



Inter- and intra-reliability of Cobb angle measurement in pediatric scoliosis using PACS (picture archiving and communication systems methods) for clinicians with various levels of experience

Christopher Lucasti, Mohammad Nadir Haider, Ian P. Marshall, Richard Thomas, Maxwell M. Scott, Michael R. Ferrick

UBMD Orthopedics and Sports Medicine, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, SUNY, Buffalo, NY, USA

Contributions: (I) Conception and design: IP Marshall, MR Ferrick, R Thomas; (II) Administrative support: MR Ferrick, R Thomas; (III) Provision of study materials or patients: MR Ferrick, R Thomas; (IV) Collection and assembly of data: IP Marshall, MR Ferrick, R Thomas, C Lucasti; (V) Data analysis and interpretation: MN Haider; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Maxwell M. Scott. Department of Orthopaedics, Jacobs School of Medicine and Biomedical Sciences, 4949 Harlem Road, Amherst, NY 14226, USA. Email: mscott8@buffalo.edu.

Background: Scoliosis is a common spinal malformation, which is typically diagnosed and assessed for severity by measuring the Cobb angle on a plain radiograph. Although manual methods are still considered the gold standard, recent studies have shown that picture archiving and communication systems (PACS) methods have comparable intraobserver and interobserver reliability levels for experienced clinicians. However, there is limited data for these reliability measures using PACS in less experienced medical personnel and radiologist. The objective of this study was to determine if inter- or intra-reliability differs for clinicians with different experience levels and if PACS provides adequate inter- or intra-reliability.

Methods: Ten radiographic images of scoliosis patients were obtained via a retrospective chart review. After a brief training exercise, Cobb angle measurements were performed twice on each radiograph, separated by one week, using the PACS method. Subject groups included medical students, orthopedic resident physicians, orthopedic surgeons and radiologists. Each subject group had two participants (R1 and R2). Each participant calculated a Cobb angle using the PACS method at Week 1 and Week 2 (W1 and W2). Intra-rater reliability was calculated by comparing each rater's W1 and W2 measurements, inter-rater reliability within each experience level was calculated by comparing R1 and R2's measurements at W1 and W2, and inter-rater agreement was found by comparing one participant's W1 measurement to participants W1 measurement from another group.

Results: Resident physicians, orthopedic surgeons, and radiologist subject groups had significant levels of interrater reliability within-groups ($r=0.946, 0.968, 0.967, P<0.01$) and between groups. Additionally, for all these three groups, individual intra-rater reliability levels were high ($r=0.973, 0.976, 0.954, 0.977, 0.974, 0.913, P<0.01$). Only inexperienced medical student raters had poor within-group and between-group interrater reliability measures ($r=-0.176, -0.188, -0.071, -0.188, P>0.05$). However, individual intra-rater reliability levels were high ($r=0.919, 0.592, P<0.05$).

Conclusions: PACS is a reliable method of measuring Cobb angles in most experience levels. For inexperienced medical students, interrater reliability within-groups and between groups are not statistically significant, highlighting medical students lack expertise in calculating cobb angles. However, individual intra-rater reliability levels were high for medical students, validating the use of PACS as a consistent method for calculating Cobb angles.

Keywords: Picture archiving and communication systems (PACS); scoliosis; Cobb angle; reliability

Received: 30 October 2021; Accepted: 17 January 2023; Published online: 13 February 2023.

doi: 10.21037/asj-21-105

View this article at: <https://dx.doi.org/10.21037/asj-21-105>

Introduction

The Cobb angle is the most commonly used measurement for assessing the magnitude of spinal curvature deformities, especially scoliosis, on radiographs (1,2). By determining the Cobb angle, orthopedists are able to take a largely quantifiable approach to evaluating severity, prognosis, and possible interventions for scoliosis patients (2-5). In order to measure the Cobb angle on a posteroanterior (PA) radiograph, the apex of the spinal curve and the two vertebrae with the greatest tilt from the midsagittal plane, one superior and one inferior to the apex, should be identified. This manual method for obtaining Cobb angle measurements is currently the gold standard for assessing scoliosis severity on plain PA radiographs, however, there are other digital measurement options that are comparable in accuracy and reliability, including picture archiving and communication systems (PACS) (1,2,6). The literature reports consistent levels of intra- and inter-observer reliability across different Cobb angle measurement modalities with average variability ranging between 4 and 8 degrees for both categories (3-5,7-11). However, there is a paucity of data regarding these values with respect to experience level of the observer, particularly in the modern era of digital radiology displayed on PACS.

