

Intra- and inter-observer reliability of the Cobb measurement by chiropractic interns using digital evaluation methods

Jesse Cracknell, BA, DC^{1,2}
 Douglas M. Lawson, MSc¹
 John A. Taylor, DC, DACBR¹

Introduction: It is important to create a body of evidence surrounding the reliability of certain diagnostic criteria. While the reliability of the Cobb measurement is well established with various licensed health care professionals, this study aims to determine the inter- and intra-observer reliability of the Cobb Measurement among chiropractic interns.

*Methods: Fourteen chiropractic interns analyzed 10 pre-selected digital spinal radiographs on a Picture Archiving and Communication System (PACS) in two separate rounds of observation. The participants indicated their choice of end vertebra and Cobb Measurement in each round of observation. Agreement on vertebral levels selected was estimated using percentage agreement. Intra-observer reliability was estimated using the Pearson *r* correlation coefficient, and inter-observer correlation was estimated using the Inter-Class Coefficient (ICC).*

*Results: The range of percentage agreement on vertebral level selection was 0.36 – 0.79. The Pearson *r**

Introduction : Il est important de créer un ensemble de preuves sur la fiabilité de certains critères de diagnostic. Bien que la fiabilité de la mesure Cobb soit bien établie chez divers professionnels de la santé, cette étude vise à déterminer la fiabilité des observations individuelles et entre ces observations de la mesure Cobb parmi des stagiaires en chiropratique.

*Méthodologie : Quatorze stagiaires en chiropratique ont analysé 10 radiographies numériques de la colonne vertébrale sur un système d'archivage et de transmission d'image (PACS) en deux séances distinctes d'observation. Les participants ont indiqué leur choix de vertèbre limite et de mesure Cobb dans chaque observation. Un consensus sur les niveaux vertébraux sélectionnés a été établi en fonction du pourcentage de concordance. La fiabilité des observations individuelles a été estimée à l'aide du coefficient de corrélation de Pearson *r*, et la corrélation entre les observations des stagiaires a été estimée à l'aide du coefficient interclasse (ICC).*

Résultats : La fourchette de pourcentage de concordance sur la sélection de niveau vertébral était

¹ D'Youville College, Buffalo, NY

² Private practice, Etobicoke, Ontario, Canada

Corresponding author: Jesse Cracknell
 jessecracknelldc@gmail.com
 T: (905) 321-8346
 3134 Highbourne Cres., Oakville, Ontario, L6M 5H2

correlation coefficient for round 1 and round 2 was 0.79. The ICC (3,1) was 0.79 (round 1), and 0.70 (round 2).

Conclusion: *Less than optimal agreement on end vertebrae selection was found between observers. Intra- and inter-observer reliability of the Cobb Measurement was 'excellent' (round 1) and 'good' (round 2).*

(JCCA. 2015;59(3):261-268)

KEY WORDS: chiropractic, Cobb measurement, scoliosis, reliability

Introduction

Accurate initial and subsequent Cobb measurements are important in scoliosis management protocols. Such protocols are determined by the degree of scoliosis curvature, and the progression of these curves.¹ It has been established in the current literature, that ± 5 degrees or more of change on successive radiographs is clinically significant.² Oda et al. emphasized that patient management is based on curve progression as observed on serial radiographs.³ This is significant because in teaching facilities, many different observers may interpret these radiographs over the course of the management period. As such, decisions may be made or altered based on progressive changes as interpreted by different observers. Because radiographs significantly influence management decisions, it is essential to understand the limits of measurement accuracy as well as limits of the measurement techniques used.³

While the reliability of the Cobb Measurement by many licensed health care professionals is well established, to our knowledge, reliability of the measurement by chiropractic interns has never been published. The purpose of this study is to evaluate the intra- and inter-observer reliability of Cobb angle measurement on digital radiographs by chiropractic interns.

Methods

The convenience sample used in this study consisted of 15 volunteer observers. Of the 15 original volunteers, one volunteer withdrew from the study before beginning while

de 0,36 à 0,79. Le coefficient de corrélation de Pearson r pour la première séance et la deuxième séance d'observations était de 0,79. Le coefficient interclasse (3,1) était de 0,79 (séance 1), et de 0,70 (séance 2).

