



A Systematic Review of the Characteristics of Ultrasound Measurements in the Evaluation of Spinal Curvature in Scoliosis Patients

ARTICLE INFO

Article Type

Systematic Review

Authors

Saberi M.^{1*} MSc

Rajabi R.¹ PhD

Minoonejad H.¹ PhD

Karimizadeh Ardakani M.¹ PhD

How to cite this article

Saberi M, Rajabi R, Minoonejad H, Karimizadeh Ardakani M. A Systematic Review of the Characteristics of Ultrasound Measurements in the Evaluation of Spinal Curvature in Scoliosis Patients. Iranian Journal of War & Public Health. 2023;15(2):181-190.

¹Department of Health and Sport Medicine, Faculty of Sport Sciences and Health, University of Tehran, Tehran, Iran

*Correspondence

Address: Department of Health and Sport Medicine, University of Tehran, Kargare Shomali Street, Tehran, Iran.
Postal Code: 88351730-37
Phone: +98 (17) 32156816
Fax: +98 (21) 88021527
mehdi.saberi@ut.ac.ir

Article History

Received: January 29, 2023

Accepted: April 7, 2023

ePublished: April 30, 2023

ABSTRACT

Aims Scoliosis is one of the secondary complications in amputation veterans, which requires periodic evaluations. The gold standard for evaluating the spine is radiography, which has caused many concerns due to the high doses of X-ray radiation. For this reason, therapists are finding for an alternative method for this tool. Ultrasound is a non-invasive tool with different methods for this purpose. The current research aimed to determine the best method for measuring the curvature of the spine of scoliosis patients by ultrasound.

Information & Methods A systematic search was performed from 2000 to 2022 in Scopus, PubMed, Web of Science, and Embase databases to find articles that examined the reliability and validity of ultrasound to assess the curvature of the spine in people with scoliosis. COSMIN checklist scale was used to survey the quality of researches.

Findings 18 articles were reviewed. In most studies, correct methodological quality was not reported to evaluate internal consistency, measurement error, construct validity, hypothesis testing, cross-cultural validity, responsiveness, and interpretability. Finally, four methods of measuring spinous process angle, spinous process-transverse process angle, transverse process-upper articular process angle, and lamina center method were found for ultrasound.

Conclusion Ultrasound by measuring the center of the lamina is a simple and very reliable method to evaluate the curvatures of the spine of people with scoliosis and can be a suitable alternative to the gold standard of radiography.

Keywords Reliability and Validity; Spinal Curvature; Ultrasound Imaging; Scoliosis; Amputee; Veterans

CITATION LINKS

[1] Scoliosis imaging: What radiologists should ... [2] Cancer mortality among women frequently exposed to radiographic examinations for ... [3] Freehand three-dimensional ultrasound system for assessment of ... [4] Could clinical ultrasound improve the fitting of ... [5] Improvement on the accuracy and reliability of ultrasound coronal ... [6] A reliability and validity study for different coronal angles using ultrasound imaging in ... [7] Evaluation of scoliosis with a commercially available ultrasound ... [8] Reliability of assessing the coronal curvature of children with scoliosis by using ... [9] Factors influencing spinal curvature measurements on ultrasound images ... [10] Designing and constructing instruments for social research ... [11] Measurement in nursing and health ... [12] Quality criteria were proposed for measurement properties ... [13] Psychometric properties of chronic pain acceptance ... [14] Methodological quality of studies on the measurement properties of neck pain and ... [15] The COSMIN checklist for assessing the methodological ... [16] The consensus-based standards for the selection of health ... [17] Rating the methodological quality in systematic reviews of studies on ... [18] Updated method guidelines for systematic reviews in the Cochrane back review ... [19] Updated method guidelines for systematic reviews in the Cochrane collaboration ... [20] Reliability and accuracy of ultrasound measurements with and without the aid of ... [21] Intra- and inter-rater reliability of coronal curvature measurement for adolescent ... [22] A reliability and validity study for Scolioscan: A radiation-free scoliosis ... [23] A novel method to measure the sagittal curvature in spinal deformities ... [24] Reliability of the axial vertebral rotation measurements of adolescent ... [25] Intra and inter-rater reliability of spinal flexibility ... [26] Intra- and interrater reliability of Cobb angle measurements ... [27] A preliminary study of estimation of Cobb's angle ... [28] Reliability and validity study of clinical ultrasound ... [29] Development of 3D ultrasound system for assessment of adolescent ... [30] Development of 3-D ultrasound system for assessment ... [31] Translation, development and psychometric properties of health ... [32] Engineering a freehand 3D ultrasound ... [33] The COSMIN study reached international consensus on ... [34] Reliability and validity of lateral curvature ... [35] Ultrasound volume projection imaging ... [36] Three-dimensional spinal curvature in ... [37] Cobb angle versus spinous process angle in adolescent ... [38] Functional anatomy of the ...

Introduction

Scoliosis is one of the complex deformities known as a three-dimensional deformity with developmental deformation of the spine and trunk. Also, in the general body changes of these people, a lateral deviation in the frontal plane, disturbance in the physiological curves of the spine in the sagittal plane, and axial rotation of the vertebrae in the transverse plane are observed [1]. Measuring spinal alignment is an essential aspect of musculoskeletal assessment that helps clinicians adequately screen for scoliosis, determine baseline data, monitor progress, and guide appropriate implementation of treatment strategies. The current gold standard for measuring scoliosis is radiography, a method that provides the Cobb angle and is performed an average of 22.9 times for an immature individual during the entire treatment and follow-up period, but it has significant limitations. Access to this method requires a lot of time and money, and on the other hand, it increases the risk of cancer in people and has caused increasing concern among therapists [2].

