method among different observers. Figure 4.2 shows the Bland-Altman plot indicating the differences between the measurements from two sessions of COP method from O1 including all three type curves. Ninety-five percent of the differences fell in the range of -0.9°±4.6°

Table 4.1 Intra-observer reliability of the three methods of each curve severity sub-group and overall for each of the three observers.

Curve severity	Analysis method	O1			O2			O3		
		Cobb	COP	SPAA	Cobb	COP	SPAA	Cobb	COP	SPAA
Mild	ICC	0.82	0.84	0.75	0.77	0.65	0.46	0.71	0.77	0.27
(n =31)	95% CI	0.65-0.91	0.70-0.92	0.34-0.90	0.57-0.88	0.39-0.81	0.12-0.70	0.48-0.85	0.57-0.88	0.05-0.55
	MAD ± SD (°)	1.8±1.5	1.8±1.7	2.1±1.8	2.2±1.9	3.16±2.30	3.7±3.7	2.3±2.1	2.0±2.1	4.1±4.0
Moderate	ICC	0.88	0.90	0.59	0.91	0.79	0.65	0.84	0.66	0.60
(n = 21)	95% CI	0.72-0.95	0.77-0.96	0.22-0.81	0.80-0.96	0.55-0.91	0.33-0.84	0.63-0.93	0.34-0.84	0.24-0.81
	MAD ± SD (°)	1.9±1.7	1.6±1.4	4.5±3.6	1.7±1.6	2.74±1.93	5.6±4.9	1.9±1.9	2.4±2.6	4.5±3.8
Severe	ICC	0.96	0.95	0.77	0.92	0.86	0.50	0.90	0.91	0.38
(n = 29)	95% CI	0.79-0.99	0.72-0.98	0.57-0.89	0.82-0.96	0.73-0.93	0.18-0.72	0.73-0.96	0.78-0.96	0.02-0.65
	MAD ± SD (°)	2.7±1.9	2.2±1.7	5.7±5.1	3.6±3.2	4.07±2.85	11.8±8.7	3.7±3.4	3.9±3.2	8.5±10.4
Overall	ICC	0.99	0.99	0.94	0.98	0.98	0.75	0.98	0.98	0.74
(n=81)	95% CI	0.97-0.99	0.98-0.99	0.89-0.96	0.97-0.99	0.96-0.98	0.64-0.83	0.96-0.99	0.97-0.99	0.62-0.83
	MAD ± SD (°)	2.1±1.7	1.9±1.6	4.0±4.0	2.6±2.5	3.38±2.46	7.0±7.1	2.7±2.7	2.8±2.8	5.8±7.2

MAD = Mean Absolute Difference, SD = Standard Deviation (mean±2SD), which indicated small variations between the two repetitions. The differences were correlated with the curve severity (p=0.02).

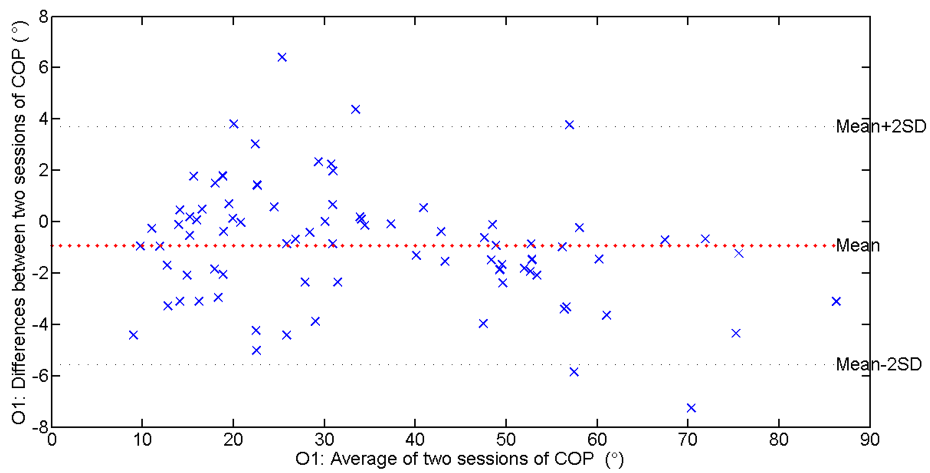


Figure 4.2 Bland-Altman plot of the COP method between two sessions from O1.

The end vertebra selections between two sessions are compared and listed in Table 4.2. For each observer, more errors existed in the mild group compared to the other two groups. Taking observer 1 as an example, only 61.3% ((64.5%+58.1%)/2) of the curve were selected at the same end vertebrae in the mild group, while 78.6% ((76.2%+81.0%)/2) and 86.2% ((82.8%+89.7%)/2) for the moderate and mild group, respectively.

Table 4.2 The number of the end vertebra differences between two sessions.

		Top end vertebra			Bottom end vertebra		
		0	±1	±2 or more	0	±1	±2 or more
O1	Mild	20 (64.5%)	9 (29.0%)	2 (6.5%)	18 (58.1%)	12 (38.7%)	1 (3.2%)
	Moderate	16 (76.2%)	4 (19%)	1 (4.8%)	17 (81.0%)	4 (19%)	0
	Severe	24 (82.8%)	5 (17.2%)	0	26 (89.7%)	3 (10.3%)	0
O2	Mild	13 (42.0%)	9 (29.0%)	9 (29.0 %)	15 (48.4%)	16 (51.6%)	0
	Moderate	9 (42.9%)	10 (47.6%)	2 (9.5%)	14 (66.7%)	7 (33.3%)	0
	Severe	14 (48.3%)	13 (44.8%)	2 (6.9%)	15 (51.7%)	13 (44.8%)	1 (3.5%)
O3	Mild	28 (90.3%)	3 (9.7%)	0	22 (71.0%)	8 (25.8%)	1 (3.2%)
	Moderate	16 (76.2%)	4 (19.0%)	1 (4.8%)	18 (85.7%)	3 (14.3%)	0
	Severe	22 (75.9%)	6 (20.7%)	1 (3.4%)	25 (86.2%)	4 (13.8%)	0

4.3.2 Inter-observer reliability

Table 4.3 shows the inter-observer reliability considering all the three observers on the three measurements methods, Cobb, COP and SPAA. For the Cobb and COP methods, the ICC values indicated good to high reliability in all three curve severity sub-groups except the mild group for COP method (0.74). Both the Cobb and COP methods showed high reliability for the overall comparisons. However, the SPAA method had a poor reliability in all three curve severity sub-groups (ICC: 0.31-0.48) and the overall ICC value of the SPAA was only 0.75. The MAD±SD values were small in all three curve severity sub-groups for both the Cobb and COP methods (<3.7°). Also, the Cobb and COP methods had similar MAD±SD values, which indicated these two methods had similar inter-observer variations. However, the SPAA method had large MAD values (up to 12.4°) and increased from the mild to severe groups.

Table 4.3 The inter-observer reliability for three methods of all observers.

Curve severity	Analysis method	Cobb	COP	SPAA
Mild	ICC (95% CI)	0.83 (0.72-0.91)	0.74 (0.50-0.87)	0.31 (0.06-0.56)
	MAD±SD(°)	1.7±1.4	2.4±2.0	4.2±3.4
Moderate	ICC (95% CI)	0.84 (0.67-0.93)	0.81 (0.60-0.92)	0.48 (0.16-0.73)
	MAD±SD(°)	2.1±1.8	2.1±2.0	6.0±4.5
Severe	ICC (95% CI)	0.92 (0.74-0.97)	0.92 (0.69-0.97)	0.39 (0.08-0.65)
	MAD±SD(°)	3.7±2.7	3.7±2.7	12.4±8.0
Overall	ICC (95% CI)	0.98 (0.96-0.99)	0.98 (0.97-0.99)	0.75 (0.50-0.86)
	MAD±SD(°)	2.5±2.3	2.4±2.2	7.6±6.8

Figure 4.3 shows the Bland-Altman plot of inter-observer comparison of COP method between O1 and O2. Ninety-five percent of the differences between the two observers were within the range of 2.4°±4.8° (mean±2SD), and were independent of the curve severity (p=0.71).

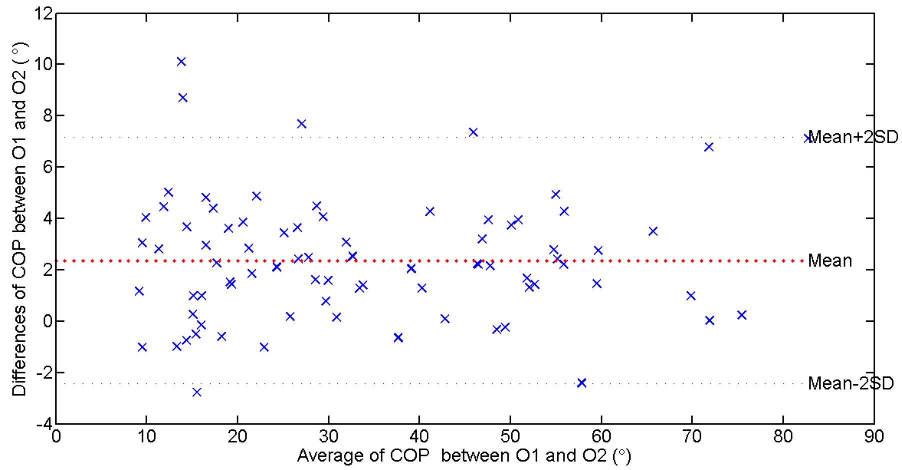


Figure 4.3 Bland-Altman plot of inter-observer comparison of COP method between O1 and O2.

4.3.3 Comparisons between methods

In terms of the inter-method reliability (Table 4.4), all observers showed better ICC values when comparing the Cobb vs. COP than the Cobb vs. SPAA in all curve groups. O1 showed excellent reliability between the Cobb and COP methods (overall ICC 0.99). O2 showed high reliability in the moderate curve group (0.88) and excellent reliability in the severe curve group (0.99), but poor reliability (0.55) in the mild curve group. The overall group also showed excellent reliability between Cobb and COP methods (0.98). O3 had moderately reliable measurements in mild (0.80) and moderate (0.77) curve groups and high reliability in the severe curve group (0.99) and overall curve group (0.99) for the comparisons of Cobb and COP methods. All observers showed poor reliability in all curve groups when comparing the Cobb to the SPAA methods, except the overall group in O1 (0.83).

The MAD±SD for the comparison of Cobb and COP methods in all three observers were in a small range with the maximum value of 3.6°±2.5°, which corresponds to the lowest ICC value in the mild curve group of O2. The MAD±SDs between the Cobb and COP on the O1, O2, and O3 for the overall measurements were 1.5°±1.3°, 2.7°±2.1°, and 2.2°±1.6°, respectively. The higher ICC values corresponded to lower MAD values for the Cobb vs. COP comparison

Table 4.4 The comparisons of Cobb vs. COP, and Cobb vs. SPAA of all three observers.

Curve severity	Analysis method	O1		O2		O3	
		Cobb vs. COP	Cobb vs. SPAA	Cobb vs. COP	Cobb vs. SPAA	Cobb vs. COP	Cobb vs. SPAA
Mild	ICC	0.88	0.31	0.55	0.08	0.80	0.12
(n =31)	95% CI	0.61-0.95	0-0.64	0.03-0.80	0-0.30	0.15-0.93	0-0.40
	MAD ± SD (°)	1.6±1.2	4.8±2.9	3.6±2.5	6.9±4.8	2.2±1.5	10.2±3.5
Moderate	ICC	0.94	0.20	0.88	0.05	0.77	0.07
(n = 21)	95% CI	0.85-0.98	0-0.55	0.69-0.95	0-0.29	0.44-0.90	0-0.30
	MAD ± SD (°)	1.3±1.1	9.7±4.2	2.1±1.6	13.9±6.8	2.4±1.6	16.4±5.6
Severe	ICC	0.98	0.49	0.99	0.27	0.99	0.16
(n = 29)	95% CI	0.98-0.99	0-0.81	0.98-0.99	0-0.63	0.98-0.99	0-0.50
	MAD ± SD (°)	1.6±1.5	12.2±6.8	2.2±1.8	20.2±10.1	2.1±1.7	27.4±8.1
Overall	ICC	0.99	0.83	0.98	0.59	0.99	0.49
(n=81)	95% CI	>0.99	0.07-0.95	0.97-0.99	0-0.83	0.97-0.99	0-0.80
	MAD ± SD (°)	1.5±1.3	8.7±5.9	2.7±2.1	13.5±9.4	2.2±1.6	17.9±9.6

MAD = Mean Absolute Difference, SD = Standard Deviation in the same curve severity groups. However, the MAD±SDs for the comparisons between the Cobb and SPAA methods were much larger, in the range from 4.8° to 27.4°, and increased from the mild group to severe group for each observer. Figure 4.4 shows the Bland-Altman plot for the differences between the measurements between Cobb and COP methods from O1. Ninety-five percent of the differences between the measurements of Cobb and COP methods fell in the range of 0.7°±3.7° (mean±2SD), and these differences were independent of the curve severity (p=0.16).

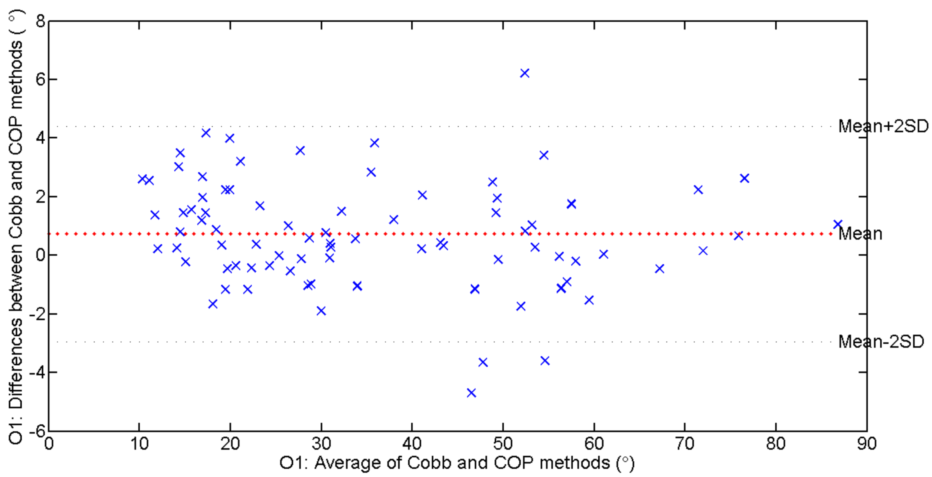


Figure 4.4 The Bland-Altman plot of the comparison between Cobb and COP methods from O1.

4.4 Discussions

The reliability of the Cobb angle had been studied by many authors (Gstoettner *et al.*, 2007; Kuklo *et al.*, 2005; Mok *et al.*, 2008) and the errors of measurement were between three and five degrees for the same observer and five and seven degrees for different observers (Carman *et al.*, 1992; Morrissy *et al.*, 1990; Shea *et al.*, 1998). In this study, the intra-observer comparisons showed excellent intra-observer reliability on both the Cobb and COP methods. The Cobb and COP methods had similar intra-observer reliability and small variation among all three observers. The SPAA method showed the least intra-observer reliability. The reason was that it was difficult to identify the spinous processes from the radiographs when there was VAR. Thus, observers needed to estimate the positions of the spinous processes based on their knowledge of vertebra anatomy when the VAR was significant.

The intra-observer variability of the Cobb method measurements indicated that all three observers reported less reliable measurements in the mild group than the moderate or severe curve groups. The observers were blinded to the selection of end vertebrae between repetitions. The variance in selection of end vertebrae was reported to be the main source of measurement errors (Mok *et al.*, 2008; Morrissy *et al.*, 1990). In the mild curve group, it was more difficulty to select the end vertebrae than the other two groups because in this group all vertebrae had only a mild tilt so there was ambiguity in choosing the vertebrae with the maximum tilt.

The inter-observer analysis showed that the COP method had similar reliability as the Cobb method. The MAD±AD values indicated the inter-observer differences of the Cobb and COP methods were in a small range. However, the SPAA showed the least inter-observer reliability with large range of variation between observers.

When comparing the proposed methods with the standard method, the COP method agreed well with the Cobb method, but the measurements between the Cobb and SPAA methods showed poor agreement in all the three observers. From the literature, it was also reported that the position of the spinous process deviated from the center of the vertebra when there was axial rotation (Cobb, 1948), and the rotation of the vertebra affected the coronal curvature estimation from the spinous processes (Herzenberg *et al.*, 1990). Although a new SPAA method was proposed to use the spinous process, the measurements had limited reliability. Therefore, the SPAA method should not be used for the estimation of coronal curvature. The MAD±AD values and the Bland-Altman plot of the Cobb and COP methods showed that most of the differences were in a small range (<4°), which was within the clinically acceptable measurement error of 5° (Morrissy *et al.*, 1990).

4.5 Summary

In this study, the COP method demonstrated comparable intra- and inter-observer reliabilities with the Cobb method, and showed excellent performance with small range of differences. However, the SPAA method had poor performance compared with the other two methods. The axial rotation affected the SPAA method as the end point of the spinous process was unclear and deviated from the center of the vertebra.

Therefore, the COP method could be applied for the assessment of the coronal curvature of AIS, but the SPAA method would not be suitable for the coronal curvature measurement. The orientations of laminae and pedicles were highly related and they shift and tilt together. Thus, the laminae were determined to be the bony landmarks for US measurements. Then the reliability of using the laminae to measure the coronal curvature on the ultrasound image must be investigated.

Chapter 5.

Reliability study of the coronal curvature measurements on ultrasound images*

Chapter 3 showed that laminae were good ultrasound landmarks with strong reflections on vertebra. Chapter 4 presented that the COP method was comparable to the Cobb method to measure the coronal curvature. Since the orientations of pedicle and lamina were highly correlated (Mehta *et al*, 2009), this chapter extended the previous studies to investigate the reliability and validity of the coronal curvature measurements using laminae as bony landmarks on ultrasound images. The center of lamina (COL) method on ultrasound images is investigated. The intra- and inter-observer reliabilities of the COL method to measure the coronal curvature of spine were investigated on both *in-vitro* and *in-vivo* ultrasound images.

5.1 *In-vitro* study

To validate the coronal measurements on ultrasound (US) images, the COL method was required to be compared with the Cobb angle measured on radiographs, which was the clinical practice method. However, the US experimental preparation was too time-consuming and this *in-vitro* experiment might take several days and water would have been needed during the US data acquisition. To minimize the transportation, the experiments were required to be performed in the same place, which made it difficult to perform this experiment in a radiographic room. Therefore, an alternative method to substitute the radiograph

---

* Part of this chapter has been published. Chen W, Lou EM, Zhang P, Le L, Hill D (2013) Reliability of assessing the coronal curvature of children with scoliosis by using ultrasound images. Journal of Children's Orthopaedics 7:521-529.

was needed. One of the proposed solutions was to measure the coronal curvature from a laser scanner captured image.

Prior to the reliability study, a comparison between the laser scanner (LS) and radiograph was performed using a Sawbones spinal column phantom described in Chapter 3.1.1 to investigate the validity of the measurements from LS image. The phantom was bent to a certain curve magnitude. Both the radiograph and the LS image of the phantom were taken. Two observers measured the horizontal levels of the lower end-plate from T3-T12 using the wooden platform as reference on the LS image and radiograph separately. One observer, who had 1.5 years' experience on measuring Cobb method, performed the measurements on LS image. The other observer, who had 15 years' experience on measuring Cobb angle, performed the measurements on radiograph. The mean absolute difference between the measurements of two images was 0.54°±0.38° (range from 0.03° to 1.07°). The results indicated that the measurements from the LS image were close to the measurements from the radiograph with small errors. Thus, the LS images could be an alternative method to assess the coronal curvature of the spinal phantom.

5.1.1 The spinal phantom

The cadaver spinal phantom (Figure 5.1) is composed of 7 cervical vertebrae (C1-C7), 12 thoracic vertebrae (T1-T12) and 5 lumbar vertebrae (L1-L5). A 4-mm-diameter metal wire threaded through the neural canal of the phantom from the cervical region to lumbar region. The metal wire acted as a support and allowed to bend the phantom in positions that mimicked a series of severity of spinal deformity. Foam filled the gap between the wire and the specimen to fix the vertebrae positions along the wire. One end of the wire was fixed to an aluminum base which was used to position the phantom. Colored stickers (red and green) were pasted on vertebrae to indicate the lamina positions and used for better identification of the vertebrae. Only vertebrae from T1 to L5, which spanned 408 mm, were captured in the study.

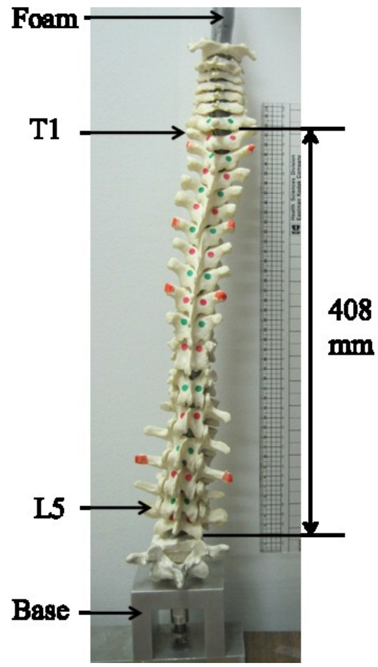


Figure 5.1 Cadaver spinal column phantom.

5.1.2 Laser scanner system

The laser scanner (LS) system consists of a laser camera (Minolta non-contact 3D Digitizer VIVID 910, Japan) and a desktop computer preinstalled with Polygon Editing Tool (PET) software. The wide light-receiving lenses of the camera were used to capture the three-dimensional information, with an accuracy of ±1.40mm, ±1.04mm, and ±0.04mm on the left-right, superior-inferior, and anterior-posterior directions, respectively. After the spinal phantom was placed inside a custom frame at 1.5 meters away from the camera, the coronal view image was captured. The scanner was set to a "FINE" imaging mode, which was able to capture the image with a resolution of 640 by 480 pixels.

5.1.3 Ultrasound system

The ultrasound system was described in Chapter 2.4. A 128-element transducer was used in this study. The dimensions of the 128-element transducer were 64 mm (length) by 10 mm (elevation) and the pitch was 0.5 mm. Fourteen elements were grouped together for depth focusing such that the group aperture was 7 mm, which was close to 7.2 mm of the 12 elements in the 64-element transducer. The groups were separated by one pitch interval. A total of 115 groups were formed to capture a 57.5 mm wide image in a single scan. The dynamic depth focusing (DDF) feature was also applied to achieve a better lateral resolution over an extended depth. The encoder was used to record the position of the transducer along the scanning direction during data acquisition. The transducer and the encoder were mounted inside a holder (Figure 5.2) to move together.

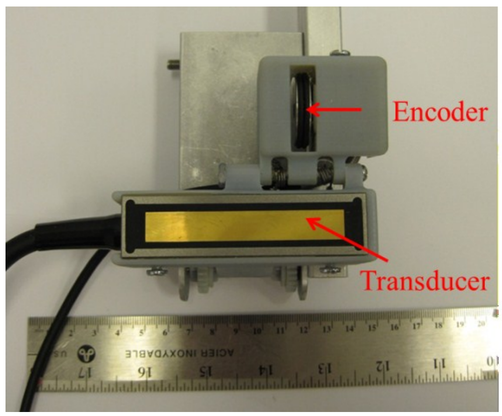


Figure 5.2 The holder with the transducer and encoder.

5.1.4 Data acquisition procedures

A total of 15 different curvature configurations were manipulated on the phantom from mild to severe cases (coronal curvature angle ranged from 10° to 50°). Among the 15 configurations, 11 were major right thoracic with left lumbar curves and 4 were major right lumbar with left thoracic curves based on a reported prevalence of 29% for the curves with structural thoracolumbar/lumbar curve (Lenke, 2005). For each configuration, a double 'S' shaped curve was created and the bottom end vertebra of the upper curve was manipulated as the top end vertebra of the lower curve. One of the configurations is displayed in Figure 5.3. Three procedures were performed to acquire the images: 1) using the LS system; 2) using the US system, and 3) using the LS system again to capture a second set of image to investigate the presence of any motion effect during the experimental procedures.

First procedure: After a spinal curvature configuration was set, the phantom was placed on the top of a table (Figure 5.3). The table was located at the center of the frame. The aluminum base of the phantom was aligned with four markers at the corners. The front view of the phantom was placed paralleled with the XY plane of the frame, which was described in Chapter 2.1.4. The upper end of the 4-mm metal wire was secured to a metal bar which was mounted at the top of the frame. The anterior side of the spine was oriented to face toward the camera to catch the image of vertebral body. The LS image was then captured as the LS1 data set.

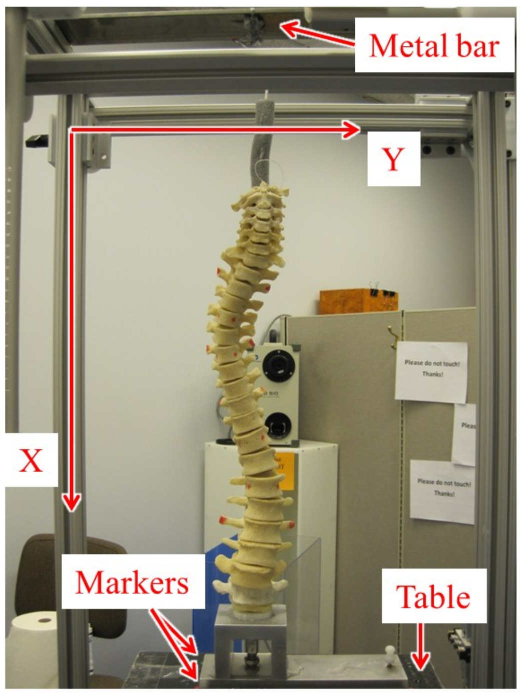


Figure 5.3 Experimental setup of the laser scanner (LS) system.