As a diagnostic and treatment-guiding tool, it is extremely important to ensure Cobb angle intra- and inter-observer reliability measures, which are defined as the consistency of serial measurements by one individual and the consistency of measurements across multiple raters, respectively, when measuring identical PA radiographs. Various studies have discussed the many methods available to obtain Cobb angle measurements for a scoliotic curve. Some of these include smartphone-assisted, ultrasound-guided, and digital measurements in addition to the manual plain radiograph method (1,6,12,13). The ease of accessibility and rapid retrieval of imaging studies has increased the role of digital imaging for the diagnosis and management of scoliosis. Digital measurement of the Cobb angle in scoliosis is not only a suitable alternative to the gold-standard manual method, but some studies have found it to be superior to the manual method. These studies show evidence for improved Cobb angle measurement speed, accuracy, and reliability when measured with digital systems, such as PACS, compared to the standard manual methods (4,14-18).

Regardless of the method used, Cobb angle measurements are a valuable tool that helps diagnose and direct treatments for patients affected by scoliosis. Although previous studies support high levels of intraobserver and interobserver reliability for a variety of Cobb angle measurement methods, this has only been established for skilled, experienced professionals like orthopedists and radiologists (15,16,18). There is limited data for these measures in less-skilled subjects, such as medical students or other members of the medical team with less training. Accordingly, the objective of this study was to determine if inter- or intra-reliability differs for clinicians with different experience levels (medical student, orthopedic resident physician, orthopedic surgeon, radiology attending) and if PACS provides adequate inter- or intra-reliability.

Determining these reliability measures for inexperienced groups may help bridge the gap between the different roles assumed by healthcare workers who have varying levels of experience and training. By focusing on this group of clinicians, we will better understand the skills less experienced clinicians need to focus on developing. If they have high intra-reliability on measurements, but poor inter-reliability, then more focus needs to be placed on teaching this group the proper measurement techniques. If there is poor intra-reliability within the medical school group, then there is a need for greater emphasis on teaching students a systematic way to approach radiographic interpretation

Highlight box

Key findings

- Intraobserver and interobserver reliabilities using picture archiving and communication systems (PACS) for PGY-3 Orthopedic Residents, Orthopedic Surgeons and Radiologists had consistent measurement agreements.
- For medical students, the intraobserver values were significantly consistent for individual subjects. Conversely, interobserver agreement amongst medical student raters and between medical students and more experienced clinicians was considered poor.

What is known and what is new?

- There is consistent levels of intra- and inter-observer reliability across different Cobb angle measurement modalities and digital systems improve Cobb angle measurement speed, accuracy, and reliability.
- This study is the first to examine the use of PACS for inter- and intra- reliability amongst less experienced clinicians.

What is the implication, and what should change now?

- PACS is a reliable alternative to standard manual Cobb angle measurements on plan radiographic film.
- Further education is needed on teaching less experienced clinicians proper measurement techniques with PACS.

to ensure repeatability. Accordingly, this project will aid in ensuring this group of less experienced healthcare workers are taught the proper measurement techniques by allowing more experienced clinicians to learn how best to educate this population.

Methods

Study design

This prospective case control study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional review board of the University at Buffalo (No. 00005660) and all appropriate consents were obtained. A convenient sample of 10 radiographic images of pediatric patients (aged 17 and under) diagnosed with a varying degree of adolescent idiopathic scoliosis were obtained from electronic medical records. Two raters of each experience level measured the Cobb angles on each patient radiograph twice, one week apart. Inter- and intra-rater agreement between different experience levels were obtained and compared. An important possible source of error in Cobb Angle measurement is the variation in production of the spinal radiograph. To control for this variability, the raters utilized the same radiographs to make their respective measurements, removing this clinically significant source of measurement error (19).