To reduce or eliminate the potential risk of radiation, various non-invasive or minimally invasive methods have been developed, such as Magnetic Resonance Imaging (MRI), scoliometry, ultrasound, topography, and the EOS 3D imaging system. Among these approaches, ultrasound is noninvasive, low-cost, and low-risk, which can provide an image of the spine and allow screening and monitoring of scoliosis in the standing or supine position during treatment and follow-up evaluations [3].

In this context, a series of studies have used different methods to evaluate the spine curves of scoliosis patients through ultrasound. Li *et al.* evaluated the angle of the spinous process using the three-dimensional clinical ultrasound method and used this method as a clinical parameter to estimate Cobb's angle. As a result, they reported a high intra-rater reliability for ultrasound imaging and also showed a high validity between this device and the radiographic Cobb angle [4]. Zheng *et al.* reported a moderate to strong correlation between the ultrasound angle and the Cobb angle for both chest and waist areas using 3D ultrasound imaging without radiation and the spinous process angle method [5]. Brink *et al.* used two methods of spinous process angle and transverse angle, to measure the coronal angles of the main thoracic and lumbar curves. Intra-class correlation coefficients for intra- and inter-rater reliability were obtained as 0.97 and 0.95, respectively. Also, both coronary ultrasound angles showed a very good linear correlation with Cobb's angles [6]. LV *et al.* introduced a new ultrasound system to replace radiography. For evaluation, they used the angle method of the spinous process and upper articular process. They obtained very high inter-rater and intra-rater reliability for ultrasound imaging system and high correlation and validity

between ultrasound and radiography methods [7]. Chen *et al.* evaluated the reliability of coronal curvature in children with scoliosis using the lamina center method on ultrasound images. As a result, the lamina center method showed high intra- and inter-observer reliability. Also, the measured difference between Cobb and lamina center methods was 0.5 ± 0.7 [8]. Zheng *et al.* used lamina center methods and standing radiographs for this study. The intra-rater reliability was obtained at the level of 0.96. Ultrasound and radiography measurements showed a high correlation coefficient [9].

As you have noticed, many of these researches based on ultrasound tools have obtained reliability and validity, but they have not fully reported all the measurement features or did not have strong measurement features. While for designing or using a tool, it is very important to pay attention to the quality of the tool's measurement features. The two main features of measuring tools are validity and reliability [10]. Some experts in the field of tools consider the third feature necessary for measuring features. This feature is called the ability to respond to changes [11]. Validity itself has three main parts, which include content, structural and formal validity. Some experts in the field of tools consider the validity of criteria as a part of validity, which should be checked if possible. The reliability of the instrument includes internal correlation, consistency, and equivalence. Some researchers also mention other criteria as measurement characteristics of the tool, which include repeatability, ceiling and floor effect, and interpretability of the results [12]. Studies show that in most of the health measurement tools that are designed or used, researchers have only focused on validity and reliability and have not considered other criteria of measurement characteristics. Reneman *et al.*, in a review study on chronic pain acceptance measurement tools, concluded that none of the tools had all the criteria of measurement characteristics [13]. Terwee *et al.* reviewed 47 studies that used neck pain questionnaires and concluded that in these questionnaires, appropriate statistical analyzes were used to measure reliability, construct validity, and standard error of measurement, but in terms of how to measure other criteria, they need improvement [14].

In Iran, no review study examines the quality of health tools, especially tools for measuring the curvature of the spine of people with scoliosis. On the other hand, most of the studies that evaluated the curvature of the scoliosis spine with ultrasound used different methods, which cause confusion in choosing the best method to evaluate these people. Ultrasound is useful in evaluating the spinal curvatures of scoliosis people and choosing the best measurement method to familiarize researchers, therapists, and reformers with the strengths and weaknesses of these methods and help them to use a method with

better measurement characteristics. Therefore, the current research aimed to determine the best method for measuring the curvature of the spine of scoliosis patients by ultrasound.

Information and Methods

Search strategy

A systematic search was performed from 2000 to 2022 in Scopus, PubMed, Web of Science, and Embase databases to find sources in four subject areas. The keywords for this search included "Scoliosis" and "Coronal Curvature" and "Lateral Curvature" and "Spinal Disorder" and "Spinal Asymmetry" OR "Ultrasound" and "Ultrasonography" OR "Measurement Properties" and "Validity and Reliability" and "Sensitivity" and "Responsiveness" OR "Assessment" and "Measure" and "Evaluation".

Extracting the data

Out of 133 articles, 18 most up-to-date and relevant articles were selected and analyzed. Then, by fully studying the methods and results of these articles, based on the inclusion and exclusion criteria of the studies, a series of articles were selected for systematic review, and the rest were excluded. Two reviewers independently assessed each study. After that, a third referee confirmed the results of these articles. Finally, the final articles were entered into the initial list in the Endnote 20 software.

Inclusion criteria

- Studies involving probability sampling.
- Studies that have sufficient data to calculate the measure.
- Studies whose findings are published as full articles.
- Studies that included people with scoliosis who did not have other obvious musculoskeletal abnormalities.
- Studies that have investigated the validity and reliability of ultrasound on the curvature of the spine in scoliosis people in an interventional way.