Second procedure: The spinal phantom was carefully removed from the table. A water tank made of plexiglass was placed on top of the table. The phantom was then put inside the tank and placed at the same location as that in first procedure based on the four corner markers. The wall of the tank was parallel with XY plane of the frame. The posterior side of the spine faced toward the transducer. The tank was then filled with water until the top of the vertebra T1 was immersed (Figure 5.4). Ultrasound gel was applied between the surface of the water tank and the transducer to ensure good coupling during data acquisition. One operator then checked if the width of the captured image could cover the horizontal deviation of the spinal curvature and adjusted the position of the vertical bar along the Y-direction. The top and bottom crossbars were engraved with a scale so that the position of the vertical bar could be adjusted. If a single scan could not cover the curved spinal phantom, two scans with 25 mm overlap along the Y-direction would be performed. During the scan, one operator controlled the horizontal arm mounted with the holder part and scanned on the surface of the water tank from T1 to L5 along the X-direction (vertical bar). The sampling frequency was reduced to 25MHz so that the maximum scanning speed changed to 36 mm/s for faster scanning. The scanning speed of the transducer was kept under 36 mm/s so that images of good quality could be obtained.

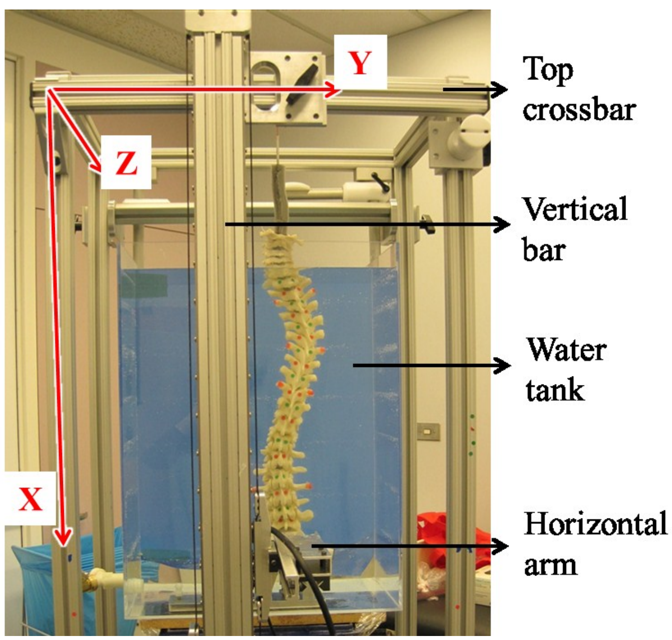


Figure 5.4 Experimental setup of the ultrasound scan.

Third procedure: After the US image was captured, the phantom was removed from the water tank first. Then the water tank was taken away from the table, and the phantom was placed back to the same position as procedure one. During the movement, the operators tried not to alter the spinal curvature configuration. A second laser scan was performed and images were recorded as

LS2. Data set LS2 was compared with LS1 to determine if there was any change on the curvature configuration due to phantom manipulation before and after US data acquisition.

5.1.5 Data processing

After the LS images (LS1 and LS2) were exported into TIFF format, the image contrast was optimized by an experienced LS operator using ImageJ software (Version 1.45, NIH, USA) on a monitor with maximum resolution of 1280 by 800 pixels. The 30 LS image files, from both LS1 and LS2, were assigned random numbers using the free service from the web site (www.random.org) one week before performing measurements on the LS images (LS1 and LS2).

The 3D US raw data (Tomoview data format) was transformed to MATLAB data format using a custom program. The XYZ-coordinate was correlated to different views of the phantom. The XY-plane indicated the coronal view, the XZ-plane was the sagittal view, and the YZ-plane was the transverse view. Then, a custom MATLAB (Version R2011a) program was developed to export the coronal image of the spinal phantom. Data sets that required two scans were combined together using the common points to both scans identified using the coded position of the probe before running the program. The data including the laminae information was selected according to the description in the previous study in Chapter 3. To export the image, the 3D ultrasound data was first projected onto the sagittal plane (XZ-plane) shown in Figure 5.5 using the MIP method, which was described in Chapter 3.1.3. Two lines were then drawn along the X-direction on the sagittal image. The range between the lines was determined based on the projected positions of the laminae, which were at deeper positions than the spinous processes and transverse processes along the Z direction. Finally, the coronal (XY-plane) image was formed from the 3D data set among the two lines placed on the sagittal image created using the MIP method.

A total of 15 US images were exported and renamed randomly (www.random.org) to reduce memory bias. The location of the T12 vertebra was labeled as the reference vertebra on the US images by three observers together prior to the measurements.

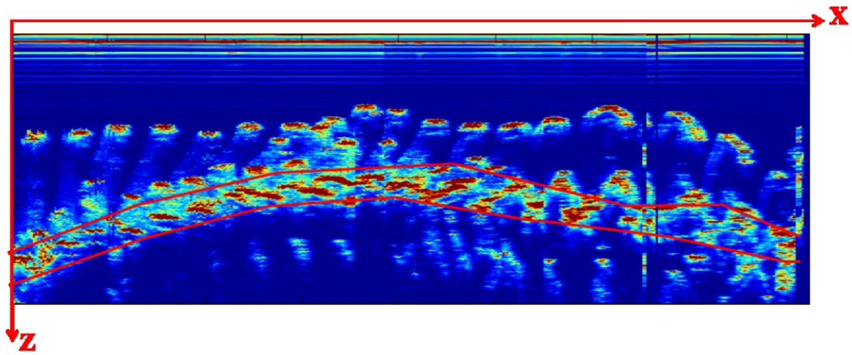


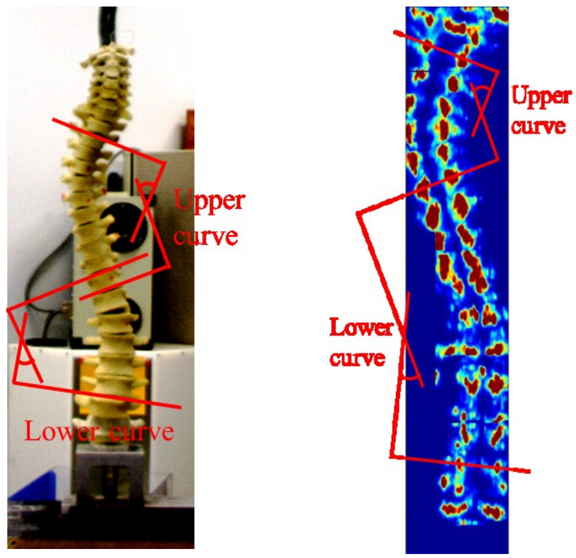
Figure 5.5 The sagittal view of the ultrasound image.

5.1.6 Observers

Three observers measured the coronal curvatures from both the LS and US images twice with one week interval to reduce the recall bias. Observer 1 (O1) had experience reviewing the US images for 2 years and Cobb method measurements for 1.5 years; observer 2 (O2) had experience on examining US images and Cobb method measurements for 1.5 years; observer 3 (O3) had no experience on neither LS nor US images. Prior to the study, O3 was trained by an expert, who had 15 years' experience on measuring Cobb angle on radiographs and practiced measurements on 10 radiographs. O3 was also trained to become familiar with the US images by O1 and practiced measurements on 5 US images.

5.1.7 Measurement methods

The LS and US images were measured at two different sessions with three days apart to blind the two assessments. ImageJ software was used to measure the curvature angles. The Cobb method was used on the LS images and the center of lamina (COL) method was applied on the US images. For the Cobb method measurements, the end vertebrae of each curve were determined and recorded first. Two lines were then drawn along the superior endplate of the upper end vertebra and inferior endplate of the lower end vertebra. The angle between these lines formed the Cobb angle (Figure 5.6a). For the COL method, the most tilted end vertebrae along a curve were recognized. Two lines were then drawn through the centers of the laminae at the end vertebrae. The angle formed by these two lines was defined as the angle measured by the COL method (Figure 5.6b).



(a) (b)

Figure 5.6 Measurement methods: (a) Cobb method, (b) Center of lamina (COL) method.

5.1.8 Statistical analysis

The intra- and inter-observer measurement differences of each method were evaluated by using the mean absolute difference (MAD) and the standard deviations (SD). The intraclass correlation coefficient (ICC) (2-way random and absolute agreement) was used to analyze the reliability between the two sessions of each observer and compare the measurements from different observers on each method (McGraw and Wong 1996; Shrout and Fleiss 1979). The ninety-five percent confidence interval (CI) was also calculated. The Currier criteria for ICC values were adopted (Currier 1990). Also, the end vertebra selections were analyzed using the error index (EI). Lower value of EI indicated less variation. For example a value of zero meant there was no difference on the end vertebra selections (Morrissy *et al.*, 1990; Oda *et al.*, 1982). The equation of the error index was

$$Error\,index = \frac{\sum_{i=1}^{n}\sqrt{(U_1 - U_2)^2 + (L_1 - L_2)^2}}{n} \tag{5-1}$$

where the U_1, U_2, L_1, and L_2 are the first and second choices of the upper and lower end vertebrae respectively, and n indicates the number of the curves.

The Bland-Altman plot (Bland and Altman, 1986; 1999; 2010) was used to indicate the differences of the intra-observer comparison of the COL method from O1, and the inter-observer differences of the Cobb and COL methods between O1 and O2. Also, the relationship between the differences and the magnitude of the measurements was investigated by Pearson's correlation coefficient. The significant level was set at $\alpha = 0.05$, that the differences was significantly correlated with magnitude when the P-value was smaller than α.

5.1.9 Motion effect

The curve configuration might change during movement of the spinal phantom between the LS and US procedures. The average of the Cobb measurements from LS1 and LS2 was used to reduce potential error introduced due to any change of curve configuration to compare with the COL measurements. The MAD±SD between the Cobb values from LS1 and LS2 among all 3 observers were 2.8°±2.4°, 3.6°±2.8°, and 2.8°±2.6° respectively.

5.1.10 Intra-observer reliability

The intra-observer reliability of both the Cobb and COL measurements of the 3 observers is shown in Table 5.1. The ICC values from both methods among all observers were higher than 0.88, which indicated that the intra-observer reliability was very high. Furthermore, the (minimum and maximum) MAD±SD values of the Cobb and COL between two sessions among all observers were (1.2°±1.1° and 2.3°±1.2°), and (2.0°±1.6° and 3.9°±2.9°), respectively. There was no difference for the intra-observer comparison of the Cobb and COL measurements and all the MAD values were less than 4°, which was within the normal acceptable error of the Cobb measurement from radiograph (5°). Figure 5.7 shows the Bland-Altman plot indicating the differences between the two sessions' measurements of the COL method from O1. Ninety-five percent of the differences fell in the range of -1.0°±4.7° (mean±2SD). There was no significant correlation between the differences and the curve severity (*p*=0.14).

Table 5.1 The intra-observer reliability of Cobb on LS and COL on US images of *in-vitro* data

Observers	Mean & SD (°)		ICC (95% CI)		MAD & SD (°)	
	Cobb (LS)	COL (US)	Cobb (LS)	COL (US)	Cobb (LS)	COL (US)
O1	33.5±10.1	35.8±9.4	0.98 (0.79-0.99)	0.97 (0.92-0.98)	1.7±1.3	2.0±1.6
O2	36.0±11.3	39.7±10.8	0.99 (0.98-0.99)	0.98 (0.94-0.99)	1.2±1.1	2.0±1.4
O3	35.1±11.3	38.5±9.4	0.98 (0.93-0.99)	0.88 (0.76-0.94)	2.3±1.2	3.9±2.9

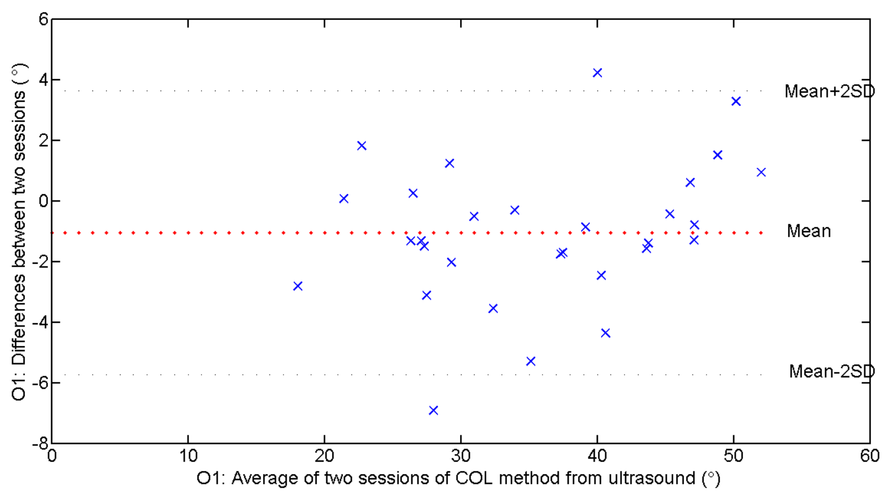


Figure 5.7 Bland-Altman plot of the intra-observer comparison of COL method on *in-vitro* US images from O1.

Table 5.2 shows the intra-observer end vertebra differences of the Cobb and COL methods. A total of 60 curvatures were measured by the Cobb method (LS1 and LS2) and 30 curvatures by the COL method. The upper end vertebra (UEV) and the lower end vertebra (LEV) were investigated separately. From both methods, the numbers of end vertebrae selection without any differences were increased and the error indices were increased from O1 to O3. O1 was the most reliable examiner to identify the same end vertebra along the curve for both methods. O2 was better than the O3 in the COL method.

Table 5.2 The intra-observer end vertebra differences of the Cobb method using LS system and the COL method using US system

		0		±1		±2 or more		Error Index
		UEV	LEV	UEV	LEV	UEV	LEV	
Cobb method	O1	44 (73.3%)	44 (73.3%)	13 (21.7%)	13 (21.7%)	3 (4.0%)	3 (4.0%)	0.68
	O2	37 (61.7%)	32 (53.3%)	15 (25.0%)	20 (33.3%)	8 (13.3%)	8 (13.3%)	1.06
	O3	32 (53.3%)	30 (50%)	21 (35.0%)	22 (36.7%)	7 (11.7%)	8 (13.3%)	1.10
COL method	O1	24 (80.0%)	25 (83.3%)	2 (3.3%)	1 (1.7%)	4 (13.3%)	4 (13.3%)	0.57
	O2	23 (76.7%)	20 (66.7%)	4 (13.3%)	6 (20.0%)	3 (10.0%)	4 (13.3%)	0.86
	O3	16 (53.3%)	12 (40.0%)	11 (36.7%)	13 (43.3%)	3 (10.0%)	5 (16.7%)	1.16

5.1.11 Inter-observer reliability

To analyze the inter-observer reliability, the average values of the measurements from two sessions of each method were used to exclude the intra-observer variations and the results are shown in Table 5.3. The ICC values of the inter-observer reliability were lower for the COL method (0.87-0.89) than the Cobb method (0.94-0.99), which was only significant (nonoverlapping 95% CI) in the comparison of O2 vs. O3. The overall values were still greater than 0.87 for the COL method. This meant the inter-observer reliabilities in all comparison pairs for both methods were very high. The MAD±SD values among all four comparisons showed the Cobb method had less variation than the COL method between observers. Furthermore, the EI results of the inter-observer comparisons were calculated from the end vertebra differences as described in the intra-observer comparison. The EI values varied in the range of 0.91-1.07 for the Cobb method and 0.94-1.26 for the COL method. Figure 5.8 and Figure 5.9 show the Bland-Altman plot of the Cobb and COL methods inter-observer comparison between O1 and O2. Ninety-five percent of the differences between the two observers were within the range of -2.5°±2.6° and -3.9°±6.0° (mean±2SD) for Cobb and COL methods, respectively. There was significant correlation between the differences and the curvature magnitude for the Cobb method (p=0.02) and COL method (p=0.03) that large curvature might related to big differences between O1 and O2.

Table 5.3 The inter-observer reliability of Cobb on LS and COL on US images on *in-vitro* data

	Mean & SD (°)		ICC (95% CI)		MAD & SD (°)		Error Index	
	Cobb (LS)	COL (US)	Cobb (LS)	COL (US)	Cobb (LS)	COL (US)	Cobb (LS)	COL (US)
O1 vs. O2	34.7±10.6	37.8±10.0	0.94 0.68-0.98	0.88 0.29-0.96	2.9±2.4	4.1±3.3	0.91	1.14
O2 vs. O3	35.5±11.2	39.1±9.9	0.98 0.96-0.99	0.89 0.78-0.94	1.6±1.1	3.7±3.2	1.07	1.26
O1 vs. O3	34.3±10.6	37.2±9.2	0.96 0.88-0.99	0.87 0.64-0.94	2.1±2.0	3.9±3.1	0.96	0.94
O1 vs. O2 vs. O3	34.8±10.8	38.0±9.9	0.96 0.90-0.98	0.88 0.74-0.94	2.2±2.0	3.9±3.2		

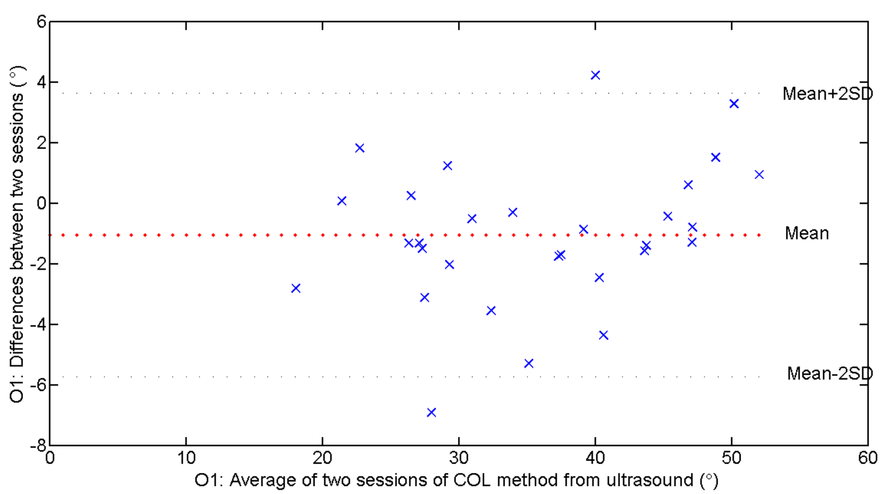


Figure 5.8 Bland-Altman plot of inter-observer comparison of the Cobb method on *in-vitro* LS images between O1 and O2.

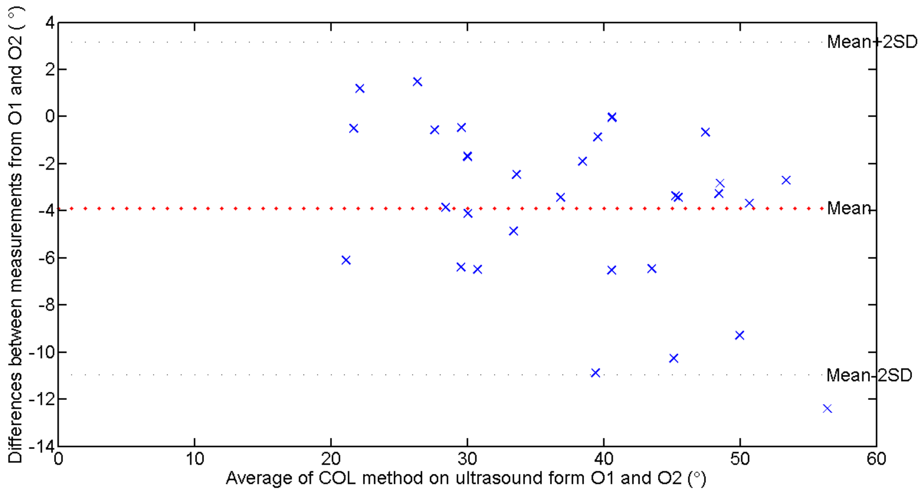


Figure 5.9 Bland-Altman plot of inter-observer comparison of the COL method on *in-vitro* US images between O1 and O2.

5.1.12 Discussions

The ICC values were commonly applied to analyze the reliability of the Cobb method. From the literature, the ranges of the ICC values of the Cobb measurements were 0.87-0.99 for the intra-observer reliability and 0.87-0.98 for the inter-observer reliability (Gstoettner *et al.*, 2007; Kuklo *et al.*, 2005; Mok *et al.*, 2008). In this study, both the Cobb method measured from the anterior side of the spine and the COL method measured from the posterior side of the spine were measured from the cadaver spinal phantom. The intra- and inter- reliabilities of the COL method were reported as high as the Cobb measurements from the literature. The MAD values of both the intra- and inter-observer comparisons were in a small range (<4°), which were within the clinic acceptable range (5°). The intra-observer reliabilities of the COL method had similar results to the Cobb method in all three observers. Although the ICC values of the inter-observer comparisons of the COL method were lower than these of the Cobb method, the ICC values were still higher than 0.87 indicating high inter-observer reliabilities for the COL method.

To evaluate the performance of the Cobb angle measurements, the end vertebrae selection was considered as one of the major error source (Mok *et al.*, 2008; Morrissy *et al.*, 1990; Oda *et al.*, 1984). The higher the EI value was, the more variations the end vertebra selection had. In this study, both the intra- and inter-observer EIs of the COL method were comparable with the literature and the EI values of the Cobb method. Selecting the end vertebrae on ultrasound images is similar to selecting end vertebrae from radiographs (Morrissy *et al.*, 1990; Oda *et al.*, 1982). Furthermore, the EI and MAD values of the COL method measurements from O1 were lower than those from O2 and O3, which showed that US experience might be related to the measurements of COL method.

5.2 *In-vivo* study

To further validate the proposed COL method, an *in-vivo* study was conducted by measuring the coronal curvatures on ultrasound images obtained from AIS patients. Intra- and inter-observer reliabilities were analyzed and the COL measurements were compared with the Cobb angles that were recorded during scoliosis clinics.

5.2.1 Study population

The inclusion criteria for the *in-vivo* study were 1) diagnosed as AIS, 2) the major Cobb angle was under 40° so that only mild and moderate curves were considered, 3) an out of brace radiographs was to be taken and no in brace radiograph was planned during the study day which would otherwise correct the spinal deformity, and 4) had no spine surgery prior to the study. Eighty consecutive volunteers were found eligible for US scanning and tested. However, the data sets from only 26 subjects (25F, 1M, age 14.5±1.8 years, Cobb angle 22°±7.4°) could be used in the present study indicating a low success rate (32.5%) for data acquisition.

5.2.2 Methods

The female subjects wore a gown with the back opened to the operator during the scanning process. Each subject sat straight inside the frame and looked forward with their arms crossed resting on a metal bar in front of their chest to avoid the shoulder blades sticking out (Figure 5.10). C7 was identified and labeled with a red sticker as the beginning point. A soft tissue biopsy synthetic material (Blue Phantom^{TM}, USA) was put on the surface of the transducer with the dimensions of 115 mm length, 35 mm width, and 10 mm thickness (Figure 5.11). This material is flexible and can be used to optimize the contact by filling the gap between the transducer and subjects' back when a patient has significant trunk rotation. Ultrasound gel was applied to both sides of the soft tissue biopsy to ensure good contact between the transducer and the back surface of the subjects. Similar to the *in-vitro* experiment, one operator, who was trained during the *in-vitro* experiments on spinal column phantom and *in-vivo* experiments on healthy volunteers, performed the scan from the C7 to L4, while the other operator controlled the laptop for data collection. It took approximately 30 seconds to perform a vertical scan along the vertical bar on one subject.

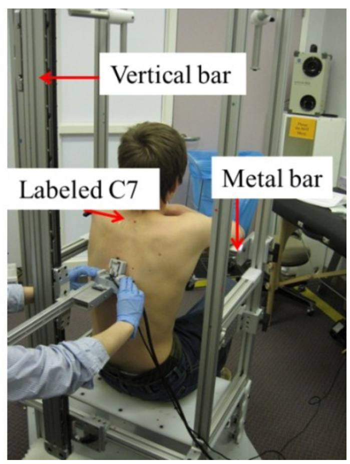


Figure 5.10 Scanning for a subject's back.

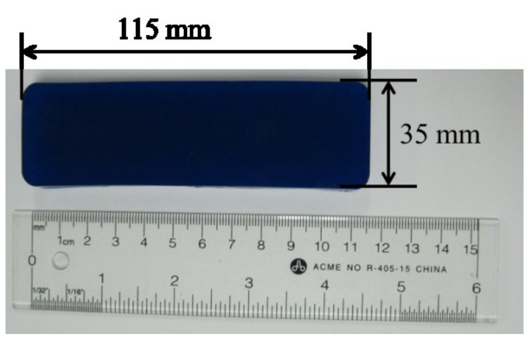


Figure 5.11 Soft tissue biopsy material. (Blue Phantom^{TM}, USA)

5.2.3 Observers

Two observers performed the coronal curvature measurements from the US images twice in two sessions. Observer 1 (O1) had experience reviewing the US images for 3 years; observer 2 (O2) had no experience on examining US images. Prior to the study, O2 was trained by O1 using the 15 US phantom images which were described in Chapter 5.1.

5.2.4 *In-vivo* ultrasound images

The projected coronal ultrasound images of the subjects were exported in bitmap image format (BMP) through Tomoview software. One of the subject images is shown in Figure 5.12. Some of the subjects had more than one curvature, thus a total of 31 curvatures were defined by O1. For each curvature, the end vertebrae were labeled before measurement by O1 to eliminate errors introduced by the end vertebrae selection. To determine the end vertebrae, the apex of each curve was first identified, and then the vertebrae with the most tilted pair of laminae were recognized as the end vertebrae. US images were assigned with random numbers using the free service from the web site (www.random.org) one week before starting the measurements. The observers were blinded to the Cobb angle from scoliosis clinical record during the measurement.

5.2.5 Measurement methods

The observers measured the US images twice with at least one week interval to reduce the recall bias. Figure 5.12 shows the *in-vivo* ultrasound image in which the spinous process was the white area in the middle. In the thoracic region, the laminae could be recognized as the red ellipse areas, but in the lumbar area, it was difficult to distinguish the laminae from the noise. Fortunately, the sagittal and transverse view images were able to help the recognition of laminae. The ImageJ software (v1.45, NIH) was used to draw two lines through the center of lamina (COL) of the end vertebrae. The angle formed between these two lines indicated the coronal curvature.