Procedure

Two medical students, 2 orthopedic surgery residents, 2 experienced orthopedic attendings and 2 radiologists measured the Cobb angle on each radiograph twice. Given the time commitment of this study and the training required for PACS, 10 clinicians was a sufficient number to initially investigate the topic. Each rater was trained using video demonstrations that were identical across all experience levels. Raters were given a folder with de-identified images. Measurement instructions were not be provided following the instructional demonstration and no questions were answered. To measure the Cobb angle, the raters first identified which vertebrae were the end vertebrae of the curve deformity, that is the vertebrae whose endplates are most tilted towards each other. Then utilizing the dedicated angle tool on PACS, lines were drawn along the endplates, and the angle where the two lines intersect was measured.

Main outcome measure

Our main outcome measure is agreement between raters on a continuous variable (degrees of curvature). Intra- and inter-class correlation coefficient (ICC) between averaged measured were calculated and a Cronbach's α obtained. An α of ≥ 0.9 is considered excellent agreement, $0.9 > \alpha \geq 0.8$ is good, $0.8 > \alpha \geq 0.7$ is acceptable, $0.7 > \alpha \geq 0.6$ is questionable and $\alpha < 0.6$ is poor or unacceptable internal consistency.

Statistical analysis

Raters from each of the four experience levels (student, resident, attending and radiologist) were labeled as Rater 1 (R1) and Rater 2 (R2) and measurements from each week were labeled as Week 1 (W1) and Week 2 (W2) respectively (total 16 groups). Univariate statistics were performed. Correlation between grouped measurements were calculated using Pearson's r . Intra-rater reliability was calculated by comparing each rater's W1 and W2 measurements, inter-rater reliability within each experience level was calculated by comparing R1 and R2's measurements at W1 and W2, and inter-rater agreement between each of the 4 experience levels was calculated by comparing one experience level's R1's W1 measurement to another experience level's R1's W1 measurement and R2's W1 measurement and R2's W1 measurement respectively. Cronbach's α with 95% confidence limits (CLs) were calculated. To account for multiple comparisons, a post-hoc Bonferroni correction was applied and a P value of 0.003 was considered significant (0.05/16 raters). Statistical analysis was performed using SAS Version 9.4 (20).

Results

All raters measured the Cobb angle at Week 1 and 2 and all were included in the analysis. Median age of patient's was 14 years (interquartile range 10.75 to 14.5 years), were 60% female, and mean Cobb angle according to both radiologist's measurements was 30.45 ± 9.2 degrees (range, 12–55 degrees). Intra-rater reliability for Rater 1 and 2 at each experience level and inter-rater reliability within each experience level is presented in *Table 1*. Medical students had variable intra-rater agreement, with one medical student (R1) having excellent internal consistency while the other student (R2) had acceptable consistency, leading to unacceptable inter-rater consistency. The other experience levels (resident, attending and radiologist) had excellent

Table 1 Intra- and inter-rater reliability of Cobb angle measurement

Medical personnel	Intra-rater reliability R1 (W1 vs. W2)			Intra-rater reliability R2 (W1 vs. W2)			Inter-rater reliability (W1 vs. W2)		
	Correlation	P value	Intra-class correlation coefficient	Correlation	P value	Intra-class correlation coefficient	Correlation	P value	Inter-class correlation coefficient
Medical student	0.919	<0.001	0.944 (0.775, 0.986)	0.592	0.071	0.722 (-0.120, 0.931)	-0.176	0.627	-0.375 (-4.536, 0.658)
Resident	0.973	<0.001	0.986 (0.943, 0.997)	0.976	<0.001	0.987 (0.950, 0.997)	0.946	<0.001	0.960 (0.839, 0.990)
Attending	0.954	<0.001	0.957 (0.827, 0.989)	0.977	<0.001	0.988 (0.952, 0.997)	0.968	<0.001	0.982 (0.927, 0.995)
Radiologist	0.974	<0.001	0.984 (0.934, 0.996)	0.913	<0.001	0.954 (0.816, 0.989)	0.967	<0.001	0.983 (0.931, 0.996)