Exclusion criteria

- Studies that have been published in various report formats other than articles.
- Review studies of the last three years
- Qualitative, review, comparative, and interventional studies
- Studies that have inconsistencies in data analysis and small sample sizes.
- Articles published in languages other than English and Persian.

Quality assessment

To determine whether the results of the included studies can be trusted, the methodological quality of the studies was assessed. To check the quality of the research included in the systematic review, the COSMIN (COnsensus-based Standards for the selection of health status Measurement INstruments) checklist scale was used, which has high validity and reliability and can be cited in the research. This checklist can be used to assess the methodological

quality of studies related to the measurement properties of health status measurement tools. The goal of COSMIN's approach is to improve the selection of measurement tools and facilitates the selection of the most appropriate index among different tools. For example, this checklist can be used to evaluate the quality of a study on a measurement tool or to compare the measurement characteristics of several tools. In this evaluation, it is important to consider the methodological quality of selected studies. If the results of high-quality studies differ from the results of low-quality studies, this is a sign of bias. This checklist has 12 boxes. Ten boxes can be used to assess whether a study meets the criteria of good methodological quality. Nine boxes contain standards and criteria for measurement characteristics, including internal consistency (box A), reliability (box B), measurement error (box C), content validity (box D), construct validity (box E), hypothesis testing (box F), inter-validity cultural (box G), criterion validity (box H) and responsiveness (box I), and a box contains criteria for studies in the field of interpretability (box J) [15, 16].

COSMIN scoring system

Four response options are defined for each item of the COSMIN checklist, indicating excellent, good, moderate, and poor methodological quality. After that, the methodological quality score in each box is obtained by taking the lowest rating from each of the items in a box (the worst score is counted). For example, if an item has a poor score in the reliability box, the methodological quality of the reliability assessment in that study is evaluated as poor. Therefore, a poor score in any case is considered a fatal flaw [17].

Levels of evidence

To summarize all the evidence on the measurement properties of the different methods, we synthesized the different studies by combining their results, taking the number and methodological quality of the studies and the consistency of their results into account. The possible overall rating for a measurement property is "positive", "indeterminate", or "negative", accompanied by levels of evidence, similarly as was proposed by the Cochrane Back Review Group (Table 1) [18, 19]. To assess whether the results of the measurement properties were positive, negative, or indeterminate, we used criteria based on Terwee *et al.* [12].

Table 1) Levels of evidence for the overall quality of the measurement property

Level	Rating	Criterion
Strong	+++ or ---	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
	++ or -	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	±	Conflicting findings
Unknown	?	Only studies of poor methodological quality

Data analysis

Due to the heterogeneity of ultrasound systems and measurement methods among the selected articles, a meta-analysis was not performed.

Also, subgroup analysis could not be performed due to the limited number of studies evaluating the same measurement method.

Therefore, a descriptive analysis was performed.

Findings

Selection of articles

The step-by-step search process of selecting articles based on the PRISMA checklist is shown in Figure 1.

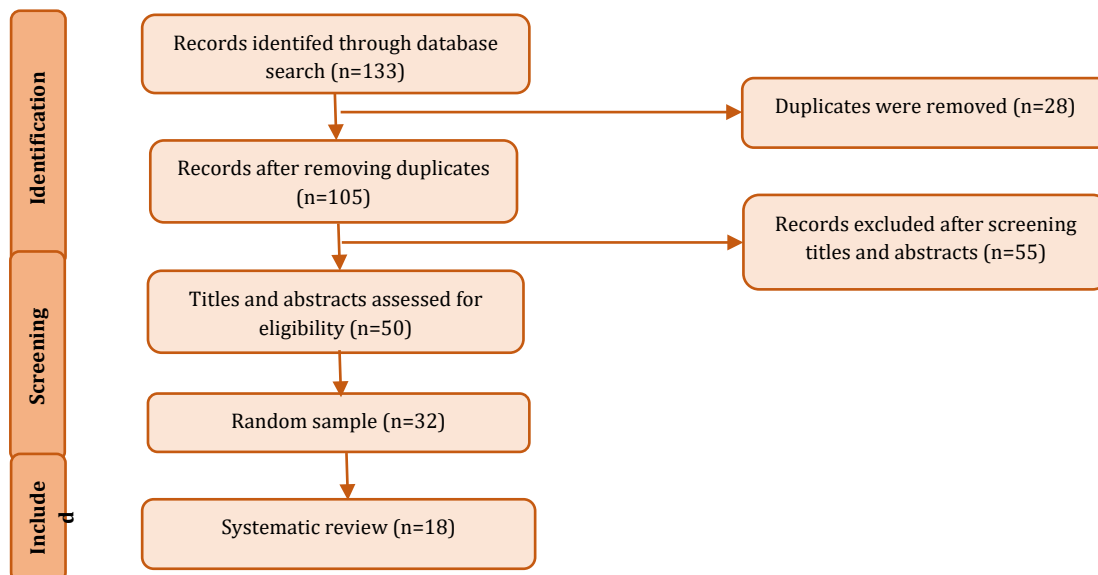


Figure 1) Flowchart of search and selection

A total of 133 cases were registered. After collecting the articles and checking the titles, unrelated, identical, and duplicate articles were removed.

After reading the abstracts of the articles, 50 articles remained, and according to the inclusion and exclusion criteria, finally, 18 full articles were read and reviewed.