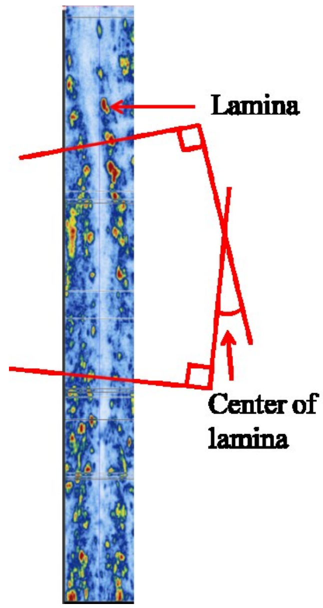


Figure 5.12 Center of lamina (COL) method measured on ultrasound image.

5.2.6 Statistical analysis

The intraclass correlation coefficient (ICC) (2-way random and absolute agreement) (Shrout and Fleiss, 1979; McGraw and Wong, 1996) and the mean absolute difference (MAD) with standard deviations (SD) were used to analyze the intra- and inter-observer reliabilities, and compare the coronal curvature measurements between the COL method and the Cobb angle from the scoliosis clinical records. The same Currier criteria for ICC values were adopted (Currier 1990). The Bland-Altman plot (Bland and Altman, 1986; 1999; 2010) was used to show the intra-, and inter-observer comparison of the coronal curvature measurements in using the COL method from both observers, and the differences between the COL method and the Cobb method. The relationship between the differences and the magnitude of the curvatures was also investigated by using the Pearson's correlation coefficient and reported as significant if $p<0.05$.

5.2.7 Intra-observer reliability

The intra-observer reliability of the COL method of the two observers is shown in Table 5.4. The ICC values for both observers were higher than 0.94 which indicated high intra-observer reliability for both observers. The MAD±SD values between the two sessions were 1.3°±1.1° and 1.9°±1.5° for O1 and O2, respectively. The means of the coronal curvature measurements from the COL method were close to the mean of the Cobb angle from the scoliosis clinical records (22°±7.4°). Figure 5.13 and Figure 5.14 show the Bland-Altman plot indicating the differences between the two sessions' measurements of the COL method from O1 and O2, respectively. Ninety-five percent of the differences fell in the range of -0.3°±3.4° and -0.2°±4.8° (mean±2SD) for O1 and O2, respectively. There was no significant correlation between the differences and the curve severity for the COL intra-observer comparisons for O1 (p=0.40) and O2 (p=0.54), respectively.

Table 5.4 The intra-observer reliability of the COL method on *in-vivo* US images

	O1	O2
ICC (95% CI)	0.97 (0.94-0.98)	0.94 (0.88-0.97)
Mean±SD (°)	22.0±6.8	21.4±6.8
MAD±SD (°)	1.3±1.1	1.9±1.5

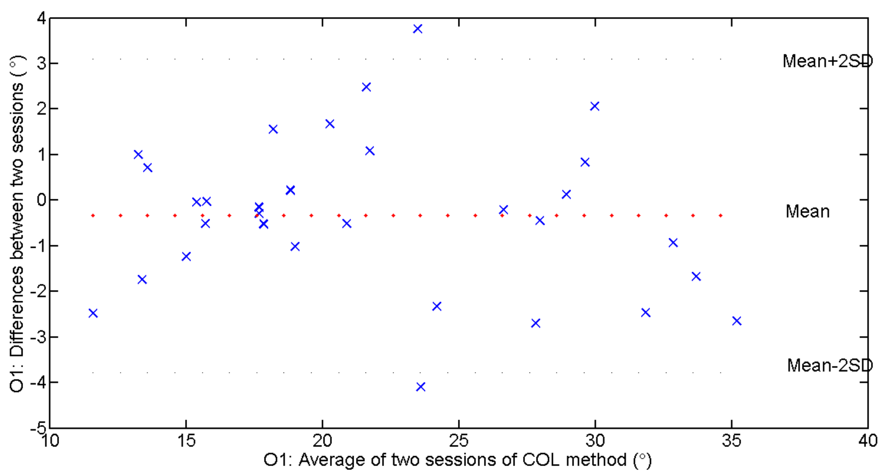


Figure 5.13 Bland-Altman plot of the intra-observer comparison of the COL method on *in-vivo* US images from O1.

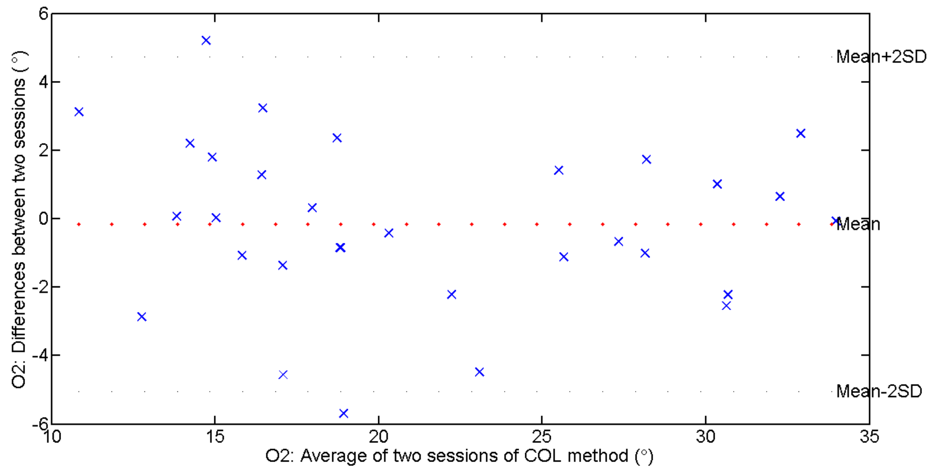


Figure 5.14 Bland-Altman plot of the intra-observer comparison of the COL method on *in-vivo* US images from O2.

5.2.8 Inter-observer reliability

To analyze the inter-observer reliability, the average values of the measurements from two sessions of each observer were used and the result is shown in Table 5.5. The ICC value (0.93) indicated high inter-observer reliability between the two observers for the COL method. The MAD±SD was 2.1°±1.3°, which showed small variation for the inter-observer comparison. Figure 5.15 shows the Bland-Altman plot of the comparison between O1 and O2. Ninety-five percent of the differences between the measurements from two observers were within the range of -0.8°±4.8° (mean±2SD). The differences were independent of the curve magnitude (p=0.94).

Table 5.5 The inter-observer reliability of the COL method on *in-vivo* US images

	O1 vs. O2
ICC (95% CI)	0.93 (0.87-0.97)
MAD±SD (°)	2.1±1.3

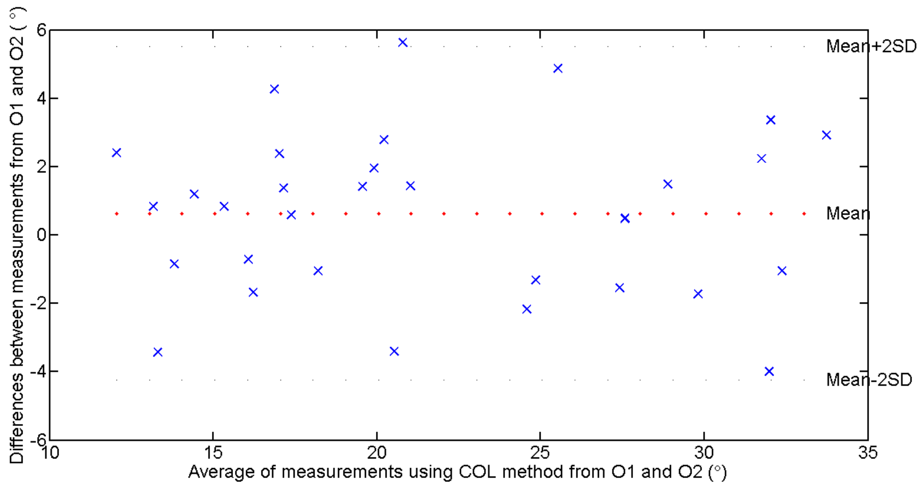


Figure 5.15 Bland-Altman plot of the inter-observer comparison of the COL method on *in-vivo* US images between O1 and O2.

5.2.9 Validity analysis

The agreement comparison results between the measurements of the COL method and the Cobb angle from local scoliosis clinical record are shown in Table 5.6. The COL method measurements agreed well with the Cobb method for both observers (ICCs>0.90). Also, the inter-method MAD±SD were 1.7°±1.2° and 2.9°±2.5° for O1 vs. clinical record and O2 vs. clinical record, respectively. Figure 5.16 and Figure 5.17 show the Bland-Altman plot of the differences between the measurements of COL and Cobb methods from O1 and O2, respectively. Ninety-five percent of the differences fell in the range of -0.1°±4.2° and -0.6°±5.7° for O1 and O2, respectively. The differences were not significantly related to the average of the curve severity of the two methods for O1 (p=0.12) and O2 (p=0.27), respectively.

Table 5.6 The agreement between measurements of the COL method on *in-vivo* US images and clinical record of Cobb method

	O1	O2
ICC (95% CI)	0.96 (0.91-0.98)	0.92 (0.84-0.96)
MAD±SD (°)	1.7±1.2	2.9±2.5

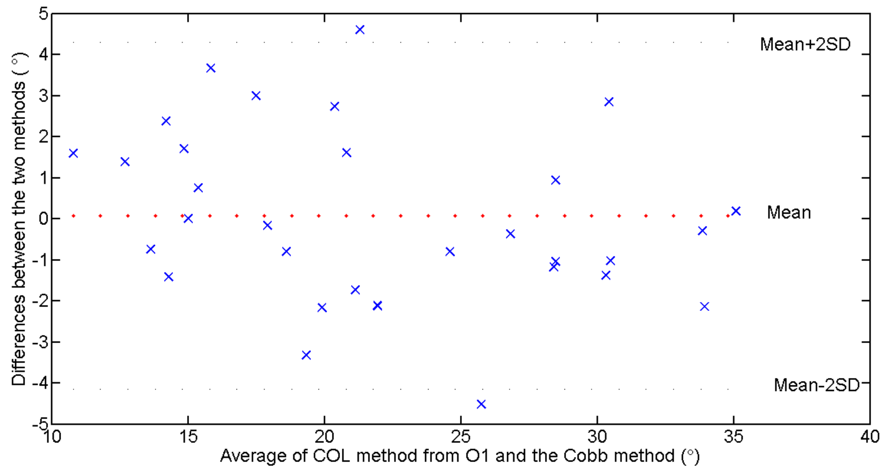


Figure 5.16 Bland-Altman plot of the comparison between the COL method on *in-vivo* US images from O1 and clinical record of Cobb method.

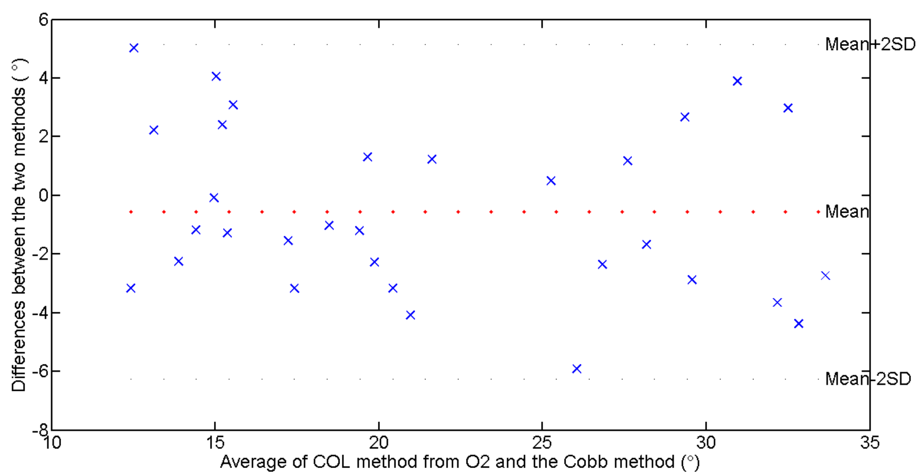


Figure 5.17 Bland-Altman plot of the comparison between the COL method on *in-vivo* US images from O2 and clinical record of Cobb method.

5.2.10 Discussions

The *in-vitro* study described in this chapter showed that the errors of the end vertebrae selections from the ultrasound images were similar to the errors from radiographs. The observer with more US experience selected the end vertebrae with less variation. The ultrasound images from the *in-vivo* study had more noise than the *in-vitro* study. Laminae were more difficult to be identified especially for an observer with no US experience. Also the Cobb angles from the clinical record were measured by the specialist, who had many years' experience on measuring Cobb angle. As the *in-vivo* study was to focus on the measurements variation and comparison analysis of the COL measurements vs. the clinical records value, the end vertebrae were pre-selected in this study by the experienced observer. Because the error source from the end-vertebrae selection was eliminated from the *in-vivo* study, the COL measurements on the *in-vivo* images, as expected, did show better reliability than that of the in-vitro study.

The ICC values indicated that the COL method measured from the AIS subjects had high intra- and inter-observer reliabilities. Compared with the ICC values of the Cobb measurements (0.87-0.99) from the literature (Gstoettner *et al.*, 2007; Kuklo *et al.*, 2005; Mok *et al.*, 2008), the COL method showed a comparable or better result (ICC>0.93). Furthermore, the MAD values were in a small range (<2.2°) for the reliability analysis which was within the clinical tolerance (5°). The comparison between the measurements of the COL method and the scoliosis clinical record of the Cobb method also showed great agreement with high ICC values and low MAD values for both observers. O1 had better performance than O2, which indicated that US experience may affect the measurement reliability of the COL method.

Although the *in-vivo* study further validated the proposed COL method, the US imaging method was limited to AIS patients who have mild to moderate deformity. The sitting posture was used in the *in-vivo* study, which was different from the standing posture of taking radiographs. The subjects were required to keep straight with the arms elevated during scanning to try to have a similar sagittal profile as in the radiograph and the frame was designed to help the subject hold the sitting posture. The supine position was not used because it was reported to correct the spinal deformity (Torell *et al.*, 1985; Yazici *et al.*, 2001) because of the different forces on the vertebrae. Although the pelvis position in the sitting posture might be different from that in the standing posture, the gravity load and the intra-discal pressure were reported to be similar for the standing and sitting with back straightened (Douglas *et al.*, 2012). Therefore, the lateral deformity of the spine would not significantly change and the validity analysis indicated that there were only small variations between the measurements using US method and radiograph method. During the recruitment, a low success rate (32.5%) was obtained for data acquisition. There were two major reasons. One was the encoder contact issue. The encoder had limited freedom to come in and out along the Z direction and the back of the subjects was not an even surface. When the patients had big kyphosis or lordosis or hump, the wheel of the encoder would not be in the same plane as the surface of the transducer. In such cases, sometimes the wheel of the encoder lost contact with the back and no US signals were transmitted. Therefore, US signals could not be acquired from the surface of the vertebra. Another reason is that one vertical scan may not cover the full spine area for some subjects with double or triple curves if laminae are deviated laterally beyond the area included in the US scan. These data sets were excluded from the study. After excluding these ineligible data sets, only 26 data sets were involved in the *in-vivo* study. For the second problem causing the low success rate, two or three scans may allow covering the full spine in the future. However, multiple scans could also create a significant challenge to merge the data sets since the 2D encoder could not track the 3D coordinate of the AIS patients' back, which may further increase the inaccuracy. To improve the success rate, a 3D GPS built-in transducer should be used so that the transducer position could be tracked during the whole scanning process without using an encoder. Also, the transducer could move along the path of the deviated the spine so that both sides of the laminae could be included in the imaging range of the transducer from the top to the bottom of the spine which could address the second problem.

5.3 Summary

The *in-vitro* and *in-vivo* results indicated that the COL method had high intra- and inter-observer reliabilities, which was compatible with the Cobb method. The measurement error of the COL was within the acceptable accuracy for the scoliosis clinic. Therefore, the proposed COL method is a promising tool to evaluate the mild to moderate coronal curvature of AIS. However, the success rate for the *in-vivo* data acquisition was low and solutions were proposed which need to be tested in the future. Also, an operator who has ultrasound experience may affect the reliability of the measurements from the COL method. After the coronal curvature investigation, the reliability of vertebral axial rotation measurements was also required to be investigated.

Chapter 6.

Reliability study of the vertebral axial rotation evaluations on ultrasound images

Besides the coronal curvature, vertebral axial rotation (VAR) is another important parameter to assess the severity of scoliosis. In Chapter 3, the transverse ultrasound (US) images of vertebrae could be acquired and laminae were identified which could be used for the VAR measurements. The present chapter reports *in-vitro* and *in-vivo* studies to evaluate the intra- and inter-observer reliabilities and the validity of the VAR measurements using the center of lamina (COL) method on US images. Also, the VAR measurements using the COL method were compared with the Stokes' method on radiographs in the *in-vitro* study.

6.1 *In-vitro* study

6.1.1 Vertebra phantoms

Three cadaver vertebrae T7, L1, and L3 were mounted onto custom rotation blocks to allow free turning for different axial rotation setup during experiments. Figure 6.1 shows the T7 vertebra mounted on top of a plastic bar linked with a sharp pointer, which is secured to a plastic platform. A transparent protractor was printed and attached on the surface of the platform for angle indication. The rotation angle of the vertebra was measured using the pointer with respect to the protractor. Each vertebra was turned from -30° to 30° with 5° increments. To investigate the limitation of the ultrasound measurement on the VAR, three extra rotations 40°, 50°, and 60° were also tested on T7 since severe VAR may occur (Easwar *et. al*, 2011). Including the three extra rotations, a total of 42 configurations were set for the three vertebrae.

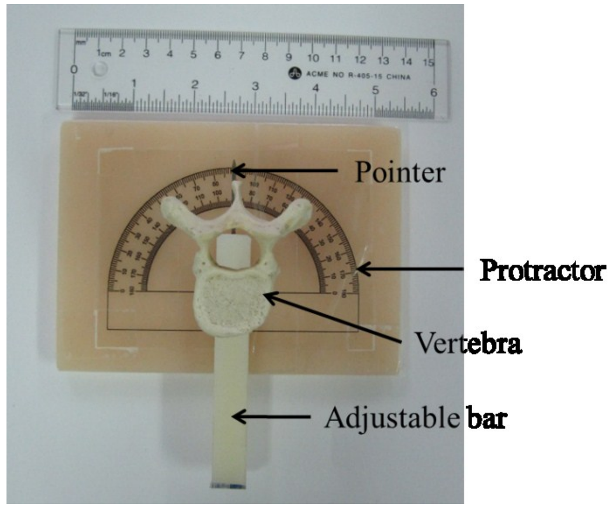


Figure 6.1 Vertebra phantom: T7.

6.1.2 Experimental setup

A custom frame described in Chapter 2.4.2 was used to control the transducer and the encoder for data collection. Each vertebra phantom was secured at the bottom of a container filled with water (Figure 6.2). Ultrasound gel was put on the surface of the container to ensure good coupling. For each rotation setting, one operator controlled the software to record US data, and the other operator scanned the vertebra from the top to the bottom along the surface of the container with a scanning speed less than 36 mm/s, which was introduced in Chapter 5.1.4, in order to obtain US data of good quality.

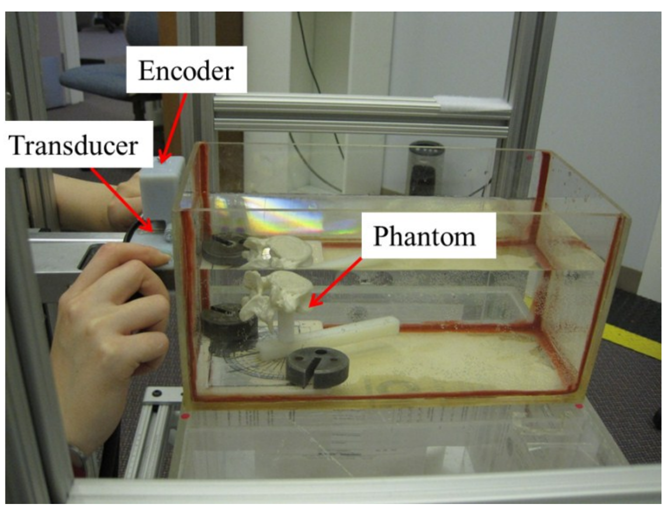


Figure 6.2 Experimental setup for the vertebra phantom.

6.1.3 Ultrasound images of vertebrae

As described in Chapter 3.1.4, the transverse image could be exported at the position where the 'W' shape of the vertebra was recognized. The ultrasound images of each configuration on the transverse plane were then exported in BMP format. Figure 6.3 shows an example of one transverse image for T7 vertebra phantom with rotation of 15°. A total of 42 US images were prepared for the measurements.

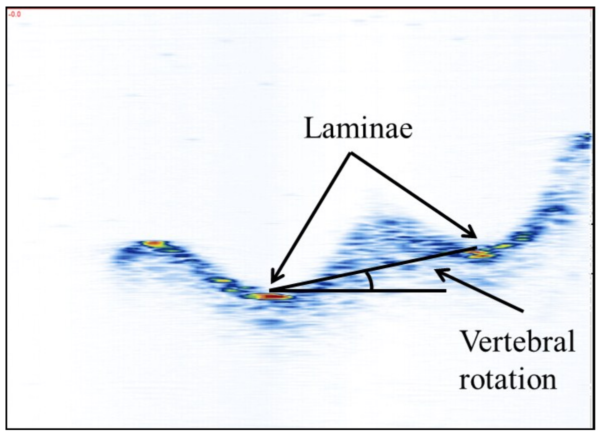


Figure 6.3 Transverse US image from the T7 phantom with rotation of 15°.

6.1.4 Radiographs of vertebrae

The posteroanterior (PA) radiographs were taken by the radiographic technician for the three cadaver vertebrae with the same 42 configurations. Different configurations were manipulated on the phantom without moving the phantom. One PA image of the T7 vertebra phantom with rotation of 0° is shown in Figure 6.4. After reviewing the 42 radiographs, the pedicles could not be identified on two images with rotations of 50° and 60°. Therefore, a total of 40 radiographs were applied for the measurements.

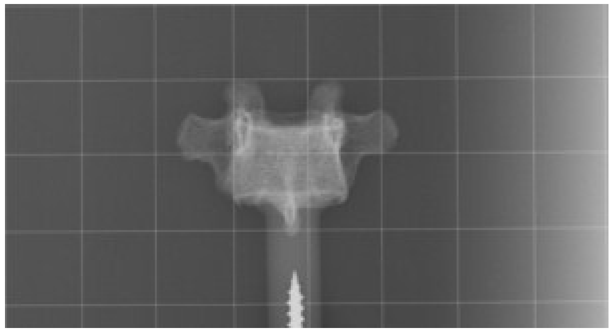


Figure 6.4 PA radiograph from the T7 phantom with rotation of 0°.

6.1.5 Arrangements of images

Each observer performed the measurements twice with one week interval to remove the memory bias. US images and the radiographs of the three cadaver vertebra phantoms were assigned random numbers using the free service from the web site (www.random.org) one week before performing the measurements.

6.1.6 Observers

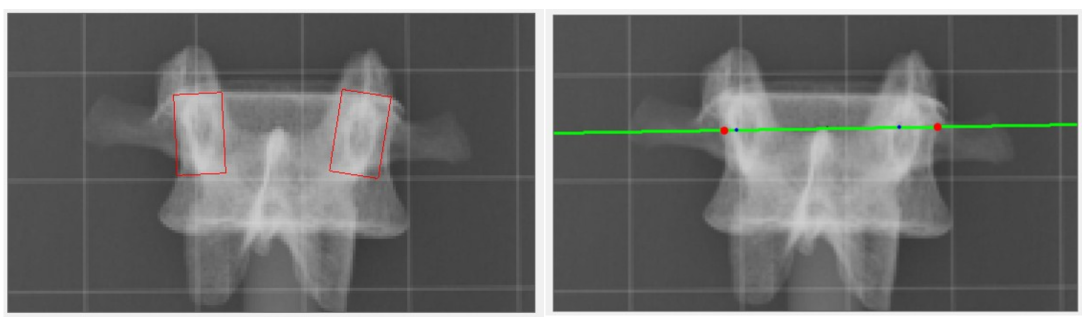
Three observers measured the phantom rotations on the US images. Observer 1 (O1) had 2-year experience on reviewing the US images, observer 2 (O2) had 1.5-year experience on examining US images, and observer 3 (O3) had no experience on US images measurements. On the other hand, the radiographs were measured by two observers, the same O1 and an observer 4 (O4). Both O1 and O4 had no experience on measuring the rotations from PA radiographs.

6.1.7 Ultrasound measurement method

Laminae were used as bony landmarks to assess the VAR. The strong reflection signals displayed as the red areas around the lowest points of the 'W' shape were the laminae. A line was then drawn through the centers of the left and right laminae. The angle formed between the line and the reference horizontal level indicated the VAR angle (Figure 6.3). ImageJ software (v1.45, NIH) was used to measure the rotations. The reference horizontal was parallel to the surface of the water tank.

6.1.8 Radiography measurement method

A custom program developed by Zhang (Zhang et al., 2010) was used to measure the vertebral axial rotation on radiographs. The vertebra name was needed to determine the width-to-depth ratio from the report of Stokes (1986). Then, the operator put the rectangular boxes to cover the pedicle areas of the vertebra (Figure 6.5a). After pedicle selections, the program used the gradient vector flow (GVF) snake model to determine the centers of pedicles. A line was automatically drawn through the two center points. The user marked the intersection points formed by the line and vertebra edges manually (red dots in Figure 6.5b). Based on the four points, the rotation angle was calculated using the Stokes' method, calculated by the pedicle offset from the vertebral body center and width-depth ratio estimation (Stokes *et al.*, 1986).



(a) (b)

Figure 6.5 Vertebral axial rotation measurements on radiograph using Stokes' method. (a) Select the pedicle areas; (b) pick the intersection points by the line and the edge of vertebral body (red dots).