Conclusion : *Le consensus sur la sélection des vertèbres limites a été moins qu'optimal entre les observateurs. La fiabilité des observations individuelles et entre ces observations de la mesure Cobb a été « excellente » (séance 1) et « bonne » (séance 2).*

(JCCA. 2015;59(3):261-268)

MOTS CLÉS : chiropratique, mesure Cobb, scoliose, fiabilité

the remaining 14 completed the study in its entirety. All volunteers were chiropractic interns studying at the same chiropractic program in the United States. This study was granted full approval by the Institutional Review Board of D'Youville College on August 20, 2013.

The study took place over a 22-day period. Of the 14 observers, 13 completed their second round of measurement 14 days after their initial round. One observer could not complete their final round of measurement for an additional 8 days, resulting in a 22 day rather than a 14 day interval between readings. This convenience sample represented more than 75% of the chiropractic interns enrolled in the institution of study.

Interns were instructed to view on PACS ten pre-selected anonymous digital radiographs, previously determined by the researchers to have scoliosis. All images were DICOM format and were displayed and measured on an AMD Catella™ PACS system with high resolution 2K monitors. Representative cases were selected from an archive database of anonymous chiropractic patients by two experienced chiropractic radiologists and the primary researcher. Inclusion criteria included a) adequate image quality; b) obvious scoliosis above a minimum of 10 degrees; and c) conspicuity of both end vertebrae. Participants were instructed to perform a Cobb Measurement on a PACS digital display program. Each participant measured the Cobb Angle on a frontal thoracic, lumbar or full spine radiograph. They identified the cephalic and caudal end vertebra defined

Table 1.
Percentage Agreement on Vertebra Selection between Observers

		Agreement of Vertebrae Selection Between Observers									
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
Cephalic End											
Vertebra	Round 1	0.57	0.43	0.50	0.50	0.50	0.50	0.64	0.43	0.50	0.57
	Round 2	0.57	0.57	0.71	0.50	0.79	0.50	0.64	0.57	0.64	0.57
Caudal End											
Vertebra	Round 1	0.64	0.50	0.36	0.64	0.57	0.71	0.64	0.43	0.43	0.36
	Round 2	0.43	0.71	0.57	0.71	0.43	0.64	0.50	0.43	0.50	0.36

Observer agreement on vertebral levels (Round 1): 52%

Observer agreement on vertebral levels (Round 2): 57%

Combined observer agreement on vertebral levels (Round 1 and Round 2): 54%

as those vertebral segments at the superior and inferior end of the curvature respectively that would result in the maximum angle. A transverse line was constructed along the superior endplate of the cephalic end vertebral body, and another transverse line was constructed along the inferior endplate of the caudal end vertebral body. The angle between the two endplate lines was then automatically calculated by the Cobb angle application on the PACS program. These Cobb angles were then recorded by the primary researcher. The researcher recorded the resultant angle, and the cephalic and caudal end vertebrae selected by each observer. These values were recorded in a spreadsheet for later analysis. The interns participated in two separate sessions, measuring the same ten radiographs once in each session. Participants were blinded to the identity of the cases, and the original findings were not disclosed during the second session.

Data were analyzed using the ICC (3,1) to determine inter-observer reliability of the Cobb Measurement. For the purposes of this study, it was necessary to utilize the form of ICC that was best for the analysis of single measurements between observers, rather than that which best evaluates the mean of several observer measurements.⁴ Gstoettner et al.⁵ suggest a grading scale using ICC results in regard to Cobb Measurement reliability, such that scores below 0.40 are regarded as poor; scores of 0.40-0.59 are considered fair; scores from 0.60-0.74 are good; and scores from 0.75-1.00 are excellent. These evaluative guidelines were used in the analysis of the results follow-

ing the observations. The confidence intervals for the ICC are reflective of the sample size and are included to assist the reader in understanding the precision of the estimate. If the variance in the sample stayed constant, increasing the sample size would reduce the confidence intervals. Please see the limitations section.