Characteristics of selected articles

The 18 selected articles were published between 2012 and 2019, mainly from Hong Kong and Canada, while the rest were from the three countries of Japan, the Netherlands, and China. This review highlighted 4 different ultrasound methods for measuring spinal curvature from ultrasound images, among which SonixTABLET and Scolioscan ultrasound systems were the most popular tools. The standing radiological Cobb method was used as the reference standard in almost all articles, and supine MRI was the reference standard in one article. The methodological quality of the studies for each method and measurement characteristics are presented in Table 2. The combination of results for each method and their level of evidence are presented in Table 3.

Assessment of measurement properties

In most studies, the aspect of stability was considered for reliability, and the internal consistency of the instrument was not considered. Internal consistency refers to the degree of uniformity in people's scoring during a test or from time to another when conducting a multi-item test. To obtain the internal consistency reliability coefficient, the test administrator must administer the test at least twice

in one day. The change in people's scores from one round to the next shows the lack of reliability of the test. The correlation between the scores in the tests is the reliability coefficient of internal consistency. Most studies reported intra- or test-retest and inter-rater or intra-rater reliability, and some obtained only intra-rater reliability [4, 5, 9, 20].

In the 18 reviewed articles, only 6 studies reported missing curves and how to manage these curves [6, 9, 20-23]. Among these, 5 studies only mentioned the number of missing curves but did not state their management method [8, 24-26]. Missing curves are the curves that were measured in radiography but not detected by ultrasound measurements [9].

Studies showed similar methodological shortcomings about average sample size (30 to 49). Out of the 18 reviewed articles, only 10 studies reported sample size, 6 of which had a medium sample size [8, 20, 21, 25, 27, 28], one had a poor sample size [24], and 3 had excellent sample size (≤ 100) [5, 9, 26].

In all the reviewed studies, at least two measurements were performed, and only six studies indicated that the measurements were independent [8, 20-23, 28]. In all the studies, there was no evidence that the patients were stable in the temporary period, which can be one of the most important reasons for the measurement error.

Out of the 18 reviewed articles, only 4 studies have compared the test conditions for the two measurements of radiography and ultrasound [9, 25, 26, 28]. Most researches have mentioned the blinding of raters.

Table 2) Methodological quality of each study with measurement characteristics of each method and measurement tool

Study	Year	Instruments	Measurement Method	Internal consistency	Measurement error	Reliability	Content validity	Structural validity	Criterion validity	Hypotheses testing	Responsiveness
Cheung et al. [3]	2015	Scolioscan system	Transverse process-superior articular process	-	-	Poor	-	-	Poor	-	-
Li et al. [4]	2012	Esaote Technos MPX ultrasound unit	Spinous process method	-	-	Fair	-	-	Fair	-	Fair
Zheng et al. [5]	2016	Scolioscan system	Spinous process method	-	-	Poor	-	-	Fair	-	-
Brink et al. [6]	2018	Scolioscan system	spinous process-transverse process angle	-	-	Fair	-	-	Good	-	-
Lv et al. [7]	2019	Logiqe9 US system	Transverse process-superior articular process	-	-	Poor	-	-	Poor	-	-
Chen et al. [8]	2013	SonixTABLET ultrasound system	Lamina center method	-	Fair	Fair	-	-	Fair	-	-
Zheng et al. [9]	2018	SonixTABLET ultrasound system	Lamina center method	-	Good	Good	-	-	Excellent	-	-
Young et al. [20]	2015	SonixTABLET ultrasound system	Lamina center method	-	Good	Fair	-	-	Good	-	Good
Zheng et al. [21]	2015	SonixTABLET ultrasound system	Lamina center method	-	Good	Good	-	-	Good	-	-
Zheng et al. [22]	2015	SonixTABLET ultrasound system	Lamina center method	-	Good	Good	-	-	Good	-	-
Lee et al. [23]	2019	Scolioscan system	Spinous process method	-	-	Poor	-	-	Poor	-	-
Chen et al. [24]	2016	SonixTABLET ultrasound system	Lamina center method	-	Fair	Fair	-	-	Fair	-	-
Khodaei et al. [25]	2018	SonixTABLET ultrasound system	Lamina center method	-	Fair	Fair	-	-	Fair	-	-
Trac et al. [26]	2019	SonixTABLET ultrasound system	Lamina center method	-	Good	Good	-	-	Excellent	-	-
Li et al. [27]	2015	Esaote Technos MPX ultrasound unit	Spinous process method	-	-	Poor	-	-	Good	-	-
Wang et al. [28]	2015	SonixTABLET ultrasound system	Lamina center method	-	Fair	Fair	-	-	Fair	-	-
Cheung et al. [30]	2013	Scolioscan system	Transverse process-superior articular process	-	-	Poor	-	-	Poor	-	-
Cheung et al. [35]	2015	Scolioscan system	Spinous process-transverse process angle	-	-	Fair	-	-	Good	-	-

Table 3) Combination of results for each method and level of evidence

Method	Internal consistency	Measurement error	Reliability	Content validity	Structural validity	Criterion validity	Hypothesis testing	Responsiveness
Spinous process	0	0	+	+	0	++	+	0
Spinous process-Transverse process	0	0	++	?	0	+++	0	0
Transverse process - superior articular process	0	0	?	0	0	?	0	0
Center of lamina method	0	++	+++	0	0	+++	0	0

The time interval between the two evaluations of radiography and ultrasound varies between one day and 3 months, and it seems that the shorter the interval, the lower the possibility of error. In most studies, the intraclass correlation coefficient with a 95% confidence interval was used for intra- and inter-rater reliability, and only 9 articles reported the standard error of measurement [5, 8, 9, 20, 21, 24-26, 28]. None of the studies provided the correlation coefficient formula. Few studies used the error-index to evaluate the accuracy of ultrasound in the selection of end vertebrae [5, 8, 9, 20]. Most studies adopted Carrier's criterion to evaluate intraclass correlation coefficient values, and only 5 studies did not mention this criterion [3, 4, 27, 29, 30].