6.1.9 Statistical analysis

The intraclass correlation coefficient (ICC) (2-way random and absolute agreement) (Shrout and Fleiss, 1979; McGraw and Wong, 1996) and the mean absolute difference (MAD) with standard deviations (SD) were used to analyze the intra- and inter-observer reliabilities of the VAR measurements using the COL and Stokes' method. Measurements using the COL and Stokes' method were also compared with the experimental setup. The Currier criteria for ICC values were adopted (Currier 1990). The Bland-Altman plot (Bland and Altman, 1986; 1999; 2010) was used to show: 1) the intra-observer comparison of the measurements of the COL and Stokes' methods from O1, 2) the inter-observer differences of the VAR measurements from the COL method between O1 and O2, and the Stokes' method between O1 and O4, and 3) the differences between the COL method and the experimental setup, and the Stokes' method with the experimental setup from O1. The relationship between the differences and the magnitude of the VAR measurements was investigated by Pearson's correlation coefficient. The significant level was set at $\alpha = 0.05$. Differences were significantly correlated with magnitude when the P-value was smaller than α.

6.1.10 Intra-observer reliability

The intra-observer reliability results of the VAR measurements using the COL method from the three observers are shown in Table 6.1. The ICC values indicated high intra-observer reliabilities from all observers (>0.99). The MAD±SD values of the VAR measurements between two sessions were 0.4°±0.4°, 0.5°±0.5°, and 0.5°±0.6°, for O1, O2, and O3 respectively. The mean±SD from the three observers (O1: 16.6°±9.9°, O2: 16.9°±9.9°, and O3: 16.2°±9.4°) were close to the mean±SD (16.8°±10.2°) of the experimental setup. Figure 6.6 shows the Bland-Altman plot indicating the differences between the two sessions versus the average VAR measurements on the US images from O1. Ninety-five percent of the differences fell in the range of -0.1°±1.0° (mean±2SD). There was no significant relationship between the differences and the VAR rotation magnitude (p=0.54).

Table 6.1 Intra-observer reliability of the COL method for VAR on *in-vitro* US images

	O1	O2	O3
ICC (95%CI)	>0.99 (>0.99)	>0.99 (>0.99)	>0.99 (>0.99)
Mean±SD (°)	16.6±9.9	16.9±9.9	16.2±9.4
MAD±SD (º)	0.4±0.4	0.5±0.5	0.5±0.6

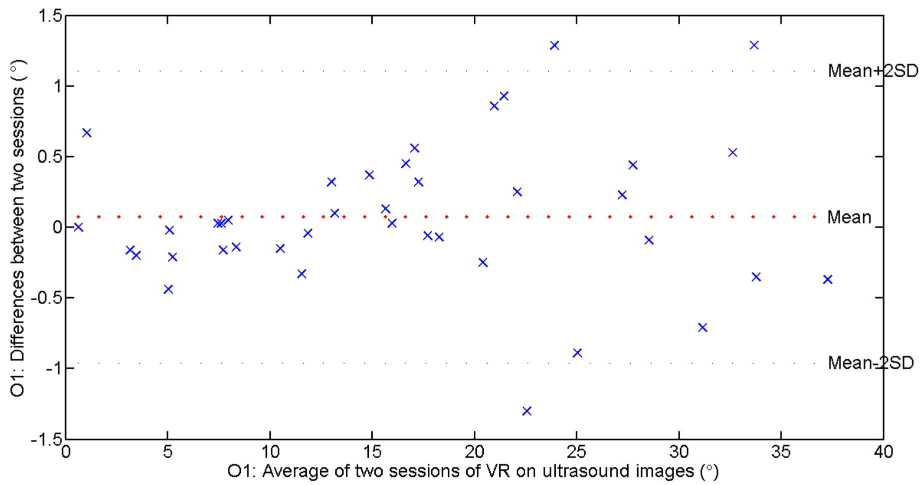


Figure 6.6 Bland-Altman plot of the intra-observer comparison of the COL method for VAR from O1 on US *in-vitro* images.

Table 6.2 shows the intra-observer reliability of the Stokes' method using Zhang's program. The ICC values indicated high reliability for both observers (>0.90). The MAD±SD were 2.3°±2.2° and 3.2°±3.8° for O1 and O4, respectively, which were higher than the MAD±SD values of the ultrasound measurements. Figure 6.7 indicates the intra-observer differences of the Stokes' method from O1. Ninety-five percent of the differences fell in the range of mean±2SD (0.3°±6.3°), which had a larger range compared with the differences shown in Figure 6.6. There was no significant correlation between the differences and the VAR magnitude (p=0.45).

Table 6.2 The intra-observer reliability of the Stokes' method for VAR on *in-vitro* radiographs

	O1	O4
ICC (95%CI)	0.96 (0.93-0.98)	0.90 (0.82-0.95)
Mean±SD (°)	14.9±11.1	15.2±11.0
MAD±SD (°)	2.3±2.2	3.2±3.8

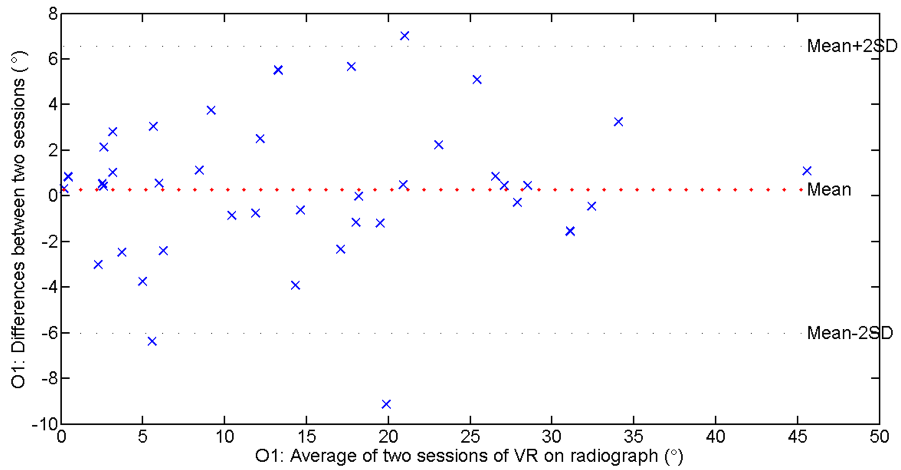


Figure 6.7 Bland-Altman plot of the intra-observer comparison of Stokes' method for VAR from O1 on *in-vitro* radiographs.

6.1.11 Inter-observer reliability

Table 6.3 summarizes the inter-observer reliability of both the COL and Stokes' methods. Average values of the two sessions' measurements were used. The ICC values of all comparisons were above 0.99 for the COL method, and 0.93 for the Stokes' method. The high ICC values indicated both methods were highly reliable. The MAD±SD values for the three pairs (O1 vs. O2, O2 vs. O3, and O1 vs. O3) of the COL method were 0.7°±0.8°, 0.8°±0.8°, and 0.7°±1.1°, respectively. The MAD±SD for comparison of Stokes' method between O1 and O4 was 3.0°±2.7°. Figure 6.8 and Figure 6.9 show the Bland-Altman plot of the inter-observer comparisons for the COL and Stokes' methods, respectively. Ninety-five percent of the differences between the O1 and O2 for the COL method and O1 and O4 for the Stokes method were within the range of -0.3°±2.0° and -0.3°±8.1° (mean±2SD), respectively. The COL method was more reliable than the Stokes' method for the inter-observer comparison. The differences were independent of the VAR magnitude for the COL method (P=0.68) and Stokes' method (*p*=0.84).

Table 6.3 The inter-observer reliability of the COL method on US images and Stokes' method on radiographs for VAR of *in-vitro* data

	COL			Stokes'
	O1 vs. O2	O2 vs. O3	O1 vs. O3	O1 vs. O4
ICC (95%CI)	>0.99 (>0.99)	>0.99 (>0.97)	>0.99 (>0.98)	0.93 (0.88-0.96)
MAD±SD (°)	0.7±0.8	0.8±0.8	0.7±1.1	3.0±2.7

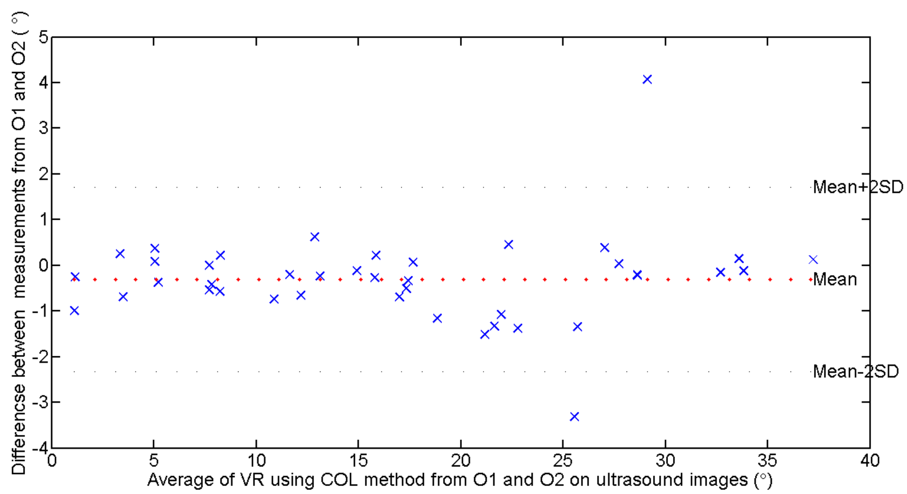


Figure 6.8 Bland-Altman plot of the inter-observer comparison of the COL method for VAR between O1 and O2 on *in-vitro* US images.

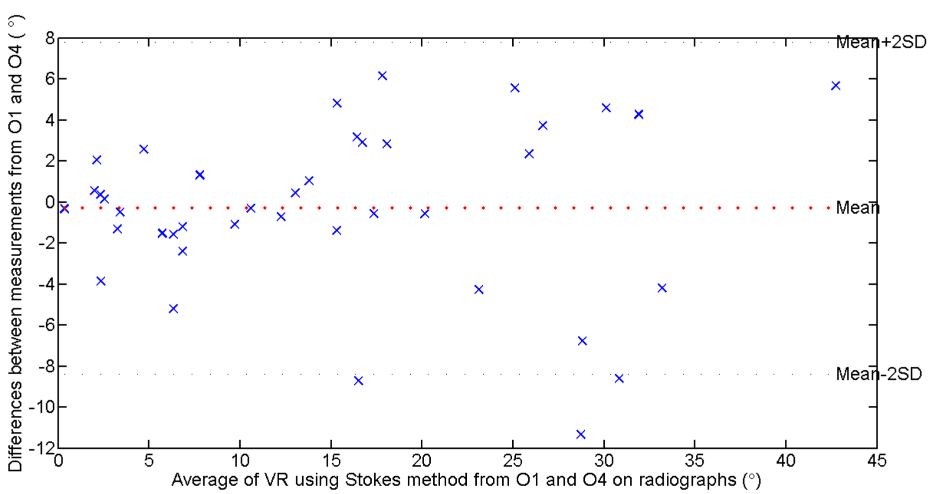


Figure 6.9 Bland-Altman plot of the inter-observer comparison of the Stokes' method for VAR between O1 and O4 on *in-vitro* radipgraphs.

6.1.12 Validity analysis

The validity analysis of the COL and Stokes' methods are shown in Table 6.4. The ICC values showed that the measurements using the COL method highly agreed with the experimental setup (ICCs>0.96), while good agreement was observed between the measurements using Stokes' method and the experimental setup (0.80<ICCs<0.90). The MAD±SD for the US validity analysis were smaller than 2.5° with the absolute difference (AD) up to 5.4°, while the MADs were above 4° for the radiograph validity analysis with the AD up to 15.3°.

Table 6.4 Validity analysis of the COL method on US images and Stokes' method on radiographs for VAR of *in-vitro* data

	COL			Stokes'	
	O1	O2	O3	O1	O4
ICC (95%CI)	0.97 (0.94-0.98)	0.97 (0.94-0.98)	0.96 (0.93-0.98)	0.89 (0.79-0.94)	0.82 (0.68-0.90)
MAD±SD (°)	2.2±1.3	2.2±1.1	2.3±1.4	4.1±3.0	4.9±4.2
AD range (°)	0.02-5.1	0.04-5	0.03-5.4	0.2-11.8	0.1-15.3

Figure 6.10 and Figure 6.11 shows the differences for the validity analysis of the COL method and Stokes' method, respectively from O1. Ninety-five percent of the differences fell in the range of -0.1°±5.2° and -1.8°±9.5° (mean±2SD), respectively for the US and radiograph validity analysis, which indicated the measurements of COL method was more accurate. The differences were independent of the VAR magnitude for the COL method (P=0.61) and Stokes' method (P=0.20).

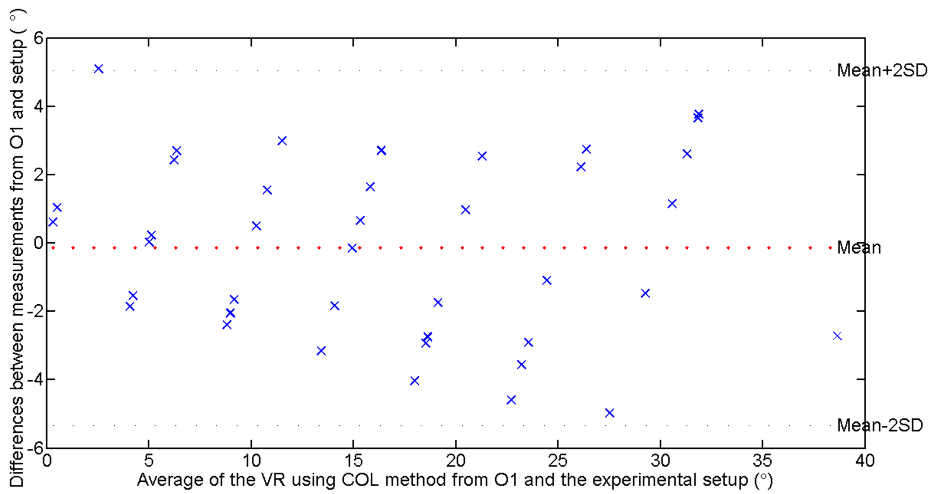


Figure 6.10 Bland-Altman plot of the validity analysis of the COL method on *in-vitro* US images for VAR from O1.

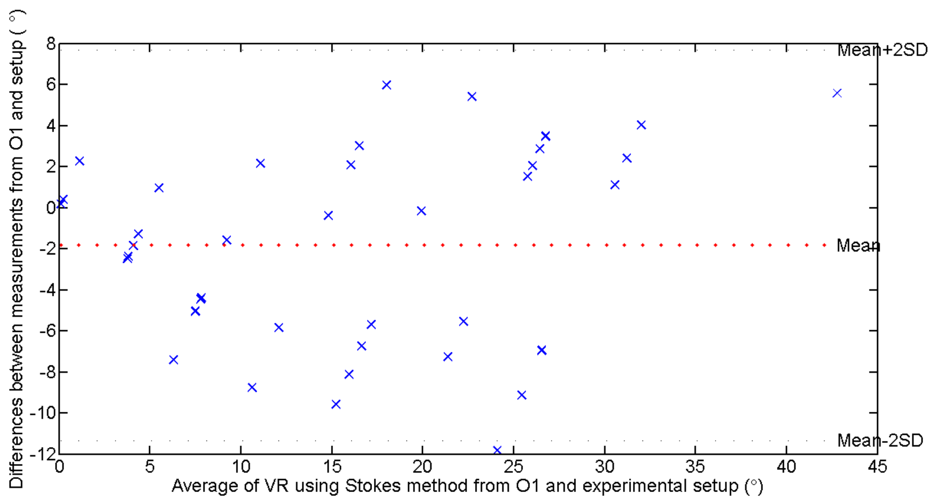
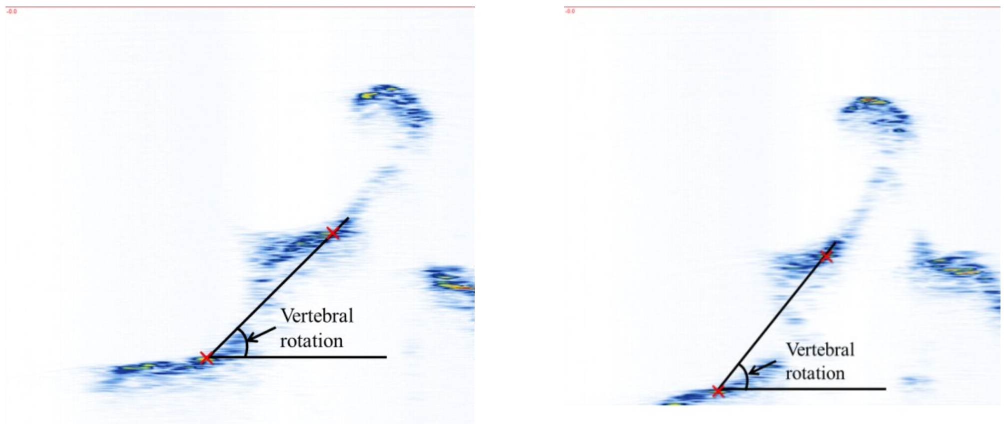


Figure 6.11 Bland-Altman plot of the validity analysis of the Stokes' method on *in-vitro* radiographs for VARfrom O1.

6.1.13 Limitation of *in-vitro* VAR measurements on ultrasound images

As mentioned in section 6.1.4, pedicles could not be recognized on the radiograph for the rotations of 50° and 60°. These two rotations were then not included in the reliability and validity analysis. Ultrasound images of these two rotations are shown in Figure 6.12. Because of the rotation, the two strong reflections moved together from the laminae, the lowest points of the 'W' shape, to the side. O1, O2, and O3 all measured the two large rotations twice with one week interval. The ADs were calculated for the intra- and inter- observer comparisons. The results are shown in Table 6.5. The small errors (<5°) indicated that it was possible to evaluate the large VAR using ultrasound.



(a) (b)

Figure 6.12 Ultrasound images of the VAR (a) of 50°, and (b) of 60°.

Table 6.5 Absolute differences for the measurements of the COL method from 50°, 60° (unit: °).

	Intra-observer comparison			Inter-observer comparison			Compare with experimental setup		
	O1	O2	O3	O1 vs. O2	O2 vs. O3	O1 vs. O3	O1	O2	O3
50°	0.1	0.2	0	2.0	2.9	1.0	3.7	1.7	4.6
60°	0.2	0.1	2.1	1.5	0.3	1.9	2.6	1.1	0.7

6.2 *In-vivo* study

The results from the *in-vitro* experiments indicated that the COL method was a reliable and accurate method to measure VAR. To verify if the proposed method is applicable to AIS subjects, an *in-vivo* study was conducted.

6.2.1 Study population

The inclusion criteria for the *in-vivo* study were 1) diagnosed as AIS, 2) the major Cobb angle was under 40°, 3) no in-brace radiographs taken during the study day, and 4) had no spine surgery prior to the study. Forty consecutive subjects from the local scoliosis clinic were eligible for the US scanning and signed the written consent before participating in this study. Only thirteen data sets (13F, age 13.7±1.8 years; Cobb angle 22°±7.7°) could be used for the study due to encoder or probe contact issues preventing the acquisition of an acceptable ultrasound image.

6.2.2 Ultrasound images

The data acquisition procedure was the same as the description presented in Chapter 5.2.2, where the subject wore a gown with the back open. Two operators were involved in the scanning procedure; one moved the transducer from C7 to L4 with the guidance of the custom frame and the other one recorded the ultrasound data. Each scan took approximately 30 seconds.

A total of 18 curvatures were recognized from the 13 subjects. For each curvature, the apical vertebra as well as the one above and the one below the apical vertebra were selected for VAR measurements. Ultrasound transverse images were exported through Tomoview software in bitmap image format (BMP) by 1:1 ratio. Multiple slices were projected to form one transverse image because the laminae from the same vertebra may not appear on one slice of the transverse image. Fifty-four vertebrae were used for the measurements. Two US transverse images missed one side lamina, and four images had poor quality because of the contact issue. These six images were then excluded from the study. Thus 48 measurements were reported. One of the US projected transverse images is shown in Figure 6.13.

6.2.3 Radiographs

The PA radiographs of the 13 subjects were exported from the local clinic database with the personal and clinical information removed. One of the radiographs is shown in Figure 6.14. The apical vertebra as well as the one above and the one below the apical vertebra of the curves were selected.

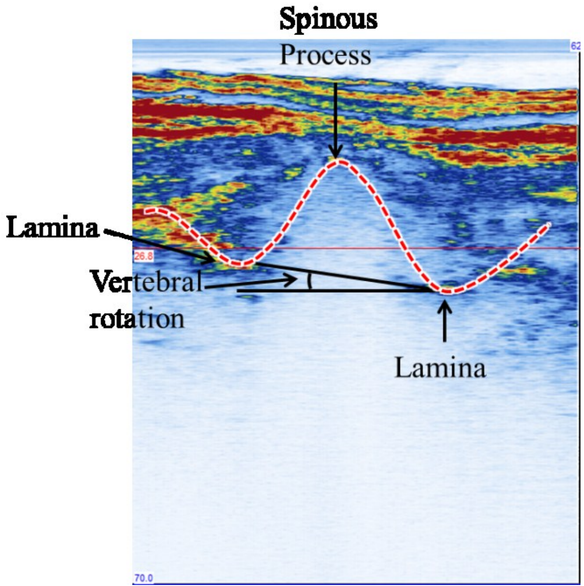


Figure 6.13 Projected US transverse image from one of the vertebrae.

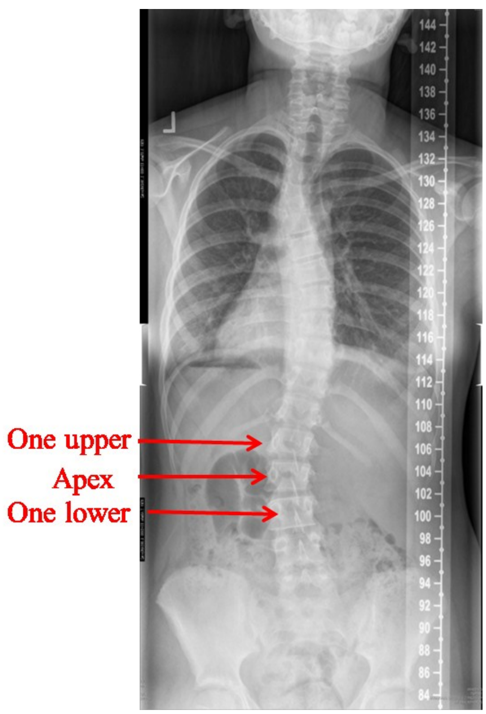


Figure 6.14 One of the PA radiographs of the AIS subject.

6.2.4 Measurements

Forty-eight US images from the AIS subjects were randomly assigned (www.random.org). However, the radiographs of the subjects were not randomly assigned because the observers needed to know the vertebra level for the width-depth ratio of the vertebra while using the Stokes' method.

The reliability of the VAR measurements from the PA radiographs using the custom program has been reported (Zhang *et al.*, 2010). The radiograph measurement results from the *in-vitro* experiments described in this chapter agreed with the results from Zhang's study (ICC: 0.82-0.86). Thus, in this clinical study, only one observer measured the VAR using the Stokes' method on the radiographs.

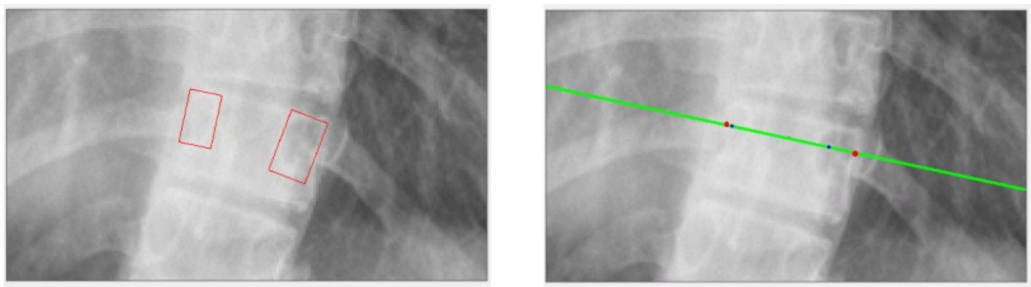
6.2.4.1 Observers

Two observers performed the US measurements in this study. Observer 1 (O1) and observer 2 (O2) both had 2-year experience on US measurements. All the measurements were performed twice with one week interval to reduce the recall bias.

6.2.4.2 Measurement methods

As shown in Figure 6.13, there was a top white layer indicating the biopsy material (the blue phantom). After that, the red layer was the soft tissue. The signals were quite noisy as it was an inhomogeneous medium. Although the 'W' shape (red dot line shown in Figure 6.13) was not as clear as the phantom image, the spinous process (SP) still could be recognized as the top points in the middle. The laminae were identified as the bottom red areas besides the SP (shadow area in white color) as the lowest points of the 'W' shape. The VAR angle was formed between the line drawing through the center points of the laminae and the line indicating the horizontal level. The horizontal level was parallel to the edge of the transducer which was assumed parallel to the surface of the back. ImageJ software (v1.45, NIH) was used to measure the VAR.

For the VAR measurements on radiographs, the same program was applied using the Stokes' method (Zhang *et al.*, 2010). The vertebra to be measured was first selected in a rectangular region from the whole spine as shown in Figure 6.15a. The procedures were then the same as the measurements of the phantom vertebra described in 6.1.8. The rotation angle was calculated using the four points (Figure 6.15b) by the Stokes' method (Stokes *et al.*, 1986).



(a) (b)

Figure 6.15 Custom program to measure the VAR on radiographs using Stokes' method. (a) Pedicle areas from the selected vertebra; (b) four points for calculation of the Stokes' method.