First listed correlation value in each comparison was calculated using Pearson's r. P values less than 0.05 were significant. The second listed correlation is Cronbach's α with 95% CLs. A post-hoc Bonferroni correction was applied and a P value less than 0.003 was considered significant (0.05/16 raters). R1, Rater 1; R2, Rater 2; W1, Week 1; W2, Week 2; CLs, confidence limits.

intra-rater and interrater consistency.

Inter-rater reliability between raters of different experience levels are presented in *Table 2*. Residents, attendings and radiologists had excellent agreement among each other, however, all three groups had variable consistency with the medical students, with medical student (R1) having excellent agreement with all other experience levels and medical student (R2) having unacceptable agreement with all other experience levels.

Discussion

The Cobb angle is considered the gold standard for the measurement of spinal deformities (1,2,6). As a clinical measurement, the Cobb angle provides important information regarding the diagnosis and severity of scoliosis and helps guide treatment decisions. Classically, pediatric orthopedic surgeons, in conjunction with their radiology colleagues, have determined Cobb angles manually with radiographs to dictate treatment plans for their patients (2,5,21). The Cobb angle, when used in concert with individual clinical symptomatology and risk of progression, is a valuable tool that helps clinicians better assess and manage cases of pediatric scoliosis. Typical cases of pediatric scoliosis are largely evaluated based on general guidelines, which provide treatment recommendations for scoliosis based on the magnitude of Cobb angle measurements; however, these guidelines are not absolute (2). Additional clinical features, including skeletal maturity level, and other patient-specific comorbidities and desires also have a role in determining final treatment decisions for patients with pediatric scoliosis (2,5,6,21).

Cobb angles ranging less than 10 degrees are not diagnostic of scoliosis and are considered to be mild spinal asymmetry without any necessary interventions. Diagnosis of scoliosis begins with Cobb angle measurement of greater than 10 degrees. If spinal deviation exceeds 10 degrees, but remains less than 20 degrees, mild scoliosis is diagnosed and only close observation and follow-up is warranted to monitor possible further curvature progression. Curves between 20 and 25 degrees are still considered to be mild scoliosis, however, some degree of bracing is occasionally recommended. This intervention is especially important for children and adolescents who have not yet reached terminal height in order to reduce the risk of worsening curvature over time. Cobb angles greater than 25 and less than 40 degrees are classified as moderate scoliosis and bracing is recommended. This is again especially important for

Table 2 Inter-rater reliability between experience levels

Medical personnel comparison	Inter-rater reliability 1 (R1 vs. R1)				Inter-rater reliability 2 (R2 vs. R2)			
	Correlation	P value	Inter-class correlation coefficient	P value	Correlation	P value	Inter-class correlation coefficient	P value
Medical student and resident	0.919	<0.001	0.953 (0.809, 0.988)	<0.001	-0.188	0.603	-0.460 (-4.876, 0.637)	0.709
Medical student and attending	0.942	<0.001	0.949 (0.793, 0.987)	<0.001	-0.071	0.846	-0.152 (-3.636, 0.714)	0.582
Medical student and radiologist	0.929	<0.001	0.956 (0.822, 0.989)	<0.001	-0.188	0.604	-0.432 (-4.778, 0.644)	0.700
Resident and attending	0.954	<0.001	0.971 (0.882, 0.993)	<0.001	0.984	<0.001	0.992 (0.967, 0.998)	<0.001
Resident and radiologist	0.958	<0.001	0.978 (0.912, 0.995)	<0.001	0.955	<0.001	0.966 (0.864, 0.922)	<0.001
Attending and radiologist	0.945	<0.001	0.968 (0.871, 0.992)	<0.001	0.938	<0.001	0.955 (0.819, 0.989)	<0.001