Among the validities mentioned in COSMIN's checklist, criterion validity was more frequent because most studies aimed to replace the new tool against the gold standard, i.e., radiography. Validity shows how well the constructed instrument is related to an external standard. If there is an external standard, criterion validity should be checked, and a correlation of at least 0.7 between the tool and the external standard is acceptable [31]. Three articles have used Pearson's correlation coefficient, 9 articles have used linear regression for the relationship between the measurements of two instruments, and 12 studies have used the Bland-Altman method to check the agreement between two assessments. The Bland-Altman method is the most

popular statistical method for evaluating the agreement between two clinical measurement methods [32].

The significance level for most analyzes was set at 0.05. None of the studies reported content validity, construct validity, cross-cultural validity, and hypothesis testing. The purpose of content validity is to ensure the ability of the tool to measure the phenomenon it claims to measure. Important aspects are whether all items are relevant to the construct, purpose, and target population. Construct validity is divided into three aspects; Construct validity is the degree to which the scores of an instrument adequately reflect the dimensions of the construct to be measured. In hypothesis testing, construct validity is examined by testing the proposed hypotheses (for example, examining the correlation or difference between known groups or the correlation of the scores of this instrument with other related instruments). The proposed hypotheses must be special. Otherwise, the results will be biased. Cultural validity is the degree to which the performance of items on a translated or culturally adapted instrument adequately reflects the original version of the instrument [33].

Among all the articles, only one sensitivity and specificity study has been reported to detect the progression of the curve in the use of ultrasound [20], which is one of the most important statistical methods in the response box. Responsiveness or sensitivity refers to the instrument's ability to detect changes over time. Responsiveness in a longitudinal context is considered an aspect of validity [33].

Also, no study reported interpretability. Interpretability refers to the ability to refer qualitative meanings to quantitative scores. Researchers should report on what changes in scores are clinically meaningful.

Level of evidence

Table 3 shows the cumulative level of evidence found for all methods. Four measurement methods were found for ultrasound, and the levels of evidence for each method will be discussed below.

The method of measuring the spinous process angle

The spinous process angle was proposed to measure scoliosis curves by identifying the spinous processes of all vertebrae along the curve and then drawing lines to join all the tips of the spinous processes. Both points form a line, and both lines form an angle. The sum of all angle values becomes the angle of the spiny process [34]. We found four studies that used the spinous process angle method [4, 22, 23, 27]. The purpose of the four studies was to introduce ultrasound to replace radiography for the evaluation of scoliosis. In all four studies, there was no proper methodological quality to assess internal consistency, measurement error, construct validity, hypothesis testing, cross-cultural validity, responsiveness, and interpretability. There is limited positive evidence for

the reliability of this method. All four studies obtained an intraclass correlation coefficient higher than 0.70 (ICC=0.91-0.97), but two studies did not obtain an interclass correlation [4, 27]. Also, moderate positive evidence was obtained for criterion validity (r=0.79-0.87).

The method of measuring the spinous process-transverse process angle

This method is used to evaluate scoliosis curves by measuring two angles of the spinous process and the transverse process. To measure the angle of the spinous process, the angle created between the tips of two adjacent spinous processes in the upper and lower regions of the spine curve is calculated, and to measure the angle between two lines of transverse processes, the most oblique (between the tips of the left and right transverse process) areas The top and bottom of the spine curve is calculated [34]. Only two studies used this method, and in both studies, proper methodological quality was not observed to assess internal consistency, measurement error, construct validity, hypothesis testing, cross-cultural validity, responsiveness, and interpretability [6, 30]. There is moderate positive evidence for the reliability of this method. Both studies obtained intra- and inter-class correlation coefficients higher than 0.70 (ICC=0.97-0.99), but positive and strong evidence for criterion validity was observed. And they reported a strong correlation between radiographic and ultrasound methods (r=0.89-0.99). Also, both methods presented a convincing argument that these methods are a good alternative to radiographic Cobb angle measurement.

Angle method of the transverse process-superior articular process

This method is used to measure the angle between the two lines of the transverse process and the upper articular process, which have the greatest curvature (between the tip of the upper articular process or the left and right transverse process) in the upper and lower regions of the spine curve [34]. Three articles used this method, and in all three studies, the correct methodological quality was not observed to evaluate internal consistency, measurement error, structural validity, hypothesis testing, cross-cultural validity, responsiveness, and interpretability [7, 29, 35]. There is inconclusive evidence of poor methodological quality for reliability. Also, unclear and weak evidence for criterion validity was observed.