The Pearson *r* correlation coefficient was used to determine the intra-observer correlation of the Cobb Measurement. The Pearson *r* is most appropriate for continuous variables within the same class.⁴

Finally, cephalic and caudal end vertebrae selection was evaluated using the percentage of agreement between observers. In addition, the standard error (SE) for each case was calculated with both the 95% confidence interval and 99% confidence interval. Once the pool of data were collected from all participants, it was transferred to a master spreadsheet. All data were then analyzed with the psych package for the statistical software R (R Core Team (2013)) in preparation for ICC calculation, and coefficient correlation.⁶

Results

The sample included 11 males, and 3 females. Five observers were between 20-25 years of age, 8 observers were between 26-30 years of age, and one observer was over 30 years of age.

Inter-observer percentage agreement of cephalic end vertebra ranged from 36-71%, and caudal end vertebra ranged from 36-79% (Table 1). With regard to vertebral

levels most commonly selected, 10 of the 14 observers agreed with each other on the same cephalic vertebra in only 2 cases, and the caudal end vertebra in 2 cases. There was no instance where at least 10 observers agreed on the same cephalic and caudal end vertebra in the same case. Inter-observer agreement was 100% on the caudal and cephalic vertebra 0/10 times in round 1, and 0/10 times in round 2. Inter-observer agreement on vertebral levels occurred in 52% of cases in the first round, and 57% in the second round. Overall, when combining the first and second round, inter-observer agreement on vertebral levels occurred in 54% of cases.

Cephalic vertebra selection: In round 1, the highest level of inter-observer agreement on cephalic vertebra selection in a single case ranged from 0.43 (cases 2 and 8) to 0.64 (Case 6). In round 2, the highest level of inter-observer agreement on cephalic vertebra selection in a single case ranged from 0.50 (case 4) to 0.79 (case 5).

Caudal vertebra selection: In both rounds 1 and 2, the

highest level of inter-observer agreement on caudal vertebrae selection in a single case was 0.71 (case 6 in round 1 and cases 2 and 4 in round 2). Also, in both rounds 1 and 2, the lowest level of inter-observer agreement on caudal vertebrae selection in a single case was 0.36 (cases 2 and 10 in round 1, and case 10 in round 2).

Intra-observer reliability: The combined round 1 and round 2 intra-observer average correlation as estimated with Pearson *r* was 0.79 (excellent).

Inter-observer reliability (Table 2): Inter-observer results of round 1 were 0.79 (excellent) (95% confidence interval between 0.62 – 0.93). Inter-observer results of round 2 were .70 (good) (95% confidence interval between 0.50 – 0.89).

Standard deviation (SD) (Table 3): The average SD, calculated for each case between observers, was 6.3 degrees. The largest SD was case 6 (10.15 degrees) and the lowest was case 4 (2.21 degrees).

Range of Cobb Measurements (Table 4): The range between the largest and smallest Cobb Measurements recorded for case 1 in round 1 was 12.56 degrees, and in round 2 it was 29.81 degrees (a difference of 17.25 degrees). The smallest range recorded (8.79 degrees) was for case 4, in round 1 and this was the only case in either round where the range was less than 10 degrees.

Standard Error (SE) (Table 3): SE was calculated for all cases within 95% and 99% confidence intervals. In round 1, the 95% confidence interval ranged from ±1.16 degrees to ±5.31 degrees. The SE within a 95% confidence interval in round 2 ranged from ±1.86 degrees to ±5.25 degrees. The 99% confidence interval in round 1 ranged from ±1.52 degrees to ±6.98 degrees and in round 2 between ±2.45 degrees and ±6.90 degrees.

Table 2.
Intra- and Inter-Observer Reliability

Cobb Angle Correlation Statistics		
		Pearson r Correlation Coefficient
Intra-observer Reliability	Round 1 & Round 2	0.79
		Inter-observer ICC (3,1)
Inter-observer Reliability	Round 1	0.79 (95% CI: 0.62 – 0.93)
	Round 2	0.70 (95% CI: 0.50 – 0.89)