First listed correlation value in each comparison was calculated using Pearson's *r*. P values less than 0.05 were significant. The second listed correlation is Cronbach's α with 95% CLs. A post-hoc Bonferroni correction was applied and a P value less than 0.003 was considered significant (0.05/16 raters). R1, Rater 1; R2, Rater 2; CLs, confidence limits.

younger patients who have not reached skeletal maturity. For adolescents near skeletal maturity who are minimally symptomatic, treatment for moderate scoliosis may monitor for progression with potential intervention at a later date. There is a high degree of variability in making treatment decisions for cases of mild and moderate scoliosis, which are impacted by several factors, including patient compliance (2,21). However, scoliosis with Cobb angles greater than 40 degrees are severe and will typically be managed with bracing, however, spinal fusion surgery is an option as well. Curves greater than 70 degrees or those severe enough to effect heart and lung function are most commonly treated with surgery (2,6,12,21).

With increasing levels of technology in healthcare, radiographs are now rarely printed and examined manually. More commonly, radiographs are stored, accessed, and examined digitally on systems like PACS, which has been shown to have comparable levels of intraobserver and interobserver reliability for highly-skilled medical professionals when compared with manual methods (1,4,14-18,21). Regardless of the method used, it is imperative to obtain accurate and consistent Cobb angle measurements to ensure proper management of all scoliosis patients. This is particularly important in cases involving growing children or adolescents as they are more susceptible to increased curvature progression as they develop to skeletal maturity. Small degrees of Cobb angle measurement variabilities can artificially alter medical impressions and ultimately treatment recommendations. Therefore, both intraobserver and interobserver reliability levels are extremely important to ensure consistent and

proper management of equally severe pediatric scoliosis patients while preventing any unnecessary treatment and morbidity (2,5,6,21).

The medical school group showed that inexperienced subjects can produce consistent Cobb angle measurements while using PACS, however, they were likely doing so inaccurately due to their lack of any formal training prior to the study. However, this study is limited in its findings by the relatively small group used for inter-observer reliability. Future studies should include more participants and try to determine the reproducibility of these findings. Moreover, future studies should aim to determine the amount of training necessary for inexperienced subjects to achieve significant levels of within-group and between-group interobserver reliability with regard to more experienced and skilled observers. Additionally, future studies should examine the levels of interobserver and intraobserver reliability for Cobb angle measurements in inexperienced subjects using manual methods to compare with these measures when using PACS.

Conclusions

Until the present study, there was limited data on the levels of interobserver and intraobserver reliability for Cobb angle measurements amongst inexperienced subjects, especially for utilizing PACS. Intraobserver and interobserver reliabilities within and between participant groups with PGY-3 or greater levels of experience had consistent measurement agreements. For medical students, the intraobserver values were significantly consistent for

individual subjects. Conversely, interobserver agreement amongst medical student raters and between medical students and more experienced clinicians was considered poor. These results demonstrate that utilization of the PACS for highly experienced workers is a reliable alternative to the standard manual Cobb angle measurement on plain radiographs. Additionally, since medical students have high intra-reliability but poor inter-reliability, further education is needed on teaching these less experienced clinicians the proper measurement techniques.

Acknowledgments

Funding: None.