Lamina center method

This method is used to measure the angle between the two lamina center lines that have the greatest slope (between the left and right lamina centers) in the upper and lower regions of the spinal curve [34]. Nine articles used this method, which was the most frequent among ultrasound methods [5, 8, 9, 20, 21, 24-26, 28]. All nine studies have good methodological quality to assess internal consistency, structural validity, hypothesis testing, cross-cultural validity, responsiveness, and interpretability. There is strong

positive evidence for the reliability of this method. All nine studies obtained intra- and inter-class correlation coefficients higher than 0.70. There was also moderate positive evidence for measurement error. All articles calculated the standard error of measurement. On the other hand, the level of positive and strong evidence for criterion validity was obtained. Some studies reported the coefficient of determination for evaluation between ultrasound and radiography, and in one study, the sensitivity and specificity were determined [20].

Discussion

A three-dimensional deformity problem of the spine in the frontal and sagittal planes is vertebral rotation, and deformation parameters at different levels may be interdependent [36]. Vertebral rotation cannot be measured directly because these radiographs do not show the exact amount of three-dimensional deformation of the spine in patients with scoliosis [5]. It is well documented that the angle measured based on spinous processes on radiographs underestimates the deformation of the spine with respect to Cobb's angle and shows that the spinous process angle is less angulated compared to Cobb's angles [37]. It has been reported that the magnitude of the axial rotation of the vertebrae is related to the lateral deviation of the vertebrae from the axis of the spine [38]. In fact, the deviations in the angle of the spinous process caused by the rotation of the vertebrae may lead to inaccuracy in the interpretation of the alignment of the vertebral bodies in spine radiography [30]. So an ultrasound scan can provide three-dimensional information about spinal deformity that a simple radiograph cannot provide. Signs such as spinous process, transverse process, and lamina have been identified in ultrasound images, and the possibility of using these signs to evaluate the curvature of the spine in three dimensions has been studied. Therefore, because the axial rotation of the vertebra is an important parameter in the study of scoliosis, a simple and accurate method can increase the use of measurement in the evaluation of scoliosis. Therefore, the lamina center method measures the axial rotation of the vertebra on transverse ultrasound images with simple methods. The results of the current study showed the strongest evidence of the lamina center method. The greatest advantage of this method was the appropriate sample size of over 100 subjects in several studies, which seems that the small sample size in most studies often leads to uncertain results.

Although the present review study showed the accuracy of ultrasound measurements is also affected by various factors. In the main region of thoracic curvature, especially for individuals with greater vertebral rotation, the curvature of the spine results in thoracic rotation, which causes one side of the subject's back to stick out. The uneven surfaces of the

back, in conjunction with the deeper mid-posterior groove in the main chest area, create gaps between the transducer and the scanning area in the middle back of the subject. This poor contact can cause loss or attenuation of the ultrasound signal and hence produce a poor ultrasound image. In addition, the rotating ribs may be placed between the transducer and the head, the reflected energy from the layer may be partially blocked or diverted, and the image quality may be compromised. It can be used in moderate scoliosis, such as Cobb's angle less than 40. In addition, the weakening of ultrasound signals from soft tissues, especially muscles, is the main factor affecting the accuracy of ultrasound measurements in the lower back. For this reason, some current studies reported the error-index because previous studies reported that the end vertebrae of the back are the main source of error in ultrasound measurements, especially in people with high body mass. In these subjects, when the ultrasound signals penetrate the subject's back, the energy decreases exponentially with the thickness of the muscle due to the attenuation and scattering effect. As a result, less energy is reflected from the thick muscle area, and the low quality of the image reduces the accuracy of the measurement. It seems that this is the main reason for missing some curves in the waist area. Missing curves are curves that are measured on radiographs but not detected by ultrasound measurements. Even these missing curves were observed in the upper chest curves. Probably, because the ultrasound transducer is convex and the contact surface in the region of the seventh cervical and first and second dorsal vertebrae is relatively small, these curves were not visible [20]. It is suggested to choose a larger sample size to compensate for and manage the curve in such research so that the number of missing curves does not affect the results. Also, the ultrasound method cannot be applied to people who have undergone spine surgery. Metal implants inside the body strongly reflect ultrasound and block the ultrasound signals from hitting the prominent points on the vertebra. On the other hand, the ultrasound imaging method is only applied to measure the frontal curvature. Sagittal curvature is a very important topic to study in ultrasound imaging of the spine in the future. The sagittal curvature measurement algorithm is still under development, and further research is needed to validate the method at present.

In addition, several important flaws in the reliability between ultrasound and radiographic measurements were observed. In most of the studies, the evaluations were not taken at the same time, and the smallest interval was reported from one hour to three months, which can be the most likely in the measurement error. Also, most of the assessments were taken in different places. Therefore, the subjects' body position may have changed between the two examinations, and this indicates that the two

measurements are not independent. On the other hand, few studies compared the test conditions for the two measurements. During the radiographic examination, people stand in a specific positioning cart, while during the ultrasound examination, people are free. Also, most of the studies did not provide a correct report on the stability of the subjects between the two tests.

In most of the reviewed studies regarding reliability, most of the intra-rater and inter-rater reliability have been reported, while no methodological studies reported correct internal consistency reliability. To obtain the internal consistency reliability coefficient, the test administrator must administer the test at least twice in one day.

The change in people's scores from one round to the next shows the lack of reliability of the test. The correlation between the scores in the tests is the reliability coefficient of internal consistency. It seems that the non-reporting of internal consistency in the reviewed studies is because some believe that internal consistency is not a component of reliability, and stability and standard error of the instrument constitute components of reliability^[36]. They believe that internal consistency is a way to prove the single factor of the instrument and shows that all the items of the instrument together measure the same concept.