6.2.4.3 Statistical analysis

The intraclass correlation coefficient (ICC) (2-way random and absolute agreement), the mean absolute difference (MAD) with standard deviations (SD), and the Bland-Altman plot were applied for the analysis of the VAR measurements on *in-vivo* data. The same Currier criteria (Currier 1990) for ICC values were adopted. The intra-observer reliability was analyzed on both the COL and Stokes' methods. The inter-observer reliability was determined on the COL method

6.2.5 Intra-observer reliability

The intra-observer reliabilities of VAR measurements from both the COL and Stokes' methods are shown in Table 6.6. The measurements from both methods had high ICC values (>0.9) and small variations (MAD±SD) (<1.5°), which indicated high intra-observer reliability. Stokes' method has a larger range of ADs (0.1°-7.6°) compared with ADs of the COL method (0°-0.9°) from O1. Figure 6.16 and Figure 6.17 show the Bland-Altman plot indicating the differences between the measurements from two sessions of the COL and Stokes' methods from O1, respectively. Ninety-five percent of the differences fell in the range of -0.1°±0.8°, and 0.3°±4.1° (mean±2SD) for the measurements of COL and Stokes' methods, respectively. There was no significant correlation between the differences and the rotation severity for the intra-observer comparisons of the COL method (P=0.21) and Stokes' method (P=0.68).

Table 6.6 The intra-observer reliability of the COL method on US images and Stokes' method on radiographs for VAR of *in-vivo* data

	O1		O2
	COL	Stokes'	COL
ICC (95%CI)	>0.99 (>0.99)	0.94 (0.89-0.96)	0.95 (0.91-0.98)
Mean±SD (°)	6.3±3.3	6.6±5.6	5.7±3.7
MAD±SD (°)	0.3±0.2	1.5±1.4	0.7±0.7
AD range (°)	0-0.9	0.1-7.6	0-3.6

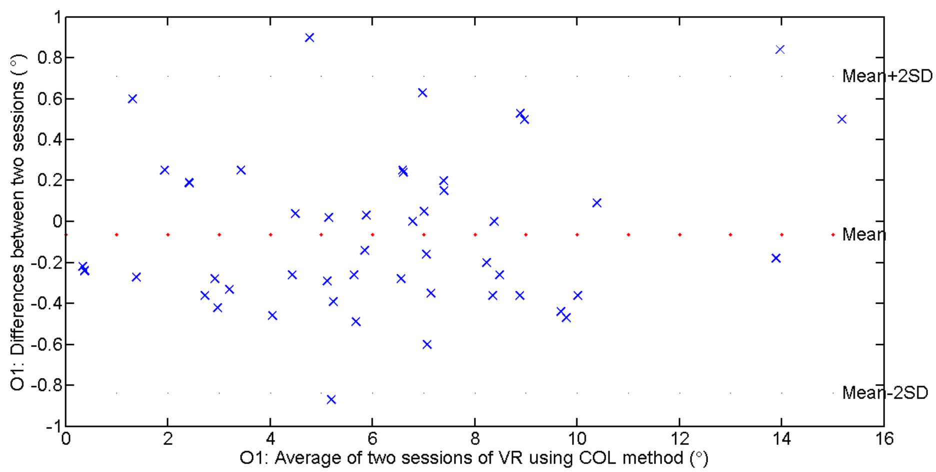


Figure 6.16 Bland-Altman plot of the intra-observer comparison of the COL method on *in-vivo* US images from O1.

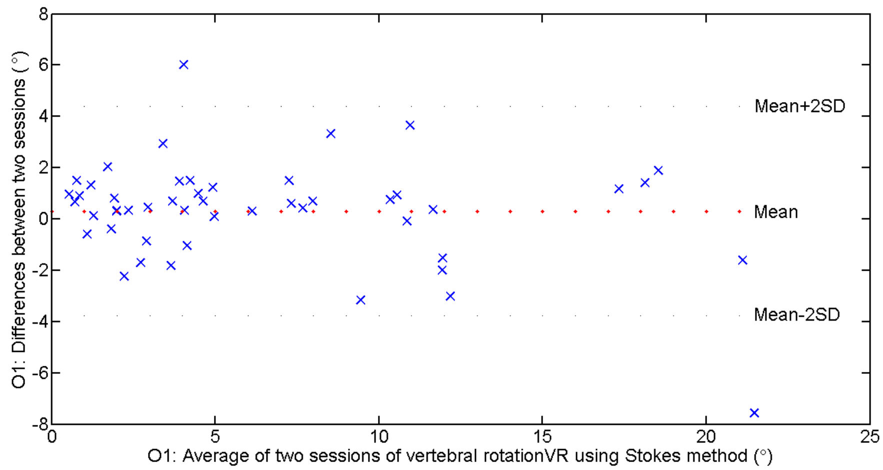


Figure 6.17 Bland-Altman plot of the intra-observer comparison of the Stokes' method on *in-vivo* radiographs from O1.

6.2.6 Inter-observer reliability

The VAR measurements using the COL method had high inter-observer reliability between O1 and O2 (ICC: 0.91 (0.82-0.96), MAD±SD: 0.9°±1.1°). Figure 6.18 shows the Bland-Altman plot of the inter-observer comparisons for the COL method. Ninety-five percent of the differences between the VAR measurements from O1 and O2 were within the range of 0.6°±2.5° (mean±2SD) and were independent of the rotation magnitude (P=0.70).

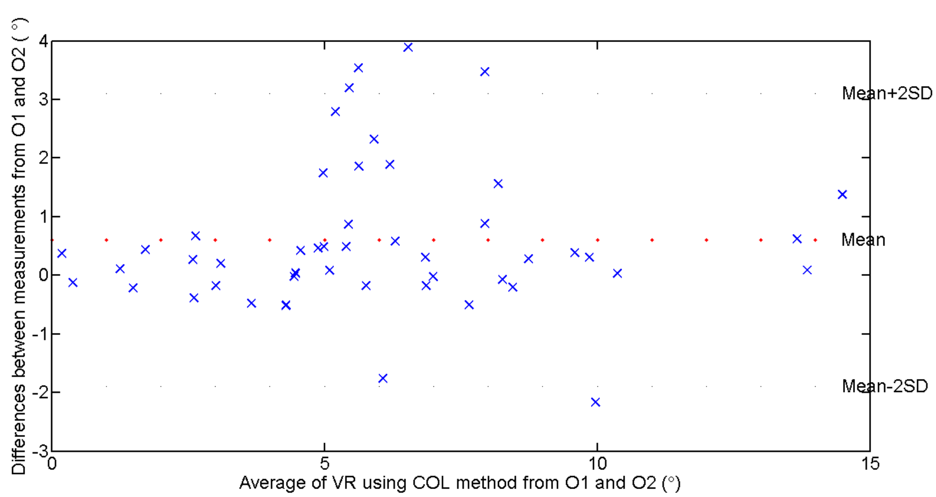


Figure 6.18 Bland-Altman plot of the inter-observer comparison of the COL method on *in-vivo* US images between O1 and O2.

6.3 Discussion

The vertebral axial rotation is an important parameter in the study of scoliosis. A simple and precise method can increase the use of this measurement in a scoliosis clinic. The reliabilities of the 2D and 3D methods to evaluate the VAR had been studied and compared (Ho *et al.*, 1992; Kuklo *et al.*, 2005; Richards, 1992; Stokes *et al.*, 1986; Vrtovec *et al.*, 2009; Weiss *et al.*, 1995). However, the radiographs methods are limited to the 2D images, which indirectly evaluate the rotation. Although the CT and MRI techniques can provide the transverse images of the vertebra for precise axial rotation assessment, these two techniques are not commonly used to diagnose AIS. Therefore, there was no standard method to quantify the VAR.

The proposed COL method could measure the VAR on ultrasound transverse images with a simple procedure. *In-vitro* and *in-vivo* studies showed high intra- and inter-observer reliabilities for the VAR measurements using the COL method. Ho's method measured from the CT transverse images had small standard deviations (1.2°-3.3°) for both the intra- and inter-observer reliabilities (Vrtovec *et al.*, 2009) and was reported as the most reliable method (Gocen *et al.*, 1998). The proposed COL method had small range of standard deviation (SD<1°). Also, the VAR measurements by the COL method were compared with the measurements by the Stokes' method on radiographs. Although Stokes' method had high reliabilities as the results reported by Zhang *et. al*, (2010), the VAR measurements using the COL method had better reliabilities and less variation (ICC >0.99, MAD: 0.4°-0.8°) than the measurements using Stokes' method (ICC: 0.90-0.96, MAD: 2.3°-3.2°).

The *in-vitro* validity of the VAR measurements from both the COL and Stokes' methods was studied by comparing with the experimental setup, which considered as the true value. The validity analysis indicated that the VAR measurements from the COL method (ICC: 0.96-0.97, MAD: 2.2°-2.3°) were more accurate compared with the measurements using the Stokes' method (ICC: 0.82-0.87, MAD: 4.1°-4.9°). Besides, the results from the VAR range investigation indicated that the proposed COL method on US images had the ability to assess bigger VAR than 50° with the errors smaller than 5°. However, the Stokes' method had limitation to evaluate the large rotations (Lam *et al.*, 2008), and the rotations greater than 50° could not be measured using the Stokes' method.

Among the three observers, two observers knew the 5° increment among the configurations, one did not know. To reduce the measurement bias, the ultrasound images were randomly assigned and the observers were required to perform the measurements according to the definition of the proposed COL method. Also, there were human errors when manipulating the phantom to different configurations. According to the results there was no significant difference among the measurements from three observers, which indicated this measurement bias did not affect the measurements significantly. Also the three observers had different US experience. Thus, the proposed COL method does not rely on the experience of the observers.

Although the validity of the proposed COL method was demonstrated in the *in-vitro* study, the validity of the *in-vivo* measurements had not been investigated. CT method could be used as the reference for the validity investigation; however, the supine posture affects the VAR measurements (Yazici *et al.*, 2001). Therefore, to validate the *in-vivo* VAR measurements using the COL method is a challenge and should be studied in the future.

6.4 Summary

The proposed COL method was a simple, precise and promising method to measure vertebral axial rotation on ultrasound transverse images. It had high intra- and inter-observer reliabilities. The validity of the VAR measurements using the COL method was demonstrated by comparing with the experimental setup. Compared with the Stokes' method, the COL method could provide more accurate and reliable VAR measurements and was able to evaluate large axial rotations.

Chapter 7.
Semiautomatic program for ultrasound measurements*

In-vitro and *in-vivo* studies described in Chapter 5 and 6 demonstrated that ultrasound imaging was a promising tool to assess the coronal curvature and vertebral axial rotation (VAR) in children with AIS. However, in the *in-vivo* study, soft tissue introduced more noise to the ultrasound data, which made some of the laminae difficult to be recognized. Also, the tilt of the vertebra in the frontal plane caused the laminae from the same vertebra might not appear on the same frame of the transverse image. In order to apply the COL method to measure the Cobb angle and VAR accurately, locations of the centers of laminae must be recognized or marked reliably. This chapter describes a semi-automatic program, which was developed to reduce the noise, export the coronal and transverse images from the areas including the lamina signals, determine the locations of the centers of laminae, and calculate the coronal curvature and VAR semi-automatically. *In-vitro* and *in-vivo* data were used to verify the performance of the program. Comparisons between the semi-automatic and manual measurements of the coronal curvature and VAR were also completed.

7.1 The semi-automatic measurement program

The developed program contained three sections: data pre-preprocessing, image reconstruction, and semi-automatic measurement.

---

*Part of the materials in this chapter has been included in an abstract accepted for the 5th International Conference on the Development of Biomedical Engineering, Vietnam, June 16-18th, 2014. Chen W, Lou EM, Le L (2014) A reliable semi-automatic program to measure the vertebral rotation using the center of lamina for adolescent idiopathic scoliosis.

7.1.1 Data pre-processing

Before processing the original 3D US data, the acquired Tomoview data (rdt file format) was transformed to MATLAB data format (mat file) using a custom program.

The original US data was a 3D matrix based on the coordinate of the data described in Chapter 2.4.2. The 3D matrix consisted of 115×601×2364 voxels. The 115 groups were determined based on the 14 elements of the transducer as a group which was explained in Chapter 5.1.3. The scanning length was set to be 60 cm, which could cover the entire spinal column virtually on all children with AIS. There were 601 frames along the scanning direction with 1 mm resolution. The sample frequency along the depth was 25 MHz and the scanning depth was set to be 7 cm. Thus, there were 2364 points of each A-line along depth direction with the sample resolution of 0.0296 mm. The pre-processing procedure is illustrated in the flow chart shown in Figure 7.1.

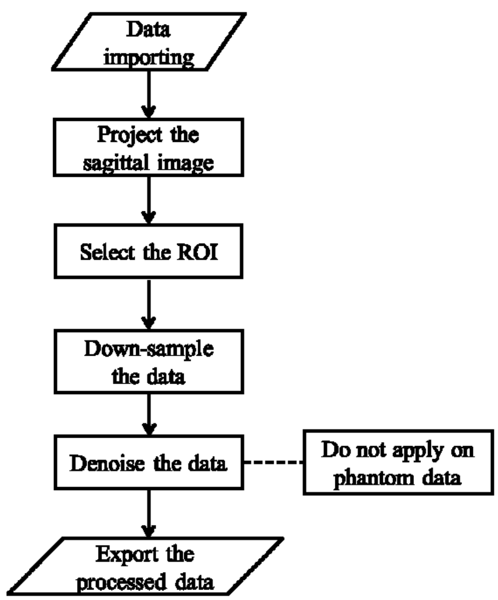


Figure 7.1 The flow chart of the pre-processing procedure.

After the 3D original data was imported into the program, the sagittal image was displayed using the maximum intensity projection (MIP) method in which the maximum values along the element index direction were used to form the sagittal image. In Figure 7.2, the blue color indicates the intensity value of zero, which means no US signals are included. The red color indicates the strong reflection signals. A region of interest (ROI) was selected in the sagittal image using a rectangular box, which included the majority of the signals along the scan length and depth. The data in blue color was excluded from the ROI. Since a soft-tissue biopsy material (described in Chapter 5.2.2) was applied, there was a layer without reflection signals on the top, which was also excluded from the ROI. This step reduced the data size to speed up the process.

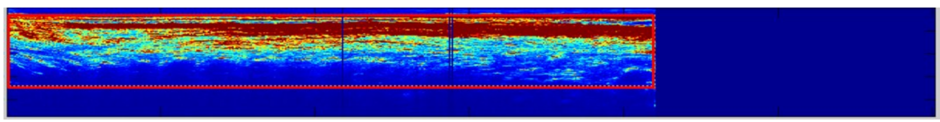


Figure 7.2 Projected sagittal image with the rectangular box to select the ROI.

The sampling resolution along the depth was 0.0296 mm, which was much smaller than the dimensions of a vertebra. To further improve the processing speed, the data was compressed. The approach was to down-sample the data; one data point was chosen from every four data points along each A-line using the MIP method. The projected sagittal image after the ROI has been selected and down-sampling has been processed is shown in Figure 7.3. The size of the data became 115×420×1716 after the ROI selection, and reduced to 115×420×429 after down-sampling, which was 12.7% of the original data set. The processing time from importing the data to down-sampling was within one minute (using a computer with Intel Core i5-2500 CPU and 8 GB RAM).

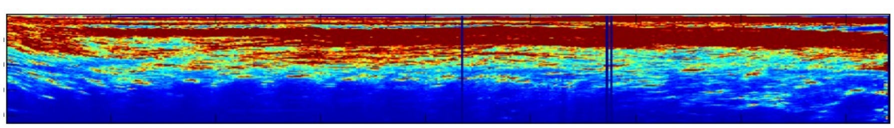


Figure 7.3 Projected sagittal image after ROI selection and down-sampling.

In general, speckle noise, which is generated by the interaction of the reflection signals and the scatters in the soft tissue, is a common problem in ultrasound images. Speckle noise reduces the quality of the image and makes the bone structure difficult to recognize. Thus, denoising is required to remove speckle noise and preserve details of the original data (Coupe *et al.*, 2009; Guo *et al.*, 2008; Gupta *et al.*, 2005). There were 420 frames along the scanning direction. The data was processed frame by frame. Before denoising, the data was first normalized.

A general model of speckle noise, which is considered as multiplicative, is given by (Achim *et al.*, 2001; Vanithamani *et al.*, 2011)

$$f_{i,j} = g_{i,j}u_{i,j} + \alpha_{i,j} \tag{7-1}$$

where $f_{i,j}$ is the observed image, $g_{i,j}$ indicates the noise free image, $u_{i,j}$ and $\alpha_{i,j}$ represent the multiplicative and additive component of the speckle noise, respectively, and (i, j) indicate the coordinates of the pixels in the image. Since the effect of the additive noise is considerably less significant than the multiplicative component, the US image can be expressed as

$$f_{i,j} \approx g_{i,j}u_{i,j} \tag{7-2}$$

In order to transform the multiplicative noise to additive one, the logarithm of the data was performed as

$$\log(f_{i,j}) = \log(g_{i,j}) + \log(u_{i,j}) \tag{7-3}$$

Then, the wavelet threshold denoising method was applied on the logarithmic data. The discrete wavelet transform (DWT) decomposes the images along rows and columns, and replaced with four blocks containing low frequency and high frequency components along each direction. The DWT first performs one step on all rows that the left side contains low pass coefficients and the right side contains high pass coefficients. Then one step is applied on all columns and obtains four types of coefficients (LL, HL, LH, and HH) (Sudha *et al.*, 2009). The LL contains the low frequency components in both directions, whereas HL, LH, and HH contain the high frequency components in rows, columns, and both directions, respectively. The LL part could be further decomposed and a two-level image decomposition using DWT is applied as shown in Figure 7.4

The speckle noise is in the high-frequency band. A soft threshold was used to remove the noise using the Symlets 4 wavelet basis (Achim *et al.*, 2001; Sudha *et al.*, 2009). The soft threshold is defined as follows:

$$D(U, t) = sgn(U)\max(0, |U| - t) \tag{7-4}$$

Where, U indicates the wavelet coefficient, t indicates the threshold, which can be determined using Stein's principle of unbiased risk estimation (Donoho and Johnstone, 1995). After thresholding, the inverse DWT was applied to get the denoised data. Then, the exponentiation operator was applied on the denoised data. Figure 7.5 shows one of the transverse images before and after denoising.

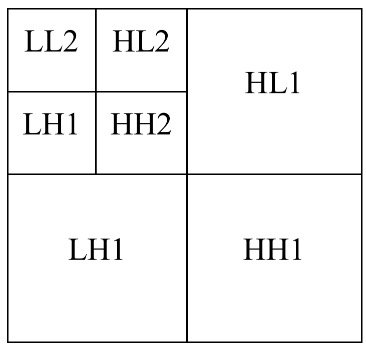
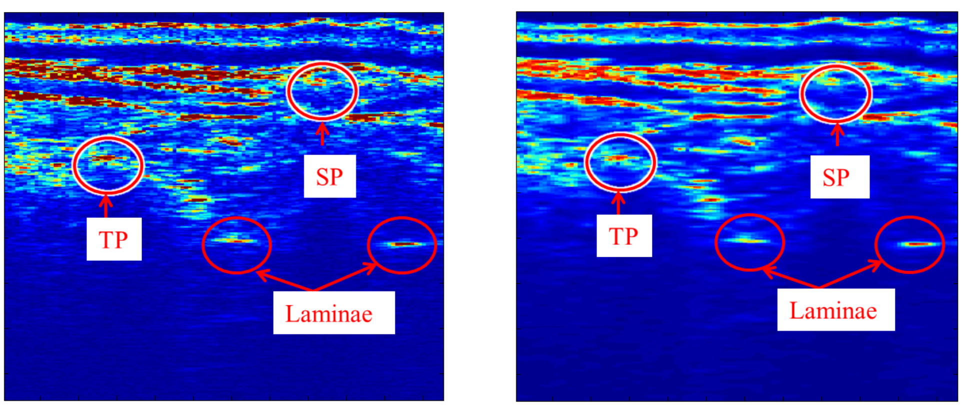


Figure 7.4 Two-level image decomposition using DWT.



(a) (b)

Figure 7.5 Transverse images of a vertebra from an AIS subject: (a) image before denoising, and (b) image after denoising.

7.1.2 Image reconstruction

The data after pre-processing was loaded for image reconstruction. In order to generate the coronal image including the laminae signals accurately, the depth range of the laminae needed to be determined. After the pre-processing, the corresponding projected sagittal image similar to Figure 7.3 was displayed using the MIP method. Since the laminae were the lowest points in a ‘W’ shape, the laminae could be identified as the lowest strong reflectors from the sagittal image. Two red lines were formed by connecting several data points, which were selected above and below the laminae (see Figure 7.6). The depth range of the laminae was determined by these two lines. The coronal image was projected from the data volume among the selected depth using the MIP method (Figure 7.7). Each point on the coronal image indicated the maximum intensity among the points along the selected depth.

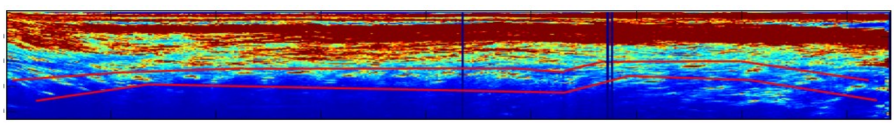


Figure 7.6 Two lines drawn on the projected sagittal image.

Figure 7.7 shows the coronal image of the selected areas. The elliptical red spots (relatively strong reflection signal areas) besides the column of the spinous process (dark blue) are the laminae. The pair of laminae from the same vertebra was perpendicular to the SP column in the center. In this image, a thoracolumbar curve is identified. To extract the reflection signals of the laminae from the same vertebra, two red lines were then drawn on top and bottom of both laminae of the same vertebra (Figure 7.7). Then the projected transverse image was formed using the MIP method (Figure 7.8). The laminae were the red areas besides the SP shadow zone at the lowest positions in the ‘W’ shape on the transverse image. Transverse images of the apex, one upper and one lower vertebra were exported for VAR measurements. Normally the vertebral rotation at the apical vertebra is maximal and orthopedic surgeons are interested in using this rotation value to predict the progression and evaluate the treatment outcomes.

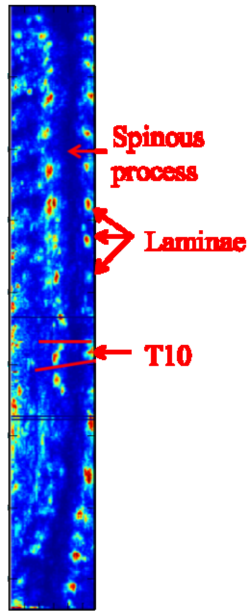


Figure 7.7 Projected coronal image for measurement. Two lines were drawn for exporting the transverse image from apex.

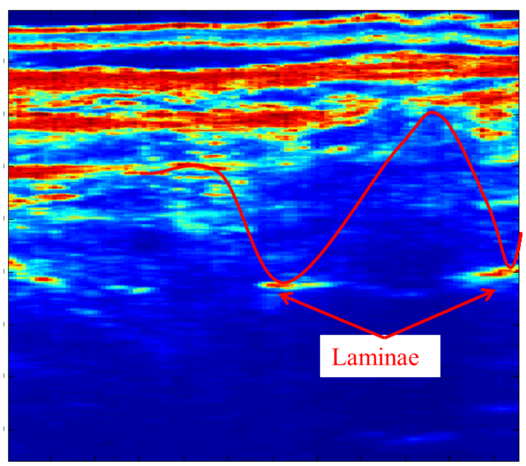


Figure 7.8 The projected transverse image.

7.1.3 Semi-automatic measurement

Semi-automatic measurements of the coronal curvature and vertebral rotation were performed on the exported images from the previous section. Both the coronal and transverse images were left-right flipped since the lateral direction of the image was opposite to the posteroanterior radiograph.

7.1.3.1 Coronal curvature calculation

Dual interpolation is a fast linear method, which interpolates three points into two points. It was performed along the element index direction to make the image smoother. After the interpolation, the size of the coronal image became 457×420 from 115×420 (Figure 7.9). Then the operator selected the laminae by using a mouse click from the end vertebrae, which were the most tilted vertebrae. A 100×20 pixel area (12.5 mm×20 mm, red rectangular boxes in Figure 7.9) around the marked point was decided based on the dimensions of the vertebrae to make sure the lamina area was fully covered. Therefore, for a single curve, four points were selected by the operator. The program then automatically determined the four lamina areas for each curvature (Figure 7.9). The rectangular boxes shown in Figure 7.9 are for illustration purpose and not displayed while the program is running.

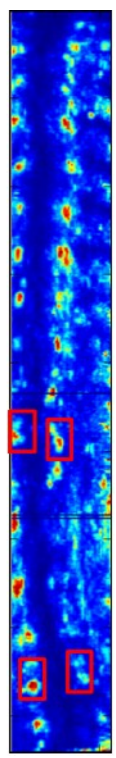
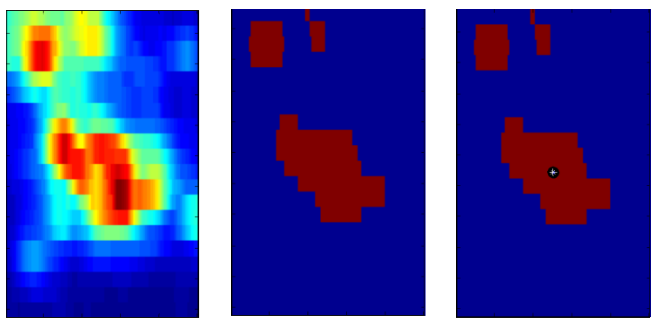


Figure 7.9 Coronal image after interpolation.

After the program segmented the lamina areas, the center points of the laminae were automatically determined. Figure 7.10 shows the center point calculation procedures using the bottom right lamina. After the program determined the areas, the intensities of the reflected signals were first normalized and a median filter was applied to remove the sparse noise (Figure 7.10a). A Laplacian filter, which was a second order derivative operator, was also applied to enhance the edge (Gonzalez, *et al.*, 2009) of the area. Then the optimum global thresholding Otsu's method segmented the image into two parts (Gonzalez, *et al.*, 2009; Otsu, 1979). The threshold values were determined by maximizing the between-class variance. The blue indicated the background signals below the threshold value, and the red indicated the objective signals above the threshold value (Figure 7.10b). In the binary image, there were several red areas, and the lamina was determined as the one including the marked point. Then the geometric center point of the lamina was calculated and marked (Figure 7.10c).



(a) (b) (c)

Figure 7.10 Procedures of determining the center point of lamina. (a) Image after normalizing and median filter, (b) image after segmentation, and (c) image with the center point of the lamina.