Footnote

Data Sharing Statement: Available at <https://asj.amegroups.com/article/view/10.21037/asj-21-105/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://asj.amegroups.com/article/view/10.21037/asj-21-105/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional review board of the University at Buffalo (No. 00005660) and all appropriate consents were obtained.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Trac S, Zheng R, Hill DL, et al. Intra- and Interrater Reliability of Cobb Angle Measurements on the Plane of Maximum Curvature Using Ultrasound Imaging Method. *Spine Deform* 2019;7:18-26.
2. Horne JP, Flannery R, Usman S. Adolescent idiopathic scoliosis: diagnosis and management. *Am Fam Physician* 2014;89:193-8.
3. Chan AC, Morrison DG, Nguyen DV, et al. Intra- and Interobserver Reliability of the Cobb Angle-Vertebral Rotation Angle-Spinous Process Angle for Adolescent Idiopathic Scoliosis. *Spine Deform* 2014;2:168-75.
4. Cheung J, Wever DJ, Veldhuizen AG, et al. The reliability of quantitative analysis on digital images of the scoliotic spine. *Eur Spine J* 2002;11:535-42.
5. Gupta MC, Wijesekera S, Sossan A, et al. Reliability of radiographic parameters in neuromuscular scoliosis. *Spine (Phila Pa 1976)* 2007;32:691-5.
6. Gstoettner M, Sekyra K, Walochnik N, et al. Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools. *Eur Spine J* 2007;16:1587-92.
7. Carman DL, Browne RH, Birch JG. Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobserver variation. *J Bone Joint Surg Am* 1990;72:328-33.
8. Loder RT, Spiegel D, Gutknecht S, et al. The assessment of intraobserver and interobserver error in the measurement of noncongenital scoliosis in children < or = 10 years of age. *Spine (Phila Pa 1976)* 2004;29:2548-53.
9. Loder RT, Urquhart A, Steen H, et al. Variability in Cobb angle measurements in children with congenital scoliosis. *J Bone Joint Surg Br* 1995;77:768-70.
10. Morrissy RT, Goldsmith GS, Hall EC, et al. Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error. *J Bone Joint Surg Am* 1990;72:320-7.
11. Tauchi R, Tsuji T, Cahill PJ, et al. Reliability analysis of Cobb angle measurements of congenital scoliosis using X-ray and 3D-CT images. *Eur J Orthop Surg Traumatol* 2016;26:53-7.
12. Langensiepen S, Semler O, Sobottke R, et al. Measuring procedures to determine the Cobb angle in idiopathic scoliosis: a systematic review. *Eur Spine J* 2013;22:2360-71.
13. Zhang J, Lou E, Shi X, et al. A computer-aided Cobb angle measurement method and its reliability. *J Spinal Disord Tech* 2010;23:383-7.
14. Kuklo TR, Potter BK, Schroeder TM, et al. Comparison of manual and digital measurements in adolescent idiopathic scoliosis. *Spine (Phila Pa 1976)* 2006;31:1240-6.
15. Shea KG, Stevens PM, Nelson M, et al. A comparison

- of manual versus computer-assisted radiographic measurement. Intraobserver measurement variability for Cobb angles. *Spine (Phila Pa 1976)* 1998;23:551-5.
16. Srinivasalu S, Modi HN, Smehta S, et al. Cobb angle measurement of scoliosis using computer measurement of digitally acquired radiographs-intraobserver and interobserver variability. *Asian Spine J* 2008;2:90-3.
 17. Tanure MC, Pinheiro AP, Oliveira AS. Reliability assessment of Cobb angle measurements using manual and digital methods. *Spine J* 2010;10:769-74.
 18. Zmurko MG, Mooney JF 3rd, Podeszwa DA, et al. Inter- and intraobserver variance of Cobb angle measurements with digital radiographs. *J Surg Orthop Adv* 2003;12:208-13.
 19. Pruijs JE, Hageman MA, Keessen W, et al. Variation in Cobb angle measurements in scoliosis. *Skeletal Radiol* 1994;23:517-20.
 20. Inc SI. SAS/ACCESS® 9.4 interface to ADABAS: reference. SAS Institute Inc., Cary, NC; 2013.
 21. Płaszewski M, Grantham W, Jespersen E. Screening for scoliosis - New recommendations, old dilemmas, no straight solutions. *World J Orthop* 2020;11:364-79.

doi: 10.21037/asj-21-105

Cite this article as: Lucasti C, Haider MN, Marshall IP, Thomas R, Scott MM, Ferrick MR. Inter- and intra-reliability of Cobb angle measurement in pediatric scoliosis using PACS (picture archiving and communication systems methods) for clinicians with various levels of experience. *AME Surg J* 2023;3:12.