On the other hand, some believe that when we run the test once or when a test is scored in a similar situation by several evaluators, internal consistency is actually checked^[30]. Although various review studies that examined the psychometrics of health instruments reported internal consistency reliability, with a glance at the previous studies and the box related to the items of the COSMIN's checklist, it seems internal consistency is more suitable for questionnaire tools and qualitative tools.

Very high reliability was obtained for most of the four ultrasound methods. In general, intra-rater reliabilities are higher than inter-rater reliabilities, indicating that a rater agrees more with his own measurements than with the assessments of others. High intra- and inter-rater reliability were also found for the methods of measuring spinous appendage angle, transverse appendage-upper articular appendage angle, and lamina center method (ICC=0.89-1.00). On the other hand, the intra-rater (ICC=0.57-0.98) and inter-rater values (ICC=0.75-0.96) were obtained relatively large for the spinous appendage-transverse appendage angle method. Criterion validity for ultrasound methods varied from moderate to excellent.

Also, some studies reported different results of ultrasound scans obtained by evaluators with different levels of experience. Experienced operator scans showed higher measurement accuracy than trainee operator scans. It seems that the strong pressure applied to maintain good contact between the transducer and the subject's back can also affect

the standing position during the ultrasound scan and lead to missing curves. As a suggestion, a custom-designed frame can standardize the position of the subject's upper body and prevent unexpected bending, tilting, and rotation of the upper body and minimize standard errors of measurement. It should be noted that most of the studies reported the experience of evaluators.

Conclusion

Ultrasound seems to be a suitable method to replace radiography in people with scoliosis. The ultrasound device is mobile and low-cost. The reliability and accuracy of the measurements of the vertebral body and axial rotation from ultrasound have been shown by many studies, and recently, the use of ultrasound imaging to identify the progression of the curve in children with idiopathic scoliosis has been reported. Therefore, using the ultrasound method to follow up non-progressive cases can reduce more than 70% of radiographs. As a result, this study showed that the method of measuring the center of the lamina is a simple and reliable method to measure the curvature of the spine of people with scoliosis without the need to expose the patient to radiation.

Acknowledgements: We thank all the mentors and advisors, as well as referees.

Ethical Permission: This study was approved by the Ethics Committee of Tehran University.

Conflict of Interests: Nothing has been reported by the authors.

Authors' Contribution: Saberi M (First Author), Main Researcher (50%); Rajabi R (Second Author), Introduction Writer/Statistical Analyst (20%); Minoonejad H (Third Author), Methodologist/Assistant Researcher (20%); Karimizadeh Ardakani M (Fourth Author), Discussion Writer (10%)

Funding/Support: Nothing has been reported by the authors.

References

- 1- Kim H, Kim HS, Moon ES, Yoon CS, Chung TS, Song HT, et al. Scoliosis imaging: What radiologists should know. *Radiographics*. 2010;30(7):1823-42.
- 2- Ronckers CM, Land CE, Miller JS, Stovall M, Lonstein JE, Doody MM. Cancer mortality among women frequently exposed to radiographic examinations for spinal disorders. *Radiat Res*. 2010;174(1):83-90.
- 3- Cheung CWJ, Zhou GQ, Law SY, Lai KL, Jiang WW, Zheng YP. Freehand three-dimensional ultrasound system for assessment of scoliosis. *J Orthop Transl*. 2015;3(3):123-33.
- 4- Li M, Ng B, Cheng J, Ying M, Zheng YP, Lam TP, et al. Could clinical ultrasound improve the fitting of spinal orthosis for the patients with AIS?. *Eur Spine J*. 2012;21(10):1926-35
- 5- Zheng R, Young M, Hill D, Le LH, Hedden D, Moreau M, et al. Improvement on the accuracy and reliability of ultrasound coronal curvature measurement on adolescent idiopathic scoliosis with the aid of previous radiographs. *Spine (Phila Pa 1976)*. 2016;41(5):404-11.
- 6- Brink RC, Wijdicks SPJ, Tromp IN, Schlosser TPC, Kruyt MC, Beek FJA, et al. A reliability and validity study for