One line was drawn automatically through the center points of the laminae at the end vertebrae. The angle formed by the two lines from the upper and lower end vertebrae was calculated and indicated the coronal curvature using the COL method (Figure 7.11). The operator could repeat the measurements and select different vertebrae as the end vertebrae to check which provided bigger lateral deviation for the curve. If there were more than one curve, multiple lines would be obtained, and the angles between any two adjacent lines were calculated.

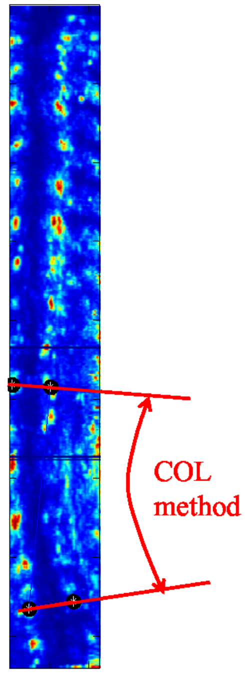


Figure 7.11 The COL method on the coronal image using the program.

7.1.3.2 Vertebral rotation calculation

A similar procedure was used for the VAR measurement. After the transverse image was regenerated from the program as described in 7.1.2, the dual interpolated method was applied to smooth the transverse image. The operator then selected the laminae by using a mouse click. A 100×30 pixel (12.5mm×3.6mm) area was determined to make sure the lamina could be covered. In each area, the lamina was automatically segmented from the background and the geometric center point of the lamina was calculated. One line was drawn through the center points of the laminae on the same vertebra and the angle between the line and the reference horizontal line indicated the VAR (Figure 7.12).

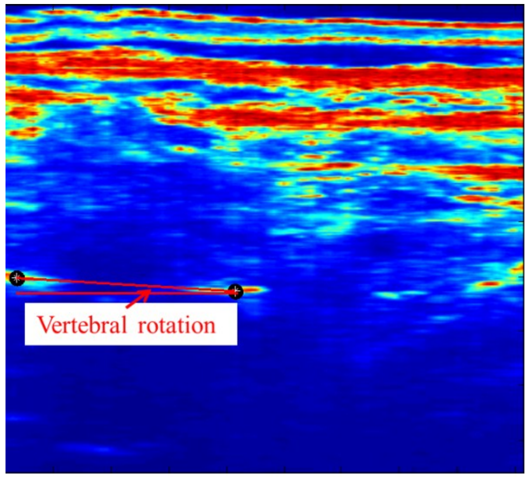


Figure 7.12 The vertebral rotation of T10.

7.1.4 Procedure of the semi-automatic program

The developed program included three parts, and several steps were followed for each part.

Section 1: Data pre-processing

Step 1: Run the program, select the data file with original US data for processing. A sagittal image was displayed.

Step 2: Draw a rectangular box to exclude the dark blue area on the displayed coronal image. The data was then down-sampled and denoised.

Step 3: Save the processed data as a mat file.

Section 2: Image reconstruction

Step 1: Run the program, import the data file saved from Section 1.

Step 2: Select several points first above then below the laminae on the displayed projected sagittal image to determine the range where the laminae can be found. Two lines were drawn through the points, and the coronal image was displayed.

Step 3: Draw two lines above and below the pair of laminae from the same vertebra. The transverse image was then formed and displayed.

Step 4: Save the coronal and transverse images as a mat file.

Section 3: Semi-automatic measurement

Step 1: Run the program, import the data file saved from Section 2.

Step 2: Perform the coronal curvature measurement. Point the laminae from upper and lower end vertebrae using a mouse. The calculation was then displayed on the command window.

Step 3: Perform the vertebral axial rotation measurement. Point the pair of laminae. The calculation was then displayed on the command window.

Step 4: Record the measurements and save the measurements results as a mat file.

7.2 Semi-automatic program evaluation

To investigate the reliabilities of the program, both the *in-vitro* and *in-vivo* data from Chapter 5 and 6 were used to perform the coronal curvature and VAR measurements using the developed semi-automatic program.

7.2.1 Observers

Two observers performed the measurements of both the coronal curvature and VAR using the program twice with one week interval. Observer 1 (O1) had experience reviewing the US images for 3 years; observer 2 (O2) only had 2 month experience on examining the US images. O1 and O2 went through the measurement procedure described in 7.1.4 using the developed program prior to the study.

7.2.2 Statistical analysis

The intraclass correlation coefficient (ICC) (2-way random and absolute agreement) (Shrout and Fleiss, 1979; McGraw and Wong, 1996) and mean absolute difference (MAD) with standard deviations (SD) were used to analyze the intra- and inter-observer reliabilities of both the coronal curvature and VAR measurements. Measurements of the COL method were compared with the scoliosis clinical record of the Cobb method. The agreement between the VAR measurements and experimental setup was also analyzed. The Currier criteria for ICC values were adopted (Currier 1990). The Bland-Altman plot was also presented to show the differences for the intra- and inter-observer comparisons (Bland and Altman, 1986; 1999; 2010). The relationship between the differences and the magnitude of the measurements was investigated by Pearson's correlation coefficient. The significant level was set at $\alpha = 0.05$. Differences were significantly correlated with magnitude when the P-value was smaller than α. The end vertebra selections were analyzed for the COL method on the phantom data using the error index (EI) (Morrissy *et al.*, 1990; Oda *et al.*, 1982).

7.2.3 Image preparation

For the coronal curvature measurements, fifteen coronal images including 30 curves from the phantom data were used. In the 26 coronal images from the *in-vivo* data, the program was not able to perform the segmentation to distinguish the laminae from the background on three of the images, and therefore 23 coronal images with 25 curves from AIS subjects were used to verify the program on coronal curvature measurements. For the VAR measurements, 40 transverse images from the phantom data and 48 images from the subjects' data were applied. All the coronal and transverse images were exported and the image numbers were randomly assigned using the free service from the web site (www.random.org) before the measurements for both sessions.

For the coronal curvature, the end vertebrae selection were investigated for the *in-vitro* study and the end vertebrae were labeled in the *in-vivo* data to focus on the variations between the COL method and Cobb method for the manual measurements. In order to compare the semi-automatic measurements with the manual measurements, the end vertebra were not determined before the measurements for the *in-vitro* study, but they were preselected for *in-vivo* data to eliminate the error effect from the selection of end vertebrae as described in the manual measurements in Chapter 5.2. The most tilted vertebrae were selected as the end vertebrae. Then the positions of the end vertebrae were included in the data file and the labels of the end vertebrae could be displayed on the coronal images when running the program.

7.2.4 Evaluation on the *in-vitro* coronal curvature measurement

The intra-observer reliability results of the semi-automatic coronal curvature measurements on the ultrasound images are shown in

Table 7.1. The ICC values from both observers were higher than 0.97 which indicated high intra-observer reliability. The mean±SD of the *in-vitro* coronal curvature measurements were 39.7°±10.4° and 40.3°±10.4° from O1 and O2, respectively. The MAD±SD between the measurements from two sessions were 1.6°±1.9° and 1.4°±1.6° for O1 and O2, respectively.

Table 7.1 Intra-observer reliability of the COL method on *in-vitro* data using the program

	O1	O2
ICC (95% CI)	0.97 (0.94-0.99)	0.98 (0.95-0.99)
Mean±SD (°)	39.7±10.4	40.3±10.4
MAD±SD (°)	1.6±1.9	1.4±1.6

Figure 7.13 and Figure 7.14 show the Bland-Altman plot indicating the differences between the two sessions' measurements from O1 and O2, respectively. Ninety-five percent of the differences fell in the range of -0.03°±5.1° and -0.9°±4.0° (mean±2SD), respectively, which indicated there were small variations between the measurements from two sessions. There was no significant correlation between the differences and the curve severity for the intra-observer comparisons of O1 (P=0.27) and O2 (P=0.27).

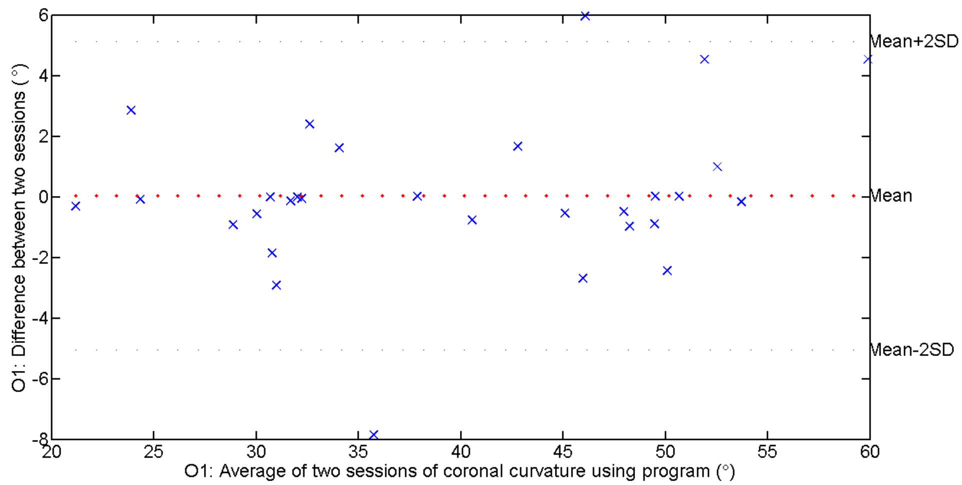


Figure 7.13 Bland-Altman plot of the intra-observer comparison of the COL method from O1 on *in-vitro* data using the program.

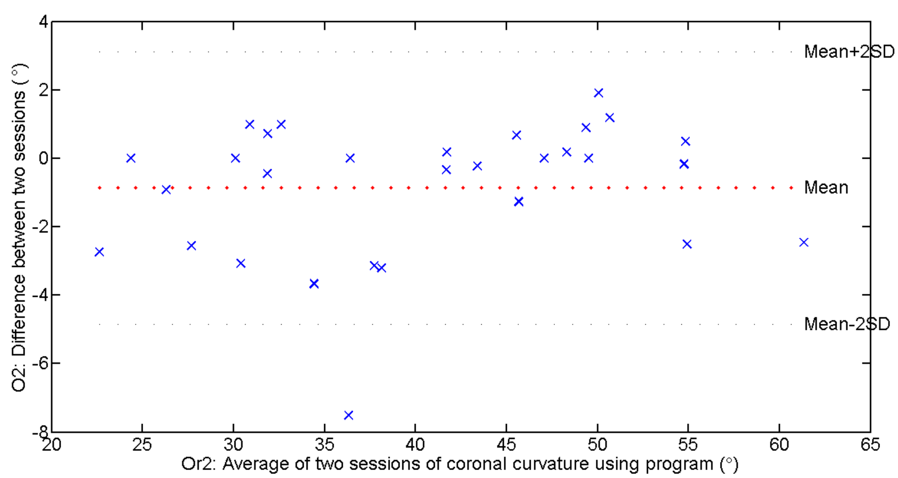


Figure 7.14 Bland-Altman plot of the intra-observer comparison of the COL method from O2 on *in-vitro* data using the program.

Table 7.2 presents intra-observer differences of the end vertebrae selections for the semi-automatic measurements on the phantom data. For the 30 curvatures, 30 upper-end vertebrae (UEV) and 30 lower-end vertebrae (LEV) were selected. All end vertebrae selections were within 2 vertebrae differences, and 75% ((83.3%+66.7%)/2) and 73.3% ((76.7%+70.0%)/2) of the end vertebrae were consistent between the two sessions for O1 and O2, respectively. The error indices (EI) were small (0.60 and 0.62), which indicated the end vertebrae could be reliably identified.

Table 7.2 The intra-observer end vertebra differences of the COL method on *in-vitro* data using the program

Observers	0		±1		±2 or more		Error Index
	UEV	LEV	UEV	LEV	UEV	LEV	
O1	25 (83.3%)	20 (66.7%)	3 (10.0%)	7 (23.3%)	2 (6.67%)	3 (10.0%)	0.62
O2	23 (76.7%)	21 (70.0%)	6 (20.0%)	6 (20.0%)	1 (3.3%)	3 (10.0%)	0.60

The average values of the measurements from the two sessions were used to analyze the inter-observer reliability. The results are shown in Table 7.3. The high ICC value (0.91) and small MAD±SD (3.1°±3.0°) indicated the high inter-observer reliability for the COL method measured by the program.

Table 7.3 Inter-observer reliability of the COL method on the *in-vitro* data using the program

	O12
ICC (95% CI)	0.91 (0.83-0.96)
MAD±SD (°)	3.1±3.0

Figure 7.15 shows the Bland-Altman plot of the comparison of the semi-automatic coronal curvature measurements between O1 and O2. Ninety-five percent of the differences between the measurements from the two observers were within the range of -0.6°±8.6° (mean±2SD). The differences were independent of the curve magnitude (P=0.94). From the Bland-Altman plot, most of the differences were within the range of 5°. However, two measurements had differences over 10°. Different red areas were chosen as the laminae on the lumbar part for these two curves.

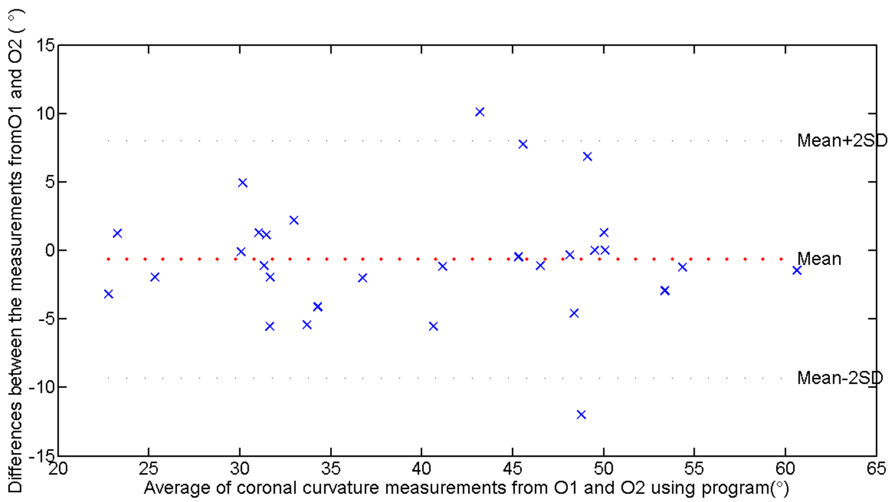


Figure 7.15 Bland-Altman plot of the inter-observer comparison of the COL method between O1 and O2 on *in-vitro* data using the program.

All the end vertebrae selections from the two sessions were included and a total of 60 UEV and 60 LEV were compared. The differences of the end vertebrae selections between the two observers are shown in Table 7.4. The two observers had 42 (70%) same UEV and 35 (58.3%) same LEV with the EI value of 0.98. The end vertebrae selection had more differences between the observers than the intra-observer selections.

Table 7.4 The inter-observer end vertebrae differences of the COL method on *in-vitro* data using the program

	0	±1	±2 or more	Error Index
UEV	42 (70.0%)	11 (18.3%)	7 (11.7%)	0.98
LEV	35 (58.3%)	17 (28.3%)	8 (13.3%)	

7.2.5 Evaluation on the *in-vivo* coronal curvature measurement

The intra-observer reliability of the semi-automatic measurements on the ultrasound images are shown in Table 7.5. Both observers presented high intra-observer reliability and O1 had higher ICC value (0.96) when compared with O2 (0.85). Also, the MAD±SD values were 1.4°±1.6° and 2.4°±2.8°, respectively. The mean±SD of the Cobb angle from the scoliosis clinical record was 22.3°±7.5°. The means of the semi-automatic US measurements using the COL method were smaller than the clinical record of the Cobb angle by 2.2° and 3.3° for O1 and O2, respectively.

Table 7.5 Intra-observer reliability of the COL method on *in-vivo* data using the program

	O1	O2
ICC (95% CI)	0.96 (0.89-0.98)	0.85 (0.70-0.93)
Mean±SD (°)	20.1±7.3	19.0±6.2
MAD±SD (°)	1.4±1.6	2.4±2.8

Figure 7.16 and Figure 7.17 show the Bland-Altman plot for the intra-observer comparisons from O1 and O2, respectively. Ninety-five percent of the differences for O1 and O2 fell in the range of 1.0°±3.7° and -0.8°±7.3° (mean±2SD), respectively. O2 had larger variation for the measurements between sessions. There was no significant correlation between the differences and the curve severity for the measurements from O1 (P=0.39) and O2 (P=0.81).

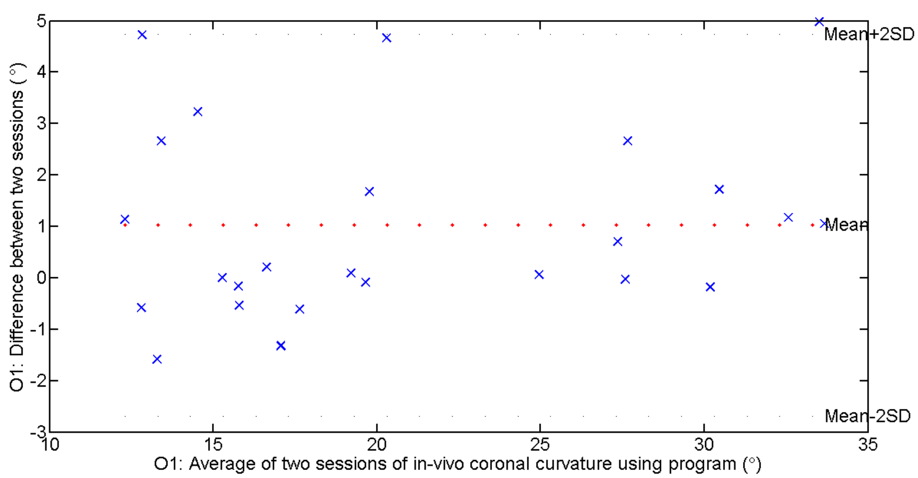


Figure 7.16 Bland-Altman plot of the intra-observer comparison of the COL method from O1 on *in-vivo* data using the program.

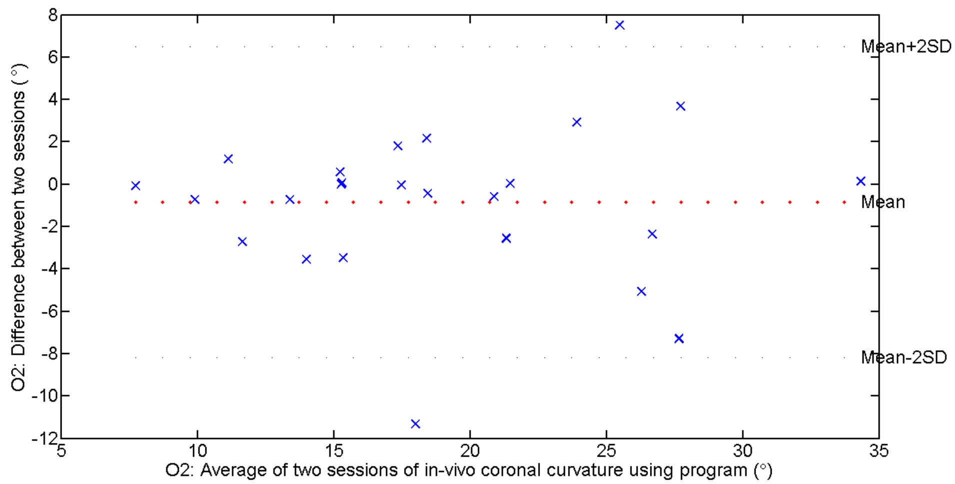


Figure 7.17 Bland-Altman plot of the intra-observer comparison of the COL method from O2 on *in-vivo* data using the program.

The average values of the measurements from the two sessions were used for the inter-observer reliability analysis. The results are shown in Table 7.6. The ICC value (0.76) indicated only fair inter-observer reliability between the two observers. The MAD±SD was 3.4°±3.6°. Figure 7.18 shows the Bland-Altman plot of the inter-observer comparison between O1 and O2. Ninety-five percent of the differences between the two observers were within the range of 2.0°±7.0° (mean±2SD). The difference between the measurements from the two observers was independent of the curve magnitude (P=0.39).

Table 7.6 Inter-observer reliability of the COL method on the *in-vivo* data using the program

	O12
ICC (95% CI)	0.76 (0.51-0.89)
MAD±SD (°)	3.4±3.6

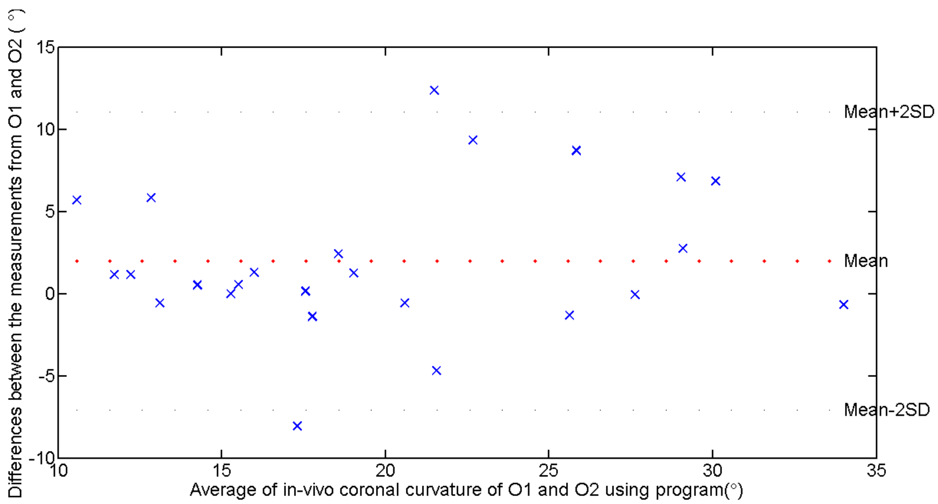


Figure 7.18 Bland-Altman plot of the inter-observer comparison of the COL method on *in-vivo* data using the program.

The semi-automatic measurements using the COL method were compared with the Cobb angle from the clinical record. Results are shown in Table 7.7. The semi-automatic measurements from O1 agreed well with the clinical record (ICC=0.89), while the measurements from O2 only showed a fair agreement with the clinical record (ICC=0.72). The MAD±SD were 2.6°±2.3° and 4.4°±3.4° for O1 and O2, respectively, and O2 had larger errors than O1.

Table 7.7 Agreement between the measurements of the COL method on US images using the program and Cobb angle from clinical record on the *in-vivo* data

	O1	O2
ICC (95% CI)	0.89 (0.77-0.95)	0.72 (0.34-0.88)
MAD±SD (°)	2.6±2.3	4.4±3.4

Figure 7.19 and Figure 7.20 show the Bland-Altman plot of the coronal curvature measurements of the COL method versus the measurements of Cobb method from O1 and O2, respectively. Ninety-five percent of the differences fell in the range of -1.1°±6.7° and -3.1°±9.2° (mean±2SD) for O1 and O2, respectively. The differences were not significant related to the average of the curve severity of the two methods for O1 (P=0.89) and O2 (P=0.35), respectively.

Both the MAD±SD and the Bland-Altman plot indicated that the measurements from O2 had more variation than the measurements from O1 when compared with the scoliosis clinical record.

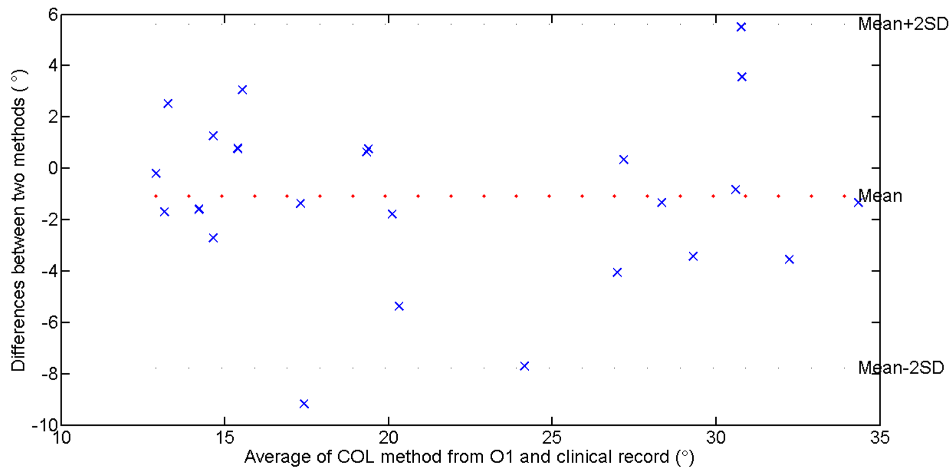


Figure 7.19 Bland-Altman plot of the comparison between the semi-automatic measurements of the COL method from O1 and clinical record of Cobb method on *in-vivo* data.

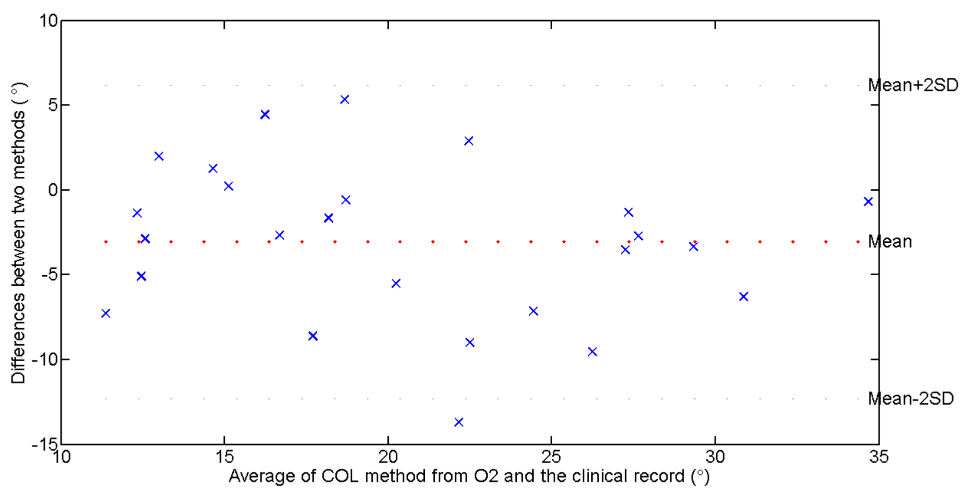


Figure 7.20 Bland-Altman plot of the comparison between the semi-automatic measurements of the COL method from O2 and clinical record of Cobb method on *in-vivo* data.

7.2.6 Discussions on the coronal curvature investigation

Semi-automatic coronal curvature measurements on the phantom data were compatible with manual measurements (Chapter 5.1). The MAD values of the intra- and inter-observer comparisons were smaller than the MAD values from the manual measurements. The program could reduce the measurement errors on phantom data. Also, the end vertebra selections using the program were more consistent for the intra-observer comparison than those of the manual measurements. When comparing the semi-automatic measurements with the manual measurements on *in-vivo* data (Chapter 5.2), O2 had lower ICC value compared to the manual measurements. The inter-observer reliability of the semi-automatic measurements was not as good as the reliability of the manual measurements. Also, when compared with the manual measurements, the semi-automatic measurements using the COL method had less agreement with the clinical record of Cobb method.

The same procedures were performed for the *in-vitro* and *in-vivo* data. The depth range selection of the laminae on the sagittal image could introduce noise that signals from other structures might project on the coronal image. Soft-tissues introduced noise and attenuated the signals, which could limit the program to accurately distinguish the lamina area. Especially when the signals from laminae are weak, the program may not be able to segment the lamina from the background. Besides, during the manual measurements, the SP was considered as a reference. The line went through the pair of laminae should perpendicular to the trend of SP, which was the middle area with the white color (Figure 5.12). However, this criterion was not considered in the semi-automatic measurements. For the same vertebra, the selection of the lamina area is crucial to the evaluation of coronal curvature. Take the lower end vertebra from Figure 7.9 as an example, there would be a big difference in the curvature angle if the line was drawn through different reflection areas (Figure 7.21). The selections of the laminae were crucial for the coronal curvature measurements using the COL method. In order to eliminate this error source, removing the noise from soft tissues and other bony structures is important and then the laminae could be selected more reliably and accurately.

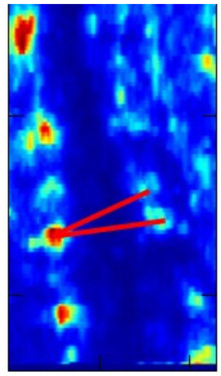


Figure 7.21 Example of different lamina selections.

The results on the subjects' data indicated that the US experience affected the reliability and that the experienced observer could provide more reliable semi-automatic measurements.

7.2.7 Evaluation on *in-vitro* vertebral rotation measurement

The intra-observer reliability results of the semi-automatic VAR measurements are shown in Table 7.8. The ICC values (>0.98) indicated high intra-observer reliability for both observers. The MAD±SD values were 0.2°±0.3° and 0.9°±1.6° for O1 and O2, respectively. The mean±SDs from the two observers were close to 16.8°±10.2° of the experimental setup value.

Table 7.8 Intra-observer reliability of the COL method for VAR on *in-vitro* data using the program

	O1	O2
ICC (95% CI)	>0.99 (>0.99)	0.99 (>0.97)
Mean±SD (°)	16.5±10.1	16.6±10.3
MAD±SD (°)	0.2±0.3	0.9±1.6

Figure 7.22 and Figure 7.23 show the Bland-Altman plot indicating the differences between the measurements from two sessions of O1 and O2, respectively. Ninety-five percent of the differences fell in the range of 0.1°±0.6° and -0.03°±3.6° (mean±2SD) for O1 and O2, respectively, which indicated small

variation. There was no significant relationship between the differences and the VAR magnitude for the intra-observer comparisons of O1 (P=0.23) and O2 (P=0.11) on subjects' data.

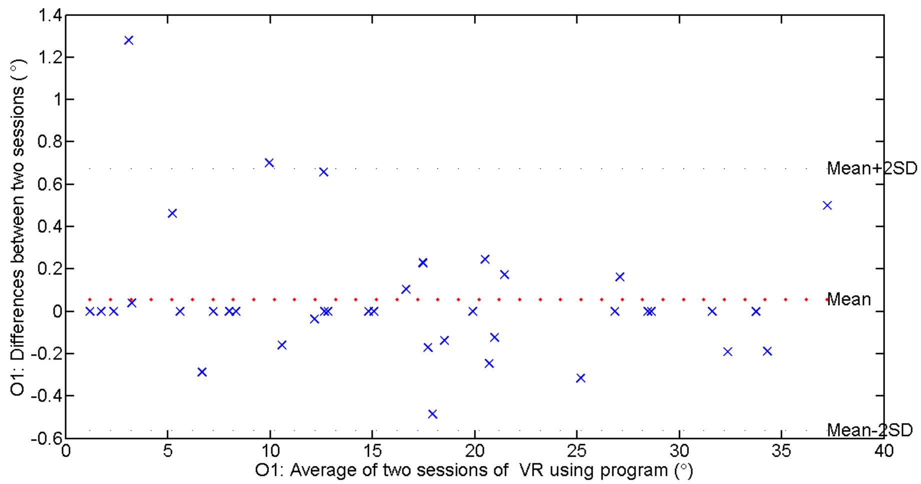


Figure 7.22 Bland-Altman plot of the VAR intra-observer comparison of the COL method from O1 on *in-vitro* data using the program.

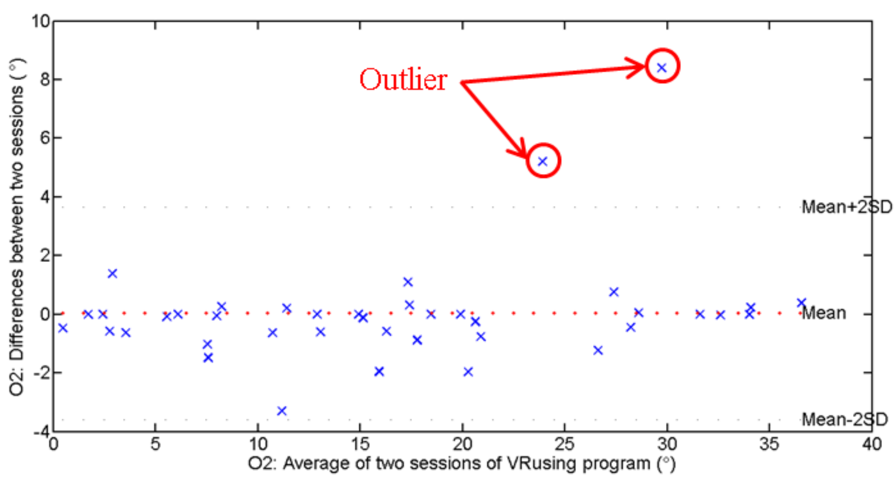


Figure 7.23 Bland-Altman plot of the VAR intra-observer comparison of the COL method from O2 on *in-vitro* data using the program.

There were two outliers with the values of 5.2° and 8.4° in the intra-observer comparison for O2 shown in Figure 7.23. That was caused by selecting different areas as the lamina. Take the one which had a 8.4° difference as an example (Figure 7.24), the observer selected two different reflection areas as lamina.

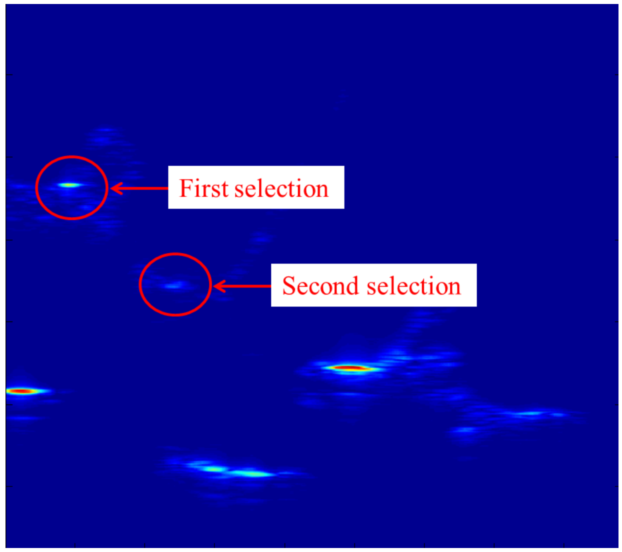


Figure 7.24 One example of the different selections of lamina areas between two sessions.

Table 7.9 summarizes the inter-observer reliability results for the semi-automatic VAR measurements on the phantom data. Average values were applied for the analysis. The results (ICC: 0.99, MAD±SD: 0.8°±1.1°) indicated very high inter-observer reliability between the two observers.

Table 7.9 The inter-observer reliability of the COL method for VAR on *in-vitro* data using the program

	O12
ICC (95% CI)	0.99 (>0.98)
MAD±SD (°)	0.8±1.1

Figure 7.25 shows the Bland-Altman plot of the inter-observer comparison for the semi-automatic VAR measurements. Ninety-five percent of the differences between the measurements from O1 and O2 were within the range of -0.2°±2.8° (mean±2SD) and were independent of the VAR measurements (P=0.36).

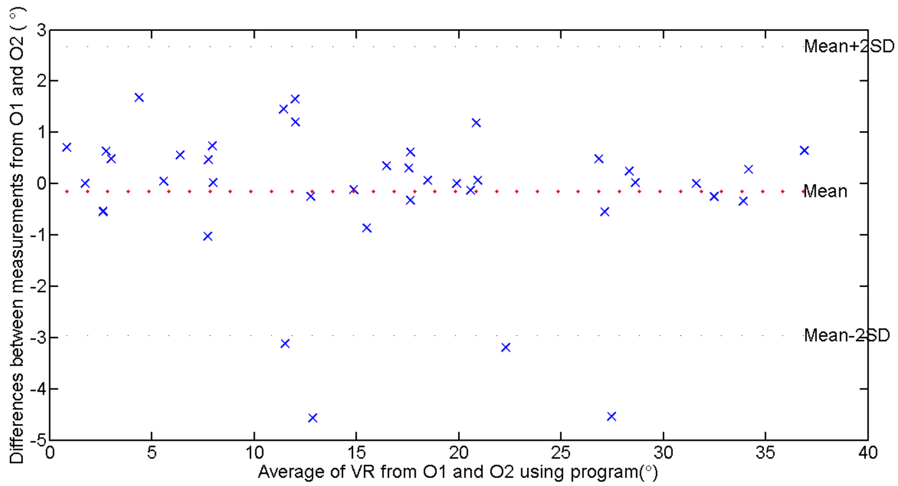


Figure 7.25 Bland-Altman plot of the VAR inter-observer comparison of the COL method on *in-vitro* data using the program.

The validity comparison results between the US measurements and the experimental setup are shown in Table 7.10. The ICC values (ICCs>0.95) indicated high agreement between the VAR measurements using the program and the experimental setup. The MAD±SD were small (<2.6°) and the absolute difference (AD) only had the maximum value of 7.22°. Figure 7.26 and Figure 7.27 show the comparisons between the semi-automatic VAR measurements and the experimental setup from O1 and O2, respectively. Ninety-five percent of the differences fell in the range of -0.3°±6.0° and -0.1°±5.9° (mean±2SD) for O1 and O2, respectively. There was no significant relationship between the differences and the VAR angles for the validity analysis of O1 (P=0.94) and O2 (P=0.71).

Table 7.10 Agreement between the experimental setup and the semi-automatic VAR measurement using the COL method on *in-vitro* data

	O1	O2
ICC (95% CI)	0.96 (0.92-0.98)	0.96 (0.92-0.98)
MAD±SD (°)	2.6±1.6	2.5±1.6
AD (°)	0.1-7.2	0.1-6.6

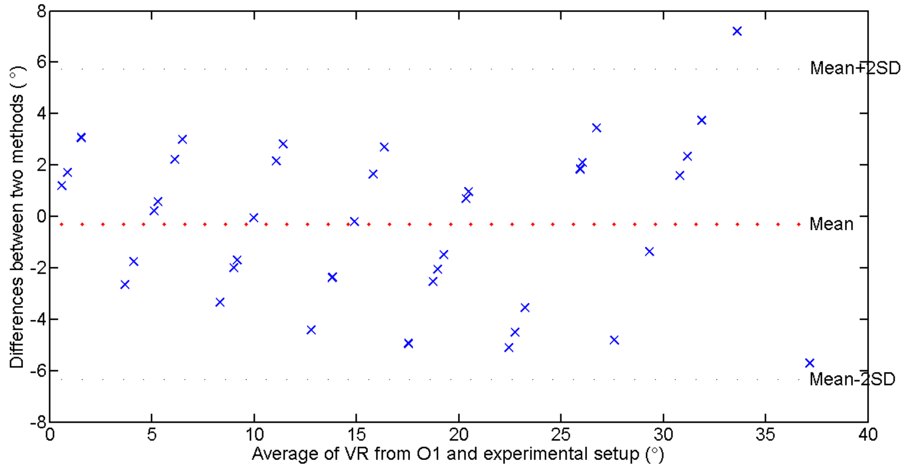


Figure 7.26 Bland-Altman plot of the comparison between the semi-automatic VAR measurement using the COL method on *in-vitro* data from O1 and the experimental setup.

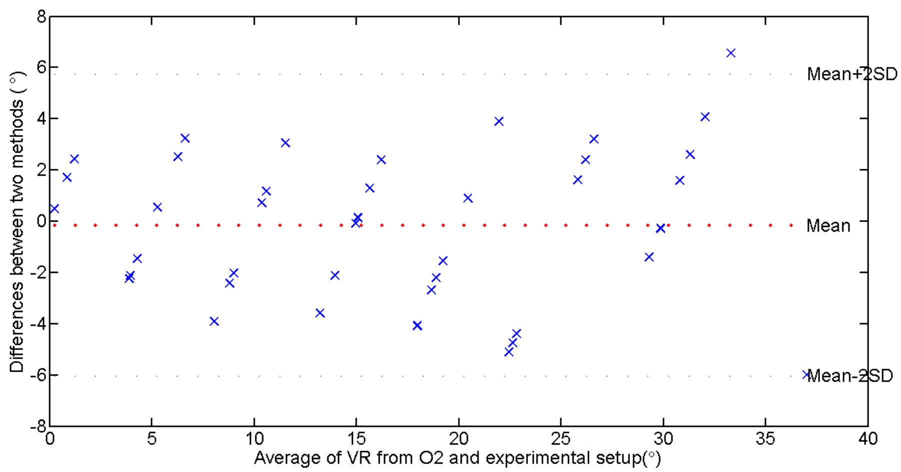


Figure 7.27 Bland-Altman plot of the comparison between the semi-automatic VAR measurement using the COL method on *in-vitro* data from O2 and the experimental setup.

7.2.8 Evaluation on *in-vivo* vertebral rotation measurement

The intra-observer reliability results of semi-automatic VAR measurements on the subjects' data are shown in Table 7.11. Both the observers showed high reliability (ICCs>0.94) with small variation (MAD±SD: 0.4°±0.4° and 0.7°±0.8° for O1 and O2, respectively).

Table 7.11 Intra-observer reliability of the COL method for VAR on *in-vivo* data using the program

	O1	O2
ICC (95% CI)	0.99 (0.98-0.99)	0.95 (0.91-0.97)
Mean±SD (°)	6.0±3.3	5.7±3.1
MAD±SD (°)	0.4±0.4	0.7±0.8

Figure 7.28 and Figure 7.29 show the Bland-Altman plot indicating the differences for the intra-observer comparisons from O1 and O2, respectively. Ninety-five percent of the differences fell in the range of 0.1°±1.1°, and 0.1°±2.1° (mean±2SD) for O1 and O2, respectively. There was no significant correlation between the differences and the VAR magnitude between the measurements from two sessions for O1 (P=0.65) and O2 (P=0.54).

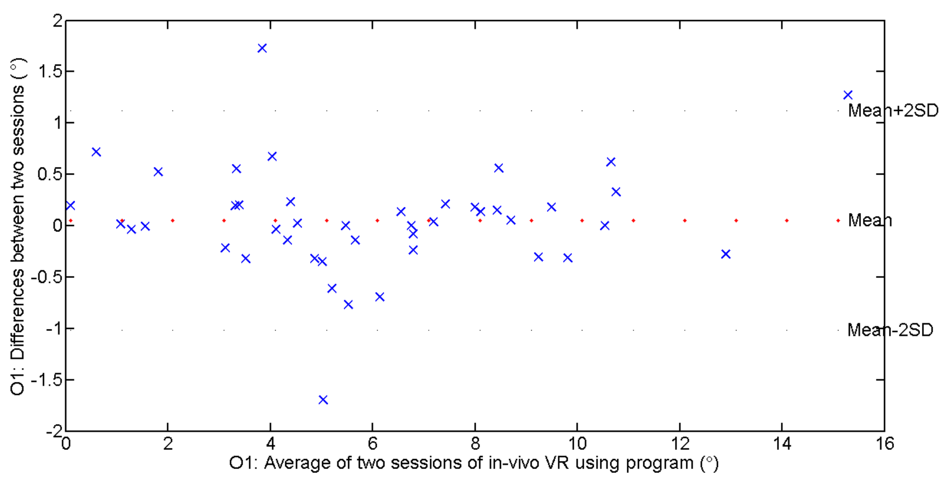


Figure 7.28 Bland-Altman plot of VAR intra-observer comparison of COL method from O1 on *in-vivo* data using the program.

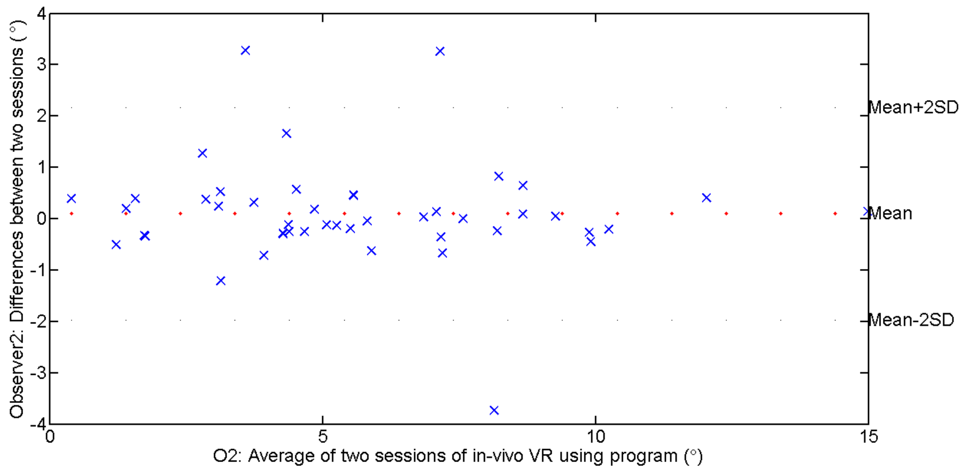


Figure 7.29 Bland-Altman plot of VAR intra-observer comparison of COL method from O2 on *in-vivo* data using the program.

The results of the inter-observer comparison of the semi-automatic VAR measurements on *in-vivo* data are shown in Table 7.12. The results (ICC: 0.97, MAD±SD: 0.5°±0.6°) indicated high inter-observer reliability between O1 and O2. Figure 7.30 shows the Bland-Altman plot of the inter-observer comparisons and ninety-five percent of the differences between O1 and O2 are within the range of 0.3°±1.5° (mean±2SD). The differences were independent of the curve magnitude (P=0.28).

Table 7.12 Inter-observer reliability of the COL method for VAR on *in-vivo* data using the program

	O12
ICC (95% CI)	0.97 (0.94-0.98)
MAD±SD (°)	0.5±0.6

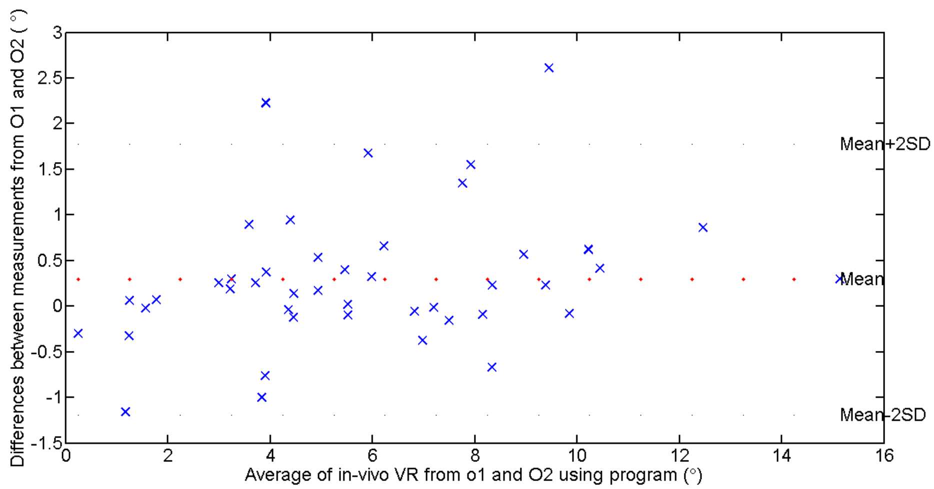


Figure 7.30 Bland-Altman plot of VAR inter-observer comparison of the COL method between O1 and O2 on *in-vivo* data using the program.

7.2.9 Discussions of the vertebral rotation investigation

Compared with the results of the manual VAR measurements on *in-vitro* and *in-vivo* data (Chapter 6.1 and 6.2), the measurements using the program had compatible intra- and inter-observer reliabilities. Besides, the program improved the inter-observer comparison on the subjects' data compared with the manual measurements. The observer with little US experience also performed well on the VAR measurements using the program. Although the semi-automatic program had good performance on the VAR measurements, it still required the operator to point the laminae. Manually pointing the laminae is the major source of the errors, which was described in 7.2.7. In order to further reduce the human errors, automatically identifying the laminae will be required.

7.3 Summary

The semi-automatic measurement program was able to export the coronal and transverse images including the lamina signals, automatically determines the centers of laminae after the operator pointed the laminae, and semi-automatically calculated the coronal curvature and vertebral rotation. Both *in-vitro* and *in-vivo* data were used to verify the program. The semi-automatic coronal curvature measurements presented compatible reliabilities with the manual measurements on phantom data and reduced the measurement errors. However, the performance of the semi-automatic measurements on the subjects' data was not as good as the performance of the manual measurements. By using the program, the lamina selection was crucial; therefore, improvement of the imaging processing might be required in order to get more accurate results. The semi-automatic VAR measurements presented compatible reliabilities with the manual measurements on both the *in-vitro* and *in-vivo* data, and improved the inter-observer reliability on the *in-vivo* data. Therefore, the program could be applied to evaluate the vertebral rotation using the COL method.

Chapter 8.
Conclusions and recommendations

8.1 Discussions and conclusions

Adolescent idiopathic scoliosis (AIS) is a 3D spinal deformity. Both the coronal curvature and the vertebral axial rotation are important parameters to assess the severity, predict the progression, and evaluate the treatment outcomes of AIS. The use of 2D radiography as the standard imaging method for scoliosis has limitations to reveal the true 3D nature of scoliosis. Also, the traditional radiographic method exposes patient to harmful ionizing radiation. Other imaging methods were studied such as computed tomography (CT), EOS, magnetic resonance imaging (MRI), surface topography, and ultrasound. CT and EOS also introduced ionizing radiation. MRI is time-consuming and expensive, which is not commonly used to diagnose AIS. Besides, CT and MRI scan patients generally in a supine posture, which affects the severity of the spinal deformity. Surface topography provides the external information of the trunk and cannot quantify the internal curvatures. US, as a non-ionizing radiation method, has the advantages of being cost-effective and portable. The literature showed that US had the potential to assess scoliosis. However, the reports were limited. The bony landmarks for assessment were not determined and the reliability and validity were not investigated for both the coronal curvature and vertebral axial rotation (VAR) measurements. Therefore, in order to reduce the exposure to ionizing radiation, the ultrasound method, was studied in this Ph.D. for the measurements of coronal curvature and VAR in children with AIS.

According to the ultrasound theory, the tissue-bone surface is a strong reflector, which makes it possible to image the posterior side of the vertebrae. Ultrasound equipment combined with the interface software was used in this project. The optimal element group for imaging and data acquisition was determined through experiments. To investigate which reflectors could be used on ultrasound images, experiments using a single cadaver vertebra and a Sawbones spinal column were performed. Three-dimensional US data was obtained and coronal, transverse, and sagittal images could be reconstructed. The spinous process (SP), transverse processes (TP) and laminae were recognized on the coronal and transverse views of the ultrasound images since these structures have relative flat surface. A non-pathological volunteer was also asked to be scanned, and results indicated that TP were more difficult to identify, especially in the thoracic region. The reason is ribs may overlap with TP. Thus, the SP and laminae were the recommended landmarks to be used in ultrasound images.

To investigate which landmark, the SP or lamina, would provide more reliable and accurate method to measure the coronal curvature, a radiographic measurement study based on the center of pedicle (COP) method and spinous process angle at apex (SPAA) method was first performed. The assumption was that pedicle and lamina are attached together and considered as a rigid body; they tilt and shift together. Results of the COP method to measure the coronal curvature showed high intra- and inter-observer reliabilities. The COP method applied on the radiographic images was then extended to the center of lamina (COL) method on the ultrasound images. Similarly, high intra- and inter-observer reliabilities were reported. The coronal curvature measurements from the COL method also agreed well with the clinical standard Cobb angle with small variation (MAD<3°). However, there were still challenges for using ultrasound to assess scoliosis. The success rate was low (32.5%) in this study because of the encoder contact issue due to the severe uneven back of the subjects and the limited scanning range of the transducer. Before applying ultrasound in clinic, a higher scanning success rate was required. Also, it was found that users who have more experience on the ultrasound measurements had higher measurement reliability, which indicated that training was needed to perform the measurements.

Vertebral axial rotation, which is another important parameter to better describe the spinal deformity, could also be measured from the transverse ultrasound image using the COL method. The VAR measurements by the COL method had high intra- and inter-observer reliabilities. In the *in-vitro* study described in this thesis, the VAR measurements by the Stokes' method from the radiographs were less accurate and reliable than the measurements by the COL method from the ultrasound images. The validity of the VAR measurements using the COL method was demonstrated in the *in-vitro* study, but not in the *in-vivo* study, which is a challenge and needs to be completed in the future.