- different coronal angles using ultrasound imaging in adolescent idiopathic scoliosis. *Spine J* 2018;18(6):979-85.
- 7- Lv P, Chen J, Dong L, Wang L, Deng Y, Li K, et al. Evaluation of scoliosis with a commercially available ultrasound system. *J Ultras Med*. 2019;39(1):1-8.
- 8- Chen W, Lou EHM, Zhang PQ, Le LH, Hill D. Reliability of assessing the coronal curvature of children with scoliosis by using ultrasound images. *J Child Orthop* 2013;7(6):521-9.
- 9- Zheng R, Hill D, Hedden D, Mahood J, Moreau M, Southon S, et al. Factors influencing spinal curvature measurements on ultrasound images for children with adolescent idiopathic scoliosis (AIS). *PLoS ONE* 2018;13(6):e0198792.
- 10- Colton D, Covert RW. Designing and constructing instruments for social research and evaluation. 1st Ed. San Francisco: Jossey-Bass;2007.
- 11- Waltz CF, Strickland OL, Lenz ER. Measurement in nursing and health research. 4 th Ed. New York: Springer Publishing Company;2010.
- 12- Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
- 13- Reneman MF, Dijkstra A, Geertzen JH, Dijkstra PU. Psychometric properties of chronic pain acceptance questionnaires: A systematic review. *Eur J Pain*. 2010;14(5):457-65.
- 14- Terwee CB, Schellingerhout JM, Verhagen AP, Koes BW, de Vet HC. Methodological quality of studies on the measurement properties of neck pain and disability questionnaires: A systematic review. *J Manipulative Physiol Ther*. 2011;34(4):261-72.
- 15- Lidwine B, Mokkink Caroline B, Terwee Donald L, Jordi Alonso P, Paul W, Stratford Dirk L, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual Life Res*. 2010;19(4):539-49.
- 16- Mokkink LB, Prinsen CAC, Bouter LM, de Vet HCW, Terwee CB. The consensus-based standards for the selection of health measurement instruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*. 2016;20(2):105-13.
- 17- Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LB, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res*. 2012;21(4):651-7.
- 18- Furlan AD, Pennick V, Bombardier C, van Tulder M. Updated method guidelines for systematic reviews in the Cochrane back review group. *Spine (Phila Pa 1976)*. 2019;34(18):1929-41.
- 19- van Tulder M, Furlan A, Bombardier C, Bouter L, Board. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. *Spine (Phila Pa 1976)*. 2003;28(12):1290-9.
- 20- Young M, Hill DL, Zheng R, Lou E. Reliability and accuracy of ultrasound measurements with and without the aid of previous radiographs in adolescent idiopathic scoliosis (AIS). *Eur Spine J*. 2015;24(7):1427-33
- 21- Zheng R, Chan AC, Chen W, Hill DL, Le LH, Hedden D, et al. Intra- and inter-rater reliability of coronal curvature measurement for adolescent idiopathic scoliosis using ultrasonic imaging method- a pilot study. *Spine Deform*. 2015;3(2):151-8.
- 22- Zheng YP, Lee TTY, Lai KKL, Yip BHK, Zhou GQ, Jiang WW, et al. A reliability and validity study for Scolioscan: A radiation-free scoliosis assessment system using 3D ultrasound imaging. *Scoliosis Spinal Disord*. 2016;11:13.
- 23- Lee TTY, Jiang WW, Cheng CLK, Lai KA, To MKT, Castelein RM, et al. A novel method to measure the sagittal curvature in spinal deformities: The reliability and feasibility of 3-d ultrasound imaging. *Ultrasound Med Biol*. 2019;45(10):2725-35.
- 24- Chen W, Le LH, Lou EHM. Reliability of the axial vertebral rotation measurements of adolescent idiopathic scoliosis using the center of lamina method on ultrasound images: In vitro and in vivo study. *Eur Spine*. 2016;25(10):3265-73.
- 25- Khodaei M, Hill D, Zheng R, Le LH, Lou EHM. Intraand inter-rater reliability of spinal flexibility measurements using ultrasonic (US) images for non-surgical candidates with adolescent Idiopathic scoliosis: A pilot study. *Eur Spine J*. 2018;27(9):2156-64.
- 26- Trac S, Zheng R, Hill DL, Lou E. Intra- and interrater reliability of Cobb angle measurements on the plane of maximum curvature using ultrasound imaging method. *Spine Deform*. 2019;7(1):18-26.
- 27- Li M, Cheng J, Ying M, Ng B, Lam TP, Wong MS. A preliminary study of estimation of Cobb's angle from the spinous process angle using a clinical ultrasound method. *Spine Deform*. 2015;3(5):476-82.
- 28- Wang Q, Li M, Lou EHM, Wong MS. Reliability and validity study of clinical ultrasound imaging on lateral curvature of adolescent idiopathic scoliosis. *PLoS ONE*. 2015;10(8):e0135264.
- 29- James CW, Law SY, Zheng YP. Development of 3D ultrasound system for assessment of adolescent idiopathic scoliosis (AIS): And System Validation. 35th Annual International Conference of the IEEE EMBS Osaka; 2013.
- 30- Cheung CWJ, Law SY, Zheng YP. Development of 3-D ultrasound system for assessment of adolescent idiopathic scoliosis (AIS): and system validation. *Annu Int Conf IEEE Eng Med Biol Soc*. 2013;2013:6474-7.
- 31- Ebadi A, Taghizadeh Z, Montazeri A, Shahvari Z, Tavousi M, Bagherzadeh R. Translation, development and psychometric properties of health related measures-Part 2: Construct validity, reliability and responsiveness. *Payesh*. 2017;16(4):445-55. [Persian]
- 32- Gee A, Prager R, Treece G, Berman L. Engineering a freehand 3D ultrasound system. *Pattern Recogn Lett*. 2003;24(4-5):757e77.
- 33- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-45.
- 34- Wu HD, Liu W, Wong MS. Reliability and validity of lateral curvature assessments using clinical ultrasound for the patients with scoliosis: A systematic review. *Eur Spine J*. 2019;29(4):717-25.
- 35- Cheung CWJ, Zhou GQ, Law SY, Mak TM, Lai KL, Zheng YP. Ultrasound volume projection imaging for assessment of scoliosis. *IEEE T Med Imaging*. 2015;34(8):1760-8.
- 36- Stokes IA, Bigalow LC, Moreland MS. Three-dimensional spinal curvature in idiopathic scoliosis. *J Orthop Res*. 1987;5(1):102-13.
- 37- Herzenberg JE, Waanders NA, Closkey RF, Schultz AB, Hensinger RN. Cobb angle versus spinous process angle in adolescent idiopathic scoliosis. The relationship of the