In order to improve image quality and reduce human measurement errors, a semi-automatic program was developed. This program was able to reduce the speckle noise, semi-automatically determine the centers of laminae, and automatically calculate the coronal curvature and VAR. However, the performance of the coronal curvature measurements using the program on the *in-vivo* data was not as good as the manual measurements. It was found that the lamina area selection was crucial, which affected the measurement results; therefore, improvement of the imaging processing might be required. Also, the end vertebrae were preselected to remove this error source in the *in-vivo* study for both the manual and semi-automatic measurements. An *in-vivo* study without preselecting the end vertebrae should be performed in the future.

As a conclusion, ultrasound was able to image the spine in AIS children. The COL method developed in this thesis is a reliable and accurate method to measure the coronal curvature and VAR from the ultrasound images. The processing time starting from scanning a subject to reporting the measurements was less than 15 minutes.

8.2 Recommendations

This PhD work demonstrates ultrasound is able to reliably measure the coronal curvature and vertebral axial rotation in children with AIS. However, it was found that using the existing ultrasound equipment (TomoScan) has some limitations. The encoder may have contact issue. The uneven back surface of the patients prevents the encoder from properly contacting the surface, in which the encoder does not rotate. When the encoder is not moved, ultrasound signals are not emitted causing missing data. Furthermore, because of the dimensions of the transducer, the scanning area is limited. In some double curve cases, a single vertical scan may not cover the laminae area at the apical vertebrae on both directions of the curves. In this study, 80 AIS subjects were scanned, but only 26 data sets could be used. The success rate of obtaining valid data for processing was quite low (32.5%). To improve this, a freehand 3D ultrasound system with a built-in GPS to track the transducer is recommended. A convex transducer instead of a linear transducer should also be used because it could cover a wider scanning area. Also, a transducer with lower frequency could be used for data acquisition that it might provide stronger reflected signals from laminae since lower frequency US attenuates less when travelling through soft tissues (Bushberg *et al.*, 2002).

Coronal curvature measurements using the semi-automatic program were not as good as the manual measurements. The lamina selections by the user were crucial. Some of the edge detection methods, such as the first-order (Sobel, Roberts, and Prewwit) and second-order (Laplacian) methods (Gonzalez, *et al.*, 2009), and segmentation methods, based on the threshold (local and global) (Otsu, 1979; Taxt *et al.*, 1989) and region growth methods (Adams and Bischof, 1994), were tested to automatically find out the laminae on the coronal image. However, because of the noise from soft tissues and other bony structures, it was difficult to automatic recognize the laminae from the coronal image. To accurately distinguish the laminae from the noise is the major challenge for the measurements using program. A new method might solve this problem by automatically recognizing the vertebra surface ‘W’ shape on the transverse image. The edge detection and pattern recognition methods can be used to identify the laminae and the noise from the soft tissues and other bony structures could be reduced. Then the projected coronal image could be formed from the processed data for the coronal curvature measurements. After that, image processing is needed to obtain better performance of the program.

1) The visualization of the image can be improved by adjusting the image contrast, modifying the intensity range and applying different filters. Interpolation method (Rohling *et al.*, 1999) or hole filling method (Purnama *et al.*, 2010) can be used to fill the gap of the missing data, where no information could be obtained from the vertebrae.

2) The lamina segmentation is important, as it relates to the measurement reliability. Image segmentation methods to separate the laminae from background, such as the histogram based method (Yanowitz and Bruckstein, 1989) and region growing method (Adams and Bischof, 1994) should be tested and combined with the thresholding method applied in this thesis.

3) The semi-automatic measurement program described in Chapter 7 still relies on users to select the laminae. Human errors may be introduced. A fully automatic measurement based on the edge detection and pattern recognition will reduce human errors further.

The ultrasound provides 3D information of the spine. In this thesis, only the coronal and transverse images were applied for the assessment of the spinal deformity. From the sagittal image, the lamina could also be identified, and the trend of laminae could be able to provide supplementary information such as the kyphosis or lordosis for the evaluation of AIS. Besides the laminae, the potential to use other bony landmarks, such as the SP and TP, should also be considered in the measurement strategy.

8.3 Summary

In conclusion, the key points from this project were:

1) The proposed ultrasound method could image the spine in 3D in AIS children such that images from the coronal, transverse, and sagittal planes could be reconstructed. A low success rate (32.5%) for *in-vivo* data acquisition was obtained using the present equipment. To improve the success rate, a GPS build-in the ultrasound system should be used to prevent the contact issues that affected data acquisition in the present study.

2) A new measurement method, the center of lamina (COL) method, was proposed on US images in this thesis. The COL method was demonstrated to be a reliable and accurate method for the measurements of both the coronal curvature and VAR. However, the coronal curvature *in-vivo* study was performed with the end vertebrae pre-selected before the measurements, which removed the error source related to selecting the end vertebrae. Therefore, an *in-vivo* study without knowing the end vertebrae should be conducted in the future.

3) A semi-automatic program was developed to perform the measurements of coronal curvature and VAR using the COL method. The processing time, starting from scanning a subject to reporting the measurements, was less than 15 minutes which is sufficiently short to possibly allow clinical implementation. The program was a promising tool to measure the VAR and reduced human errors. However, the coronal curvature measurements need to be further improved. A new method, which could accurately identify the laminae from noise, is needed to reduce the variations when calculating the coronal curvature using the COL method.

References

Aaro S, Dahlborn M (1981) Estimation of vertebral rotation and the spinal and rib cage deformity in scoliosis by computer tomography. Spine 6:460-467

Achim A, Bezerianos A, Tsakalides P (2001) Wavelet-based ultrasound image denoising using an alpha-stable prior probability model. Conf Proc IEEE Int Conf Image Process:221-224

Adams R, Bischof L (1994) Seeded Region Growing. IEEE Trans. Pattern Anal. Machine Intell. 16:641-647

Asghar J, Samdani AF, Pahys JM, D'Andrea LP, Guille JT, Clements DH, Betz RR, Grp HS (2009) Computed tomography evaluation of rotation correction in adolescent idiopathic scoliosis: a comparison of an all pedicle screw construct versus a hook-rod system. Spine 34:804-807

Bayley E, Zia Z, Kerslake R, Boszczyk B (2010) The ipsilateral lamina-pedicle angle: can it be used to guide pedicle screw placement in the sub-axial cervical spine? Eur Spine J 19:458-463

Birchall D, Hughes DG, Hindle J, Robinson L, Williamson JB (1997) Measurement of vertebral rotation in adolescent idiopathic scoliosis using three-dimensional magnetic resonance imaging. Spine 22:2403-2407

Birchall D, Hughes D, Gregson B, Williamson B (2005) Demonstration of vertebral and disc mechanical torsion in adolescent idiopathic scoliosis using three-dimensional MR imaging. Eur Spine J 14:123-129

Bland MJ, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327:307-310

Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. Stat Methods Med Res 8:135-160

Bland JM, Altman DG (2010) Statistical methods for assessing agreement between two methods of clinical measurement. Int J Nurs Stud 47:931-936

Brenner DJ, Hall EJ (2007) Computed tomography -- an increasing source of radiation exposure. N Engl J Med 357:2277-2284

Bunnell WP (1988) The natural history of idiopathic scoliosis. Clin Orthop Relat Res:20-25

Bushberg J, Seibert J, Leidholdt E (2002) The essential physics of medical imaging, 2nd edn. Lippincott Williams & Wilkins, Philadelphia, pp 469-553

Carman DL, Browne RH, Birch JG (1990) Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation. J Bone Joint Surg Am 72 (3):328-333

Chen W, Lou EHM, Le LH (2011) Using ultrasound imaging to identify landmarks in vertebra models to assess spinal deformity. Conf Proc IEEE Eng Med Biol Soc:8495-8498

Chen W, Le L, Lou E (2012) Ultrasound imaging of spinal vertebrae to study scoliosis. Open J Acoust 2:95-103

Cheung C-WJ, Zheng Y, Lim CT, Goh JCH (2010) Development of 3-D ultrasound system for assessment of adolescent idiopathic scoliosis (AIS). World Congr Biomech:584-587.

Chi WM, Cheng CW, Yeh WC, Chuang SC, Chang TS, Chen JH (2006) Vertebral axial rotation measurement method. Comput Methods Programs Biomed 81:8-17

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

Cobb J (1948) Outline for the study of scoliosis. Instr Course Lect 5:261-275

Coupe P, Hellier P, Kervrann C, Barillot C (2009) Nonlocal means-based speckle filtering for ultrasound images. IEEE Trans Image Process 18:2221-2229

Currier DP (1990) Elements of research in physical therapy, 3rd ed. Williams & Wilkins, Baltimore, pp 160-171

Daruwalla JS, Balasubramaniam P (1985) Moire topography in scoliosis. Its accuracy in detecting the site and size of the curve. J Bone Joint Surg Br 67:211-213

Deacon P, Flood BM, Dickson RA (1984) Idiopathic scoliosis in three dimensions. A radiographic and morphometric analysis. J Bone Joint Surg Br 66:509-512

Deschenes S, Charron G, Beaudoin G, Labelle H, Dubois J, Miron M-C, Parent S (2010) Diagnostic imaging of spinal deformities: reducing patients radiation dose with a new slot-scanning X-ray imager. Spine 35:989-994

Do T, Fras C, Burke S, Widmann RF, Rawlins B, Boachie-Adjei O (2001) Clinical value of routine preoperative magnetic resonance imaging in adolescent idiopathic scoliosis. A prospective study of three hundred and twenty-seven patients. J Bone Joint Surg Br 83A:577-579

Donoho DL, Johnstone IM (1995) Adapting to unknown smoothness via wavelet shrinkage. J Amer Statist Assoc 90:1200-1224

Doody MM, Lonstein JE, Stovall M, Hacker DG, Luckyanov N, Land CE (2000) Breast cancer mortality after diagnostic radiography: findings from the U.S. Scoliosis Cohort Study. Spine 25:2052-2063

Douglas G. Orndorff, MD, Morgan A. Scott, Katie A. Patty, MS (2012) Force transfer in the spine. J Spinal Research Found 7: 30-35

Drerup B (1985) Improvements in measuring vertebral rotation from the projections of the pedicles. J Biomech 18:369-378

Easwar TR, Hong JY, Yang JH, Suh SW, Modi HN (2011) Does lateral vertebral translation correspond to Cobb angle and relate in the same way to axial vertebral rotation and rib hump index? A radiographic analysis on idiopathic scoliosis. Eur Spine J 20:1095-1105

Furness G, Reilly MP, Kuchi S (2002) An evaluation of ultrasound imaging for identification of lumbar intervertebral level. Anaesthesia 57:277-280

Glaser DA, Doan J, Newton PO (2012) Comparison of 3-dimensional spinal reconstruction accuracy: biplanar radiographs with EOS versus computed tomography. Spine 37:1391-1397

Gocen S, Aksu MG, Baktiroglu L, Ozcan O (1998) Evaluation of computed tomographic methods to measure vertebral rotation in adolescent idiopathic scoliosis: an intraobserver and interobserver analysis. J Spinal Disord 11:210-214

Goldberg CJ, Kaliszer M, Moore DP, Fogarty EE, Dowling FE (2001) Surface topography, Cobb angles, and cosmetic change in scoliosis. Spine 26:E55-63

Gonzalez R, Woods R, Eddins S (2009) Digital image processing using MATLAB, 2nd edn., Gatesmark, LLC, Natick, pp 109-127, 557-571

Gstoettner M, Sekyra K, Walochnik N, Winter P, Wachter R, Bach CM (2007) Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools. Eur Spine J 16:1587-1592

Gstoettner M, Lechner R, Glodny B, Thaler M, Bach CM (2011) Inter- and intraobserver reliability assessment of computed tomographic 3D measurement of pedicles in scoliosis and size matching with pedicle screws. Eur Spine J 20:1771-1779

Guo Y, Cheng HD, Tian J, Zhang Y (2009) A novel approach to speckle reduction in ultrasound imaging. Ultrasound Med Biol 35:628-640

Gupta N, Swamy MNS, Plotkin E (2005) Despeckling of medical ultrasound images using data and rate adaptive lossy compression. IEEE Trans Med Imaging 24:743-754

Helenius I, Remes V, Yrjonen T, Ylikoski M, Schlenzka D, Helenius M, Poussa M (2003) Harrington and Cotrel-Dubousset instrumentation in adolescent idiopathic scoliosis. Long-term functional and radiographic outcomes. J Bone Joint Surg Am 85A:2303-2309

Heo DH, Cho YJ, Sheen SH, Hong MS, Cho SM, Park SH (2010) 3D reconstructions of spinal segmental arteries using CT angiography: applications in minimally invasive spinal procedures. AJNR Am J Neuroradiol 31:1635-1639

Herzenberg JE, Waanders NA, Closkey RF, Schultz AB, Hensinger RN (1990) Cobb angle versus spinous process angle in adolescent idiopathic scoliosis. The relationship of the anterior and posterior deformities. Spine 15:874-879

Ho EK, Upadhyay SS, Ferris L, Chan FL, Bacon-Shone J, Hsu LC, Leong JC (1992) A comparative study of computed tomographic and plain radiographic methods to measure vertebral rotation in adolescent idiopathic scoliosis. Spine 17:771-774

Ho EK, Upadhyay SS, Chan FL, Hsu LC, Leong JC (1993) New methods of measuring vertebral rotation from computed tomographic scans. An intraobserver and interobserver study on girls with scoliosis. Spine 18:1173-1177

Hoffman DA, Lonstein JE, Morin MM, Visscher W, Harris BS, 3rd, Boice JD, Jr. (1989) Breast cancer in women with scoliosis exposed to multiple diagnostic x rays. J Natl Cancer Inst 81:1307-1312

Hong JY, Suh SW, Easwar TR, Modi HN, Yang JH, Park JH (2011) Evaluation of the three-dimensional deformities in scoliosis surgery with computed tomography: efficacy and relationship with clinical outcomes. Spine 36:E1259-1265

Humbert L, De Guise JA, Aubert B, Godbout B, Skalli W (2009) 3D reconstruction of the spine from biplanar X-rays using parametric models based on transversal and longitudinal inferences. Med Eng Phys 31:681-687

Ilharreborde B, Steffen JS, Nectoux E, Vital JM, Mazda K, Skalli W, Obeid I (2011) Angle measurement reproducibility using EOS three-dimensional reconstructions in adolescent idiopathic scoliosis treated by posterior instrumentation. Spine 36:E1306-E1313

Illes T, Tunyogi-Csapo M, Somoskeoy S (2011) Breakthrough in three-dimensional scoliosis diagnosis: significance of horizontal plane view and vertebra vectors. Eur Spine J 20:135-143

Katz DE, Richards BS, Browne RH, Herring JA (1997) A comparison between the Boston brace and the Charleston bending brace in adolescent idiopathic scoliosis. Spine 22:1302-1312

Knott P, Mardjetko S, Nance D, Dunn M (2006) Electromagnetic topographical technique of curve evaluation for adolescent idiopathic scoliosis. Spine 31:E911-915

Kuklo TR, Potter BK, Lenke LG (2005) Vertebral rotation and thoracic torsion in adolescent idiopathic scoliosis: what is the best radiographic correlate? J Spinal Disord Tech 18:139-147

Kuklo TR, Potter BK, O'Brien MF, Schroeder TM, Lenke LG, Polly DW, Jr. (2005) Reliability analysis for digital adolescent idiopathic scoliosis measurements. J Spinal Disord Tech 18:152-159

Kuklo TR, Potter BK, Polly DW, Jr., O'Brien MF, Schroeder TM, Lenke LG (2005) Reliability analysis for manual adolescent idiopathic scoliosis measurements. Spine 30:444-454

Kuraishi S, Takahashi J, Hirabayashi H, Hashidate H, Ogihara N, Mukaiyama K, Kato H (2013) Pedicle morphology using computed tomography-based navigation system in adolescent idiopathic scoliosis. J Spinal Disord Tech 26:22-28

Lam GC, Hill DL, Le LH, Raso JV, Lou EH (2008) Vertebral rotation measurement: a summary and comparison of common radiographic and CT methods. Scoliosis 3:16

Laurent A, Mistretta F, Bottigioli D, Dahel K, Goujon C, Nicolas JF, Hennino A, Laurent PE (2007) Echographic measurement of skin thickness in adults by high frequency ultrasound to assess the appropriate microneedle length for intradermal delivery of vaccines. Vaccine 25:6423-6430

Le LH, Gu YJ, Li YP, Zhang C (2010) Probing long bones with ultrasonic body waves. Appl Phys Lett 96:114102

Lenke LG (2005) Lenke classification system of adolescent idiopathic scoliosis: treatment recommendations. Instr Course Lect 54:537-542

Li M, Cheng J, Ying M, Ng B, Zheng YP, Lam TP, Wong WY, Wong MS (2010) Application of 3-D ultrasound in assisting the fitting procedure of spinal orthosis to patients with adolescent idiopathic scoliosis. Stud Health Technol Inform 158:34-37

Li M, Cheng J, Ying M, Ng B, Zheng Y, Lam T, Wong W, Wong M (2012) Could clinical ultrasound improve the fitting of spinal orthosis for the patients with AIS? Eur Spine J 21 (10):1926-1935

McGraw KO, Wong SP (1996) Forming inferences about some intraclass correlation coefficients. Psychol Methods 1:30-46

McKenna C, Wade R, Faria R, Yang H, Stirk L, Gummerson N, Sculpher M, Woolacott N (2012) EOS 2D/3D X-ray imaging system: a systematic review and economic evaluation. Health Technol Assess 16:1-188

Mehta SS, Modi HN, Srinivasalu S, Chen T, Suh SW, Yang JH, Song HR (2009) Interobserver and intraobserver reliability of Cobb angle measurement: endplate versus pedicle as bony landmarks for measurement: a statistical analysis. J Pediatr Orthop 29:749-754

Mok JM, Berven SH, Diab M, Hackbarth M, Hu SS, Deviren V (2008) Comparison of observer variation in conventional and three digital radiographic methods used in the evaluation of patients with adolescent idiopathic scoliosis. Spine 33:681-686

Moran CM, Bush NL, Bamber JC (1995) Ultrasonic propagation properties of excised human skin. Ultrasound Med Biol 21:1177-1190

Morrissy RT, Goldsmith GS, Hall EC, Kehl D, Cowie GH (1990) Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error. J Bone Joint Surg Am 72A:320-327

Morrissy RT, Weinstein SL (2006) Lovell & Winter's pediatric orthopedics. 6th edn. Lippincott Williams & Wilkins, Philadelphia, pp 694-762

Nakahara D, Yonezawa I, Kobanawa K, Sakoda J, Nojiri H, Kamano S, Okuda T, Kurosawa H (2011) Magnetic resonance imaging evaluation of patients with idiopathic scoliosis A prospective study of four hundred seventy-two outpatients. Spine 36:E482-E485

Nash CL Jr, Moe JH (1969) A study of vertebral rotation. J Bone Joint Surg Am 51:223-229

Oda M, Rauh S, Gregory PB, Silverman FN, Bleck EE (1982) The significance of roentgenographic measurement in scoliosis. J Pediatr Orthop 2:378-382

Otsu N (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst, Man, Cyber 9:62-66

Oxborrow NJ (2000) Assessing the child with scoliosis: the role of surface topography. Arch Dis Child 83:453-455

Parent E, Watkins E, Emrani M, Hill D (2010) Differences in measures of full-torso surface topography among healthy teenagers are independent of growth indicators. Scoliosis 5:O9

Parisini P, Lolli F, Greggi T, Di Silvestre M, Cioni A, Giacomini S, Bakaloudis G (2006) An innovative diagnostic procedure of vertebral deformities without exposure to x-rays. Stud Health Technol Inform 123:527-532

Pazos V, Cheriet F, Song L, Labelle H, Dansereau J (2005) Accuracy assessment of human trunk surface 3D reconstructions from an optical digitising system. Med Biol Eng Comput 43:11-15

Perdriolle R, Vidal J (1985) Thoracic idiopathic scoliosis curve evaluation and prognosis. Spine 10:785-791

Purnama IKE, Wilkinson MHF, Veldhuizen AG, van Ooijen PMA, Sardjono TA, Lubbers J, Verkerke GJ (2009) Following scoliosis progression in the spine using ultrasound imaging. IFMBE Proc:600-602

Purnama KE, Wilkinson MHF, Veldhuizen AG, Ooijen PMAv, Lubbers J, Burgerhof JGM, Sardjono TA, Verkerke GJ (2010) A framework for human spine imaging using a freehand 3D ultrasound system. Technol Health Care 18:1-17

Richards BS (1992) Measurement error in assessment of vertebral rotation using the Perdriolle torsionmeter. Spine 17:513-517

Rohling R, Gee A, Berman L (1999) A comparison of freehand three-dimensional ultrasound reconstruction techniques. Med Image Anal. 3:339-359

Samei E, Lo JY, Yoshizumi TT, Jesneck JL, Dobbins JT, 3rd, Floyd CE, Jr., McAdams HP, Ravin CE (2005) Comparative scatter and dose performance of slot-scan and full-field digital chest radiography systems. Radiology 235:940-949

Schmitz A, Jaeger UE, Koenig R, Kandyba J, Wagner UA, Giesecke J, Schmitt O (2001) A new MRI technique for imaging scoliosis in the sagittal plane. Eur Spine J 10:114-117

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

Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86:420-428

Song L, Bourassa Y, Beauchamp D, Lemelin G (2002) Optical 3D digitizer, system and method for digitizing an object. US Patent 6,493,095

Srinivasalu S, Modi HN, Smehta S, Suh S-W, Chen T, Murun T (2008) Cobb angle measurement of scoliosis using computer measurement of digitally acquired radiographs. Intraobserver and interobserver variability. Asian Spine J 2:90-93

Stokes IA, Bigalow LC, Moreland MS (1986) Measurement of axial rotation of vertebrae in scoliosis. Spine 11:213-218

Stokes IA (1989) Axial rotation component of thoracic scoliosis. J Orthop Res 7:702-708

Stokes IA, Aronsson DD (2002) Identifying sources of variability in scoliosis classification using a rule-based automated algorithm. Spine 27:2801-2805

Stokes IA, Aronsson DD (2006) Computer-assisted algorithms improve reliability of King classification and Cobb angle measurement of scoliosis. Spine 31:665-670

Stratford P (2004) Getting more from the literature: Estimating the standard error of measurement from reliability studies. Physiother Can 56:27-30

Sudha S, Suresh GR, Sukanesh R (2009) Speckle noise reduction in ultrasound images by wavelet thresholding based on weighted variance. Int J Comput Theor Eng 1:7-12

Suzuki S, Yamamuro T, Shikata J, Shimizu K, Iida H (1989) Ultrasound measurement of vertebral rotation in idiopathic scoliosis. J Bone Joint Surg Br 71 (2):252-255

Tan KJ, Moe MM, Vaithinathan R, Wong HK (2009) Curve progression in idiopathic scoliosis: follow-up study to skeletal maturity. Spine 34:697-700

Taxt T, Flynn PJ, Jain AK (1989) Segmentation of Document Images. IEEE Trans. Pattern Anal. Machine Intell. 11:1322-1329

Thomsen M, Abel R (2006) Imaging in scoliosis from the orthopaedic surgeon's point of view. Eur J Radiol 58:41-47

Torell G, Nachemson A, Haderspeck-Grib K, Schultz A (1985) Standing and supine Cobb measures in girls with idiopathic scoliosis. Spine 10:425-427

Turner-Smith AR, Harris JD, Houghton GR, Jefferson RJ (1988) A method for analysis of back shape in scoliosis. J Biomech 21:497-509

Van Goethem J, Van Campenhout A, van den Hauwe L, Parizel PM (2007) Scoliosis. Neuroimaging Clin N Am 17:105-115.

Vanithamani R, Umamaheswari G (2011) Wavelet based Despeckling of Medical Ultrasound Images with Bilateral filter. IEEE Region 10 Conference TENCON:389-393

Vrtovec T, Pernus F, Likar B (2009) A review of methods for quantitative evaluation of axial vertebral rotation. Eur Spine J 18:1079-1090

Weinstein SL, Dolan LA, Wright JG, Dobbs MB (2013) Effects of Bracing in Adolescents with Idiopathic Scoliosis. N Engl J Med 369:1512-1521

Weiss HR (1995) Measurement of vertebral rotation: Perdriolle versus Raimondi. Eur Spine J 4:34-38

Weisz I, Jefferson RJ, Turner-Smith AR, Houghton GR, Harris JD (1988) ISIS scanning: a useful assessment technique in the management of scoliosis. Spine 13:405-408Yazici M, Acaroglu ER, Alanay A, Deviren V, Cila A, Surat A (2001) Measurement of vertebral rotation in standing versus supine position in adolescent idiopathic scoliosis. J Pediatr Orthop 21:252-256

Yanowitz SD, Bruckstein AM (1989) A New Method for Image Segmentation. Comput Vision Graph 46:82-95

Zhang C, Le LH, Zheng R, Ta D, Lou E (2011) Measurements of ultrasonic phase velocities and attenuation of slow waves in cellular aluminum foams as cancellous bone-mimicking phantoms. J Acoust Soc Am 129:3317-3326

Zhang JH, Lou E, Le LH, Hill DL, Raso JV, Wang YY (2009) Automatic Cobb measurement of scoliosis based on fuzzy Hough transform with vertebral shape prior. J Digit Imaging 22 (5):463-472

Zhang JH, Lou E, Shi XL, Wang YY, Hill DL, Raso JV, Le LH, Lv LA (2010) A computer-aided Cobb angle measurement method and its reliability. J Spinal Disord Tech 23:383-387

Zhang J, Lou E, Hill D, Raso J, Wang Y, Le L, Shi X (2010) Computer-aided assessment of scoliosis on posteroanterior radiographs. Med Biol Eng Comput 48:185-195

Zheng R, Le LH, Sacchi MD, Ta D, Lou E (2007) Spectral ratio method to estimate broadband ultrasound attenuation of cortical bones in vitro using multiple reflections. Phys Med Biol 52:5855-5869