Table 3.
Standard Error (SE) and Standard Deviation of Observer Cobb Angles

		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
Round 1 SE	95% CI	±2.37°	±3.27°	±3.68°	±1.16°	±2.90°	±5.31°	±3.55°	±3.92°	±3.37°	±3.48°
	99% CI	±3.12°	±4.30°	±4.84°	±1.52°	±3.81°	±6.98°	±4.66°	±5.15°	±4.43°	±4.58°
Round 2 SE	95% CI	±4.37°	±2.94°	±4.00°	±3.84°	±3.29°	±5.25°	±1.86°	±4.33°	±2.08°	±2.33°
	99% CI	±5.74°	±3.86°	±5.25°	±5.05°	±4.33°	±6.90°	±2.45°	±5.69°	±2.73°	±3.06°
<i>Standard Deviation</i>	<i>Average</i>	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
	6.30	4.54	6.23	7.02	2.21	5.52	10.15	6.77	7.48	6.43	6.68

Table 4.
Range of Observer Cobb Measurements

		Range of Observer Cobb Measurements									
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
Round 1	Greatest Angle	39.16°	48.22°	46.48°	22.92°	26.28°	53.56°	31.48°	53.21°	43.75°	44.83°
	Lowest Angle	26.60°	29.94°	28.39°	14.13°	7.57°	18.09°	6.02°	23.58°	23.85°	19.86°
	Range	12.56°	18.28°	18.09°	8.79°	18.71°	35.47°	25.46°	29.63°	19.90°	24.97°
Round 2	Greatest Angle	41.14°	51.31°	49.60°	32.56°	36.75°	57.32°	32.83°	54.34°	43.10°	43.37°
	Lowest Angle	11.33°	30.60°	24.98°	7.71°	14.94°	29.12°	20.09°	28.93°	30.01°	28.01°
	Range	29.81°	20.71°	24.62°	24.85°	21.81°	28.20°	12.74°	25.41°	13.09°	15.36°

Discussion

In a 2003 survey, 66.9% of chiropractors reported that they had diagnosed a structural scoliosis and 66.0% reported that they had diagnosed a functional scoliosis in their previous year of practice.⁷ The Scoliosis Research Society has established the Cobb method as the standard of measurement to evaluate scoliotic curves and their progression, because it is both simple to perform, and accurate when evaluating repeated measurements.³ A large body of literature addressing the issues of Cobb Measurement variability and measurement reliability both on an intra- and inter-observer level has been published. This literature offers insight into the Cobb Measurement, variables that affect the proficiency and accuracy of the measurement, and the variability between measurements (inter- and intra-observer reliability).

An accurate Cobb Measurement is important because of the implications that the Cobb Measurement may have in management protocol, which is determined by the degree of curve progression between radiographs.¹ Because digital radiography is rapidly replacing conventional radiography in clinical practice, we used digital radiography to examine the reliability of the Cobb Measurement by chiropractic interns. All areas of investigation included cephalic end vertebrae selection, caudal end vertebrae selection, as well as intra- and inter-observer reliability analysis of actual Cobb angles.

Oda et al.³ identified variation in measurement attributed to the selection of end vertebra, measurement accuracy, and variability in measurement technique. The results of the study point to true error of measurement between radio-

graphs on repeated readings to be ± 9 degrees, attributing the wider range of variability to end vertebra selection by the observer.³ In this study, intra-observer and inter-observer error was 12.61 degrees, and 7.57 degrees respectively.

In situations where the selection of end vertebra was left to be determined by the observer, it was found that 4.2% of Cobb Measurements had more than 5 degrees of variation.² In a study of intra-observer and inter-observer variation of scoliosis by Carman et al., participants included four orthopedic surgeons and one physical therapist who observed 8 scoliosis images.¹ The participants measured each radiograph randomly in two sessions with a two-week interval between sessions.¹ While the degree of variability in this study resembled the Oda et al.³ findings, variations were not quite as high. Carman et al.¹ determined the mean SD to be 2.97 degrees, compared to Oda et al.³, which was 4.49 degrees.

In one reliability study, Gsteottner et al.⁵ evaluated and compared the Cobb Measurement and end vertebra selection on conventional radiographs and digital radiographs. The Gsteottner study found that Cobb Measurement coefficient variance (CV) was dependent on which medium the measurement was obtained and that measurement reliability varied depending on whether the measurements were performed on conventional or digital radiographs.⁵ Of special relevance to this study, Gsteottner et al.⁵ found that intra-observer selection of end vertebra on conventional radiographs was 'excellent', while it was only 'good' digitally. Inter-observer reliability was found to be 'good' on conventional radiographs and 'excellent' when measured on digital radiographs.⁵

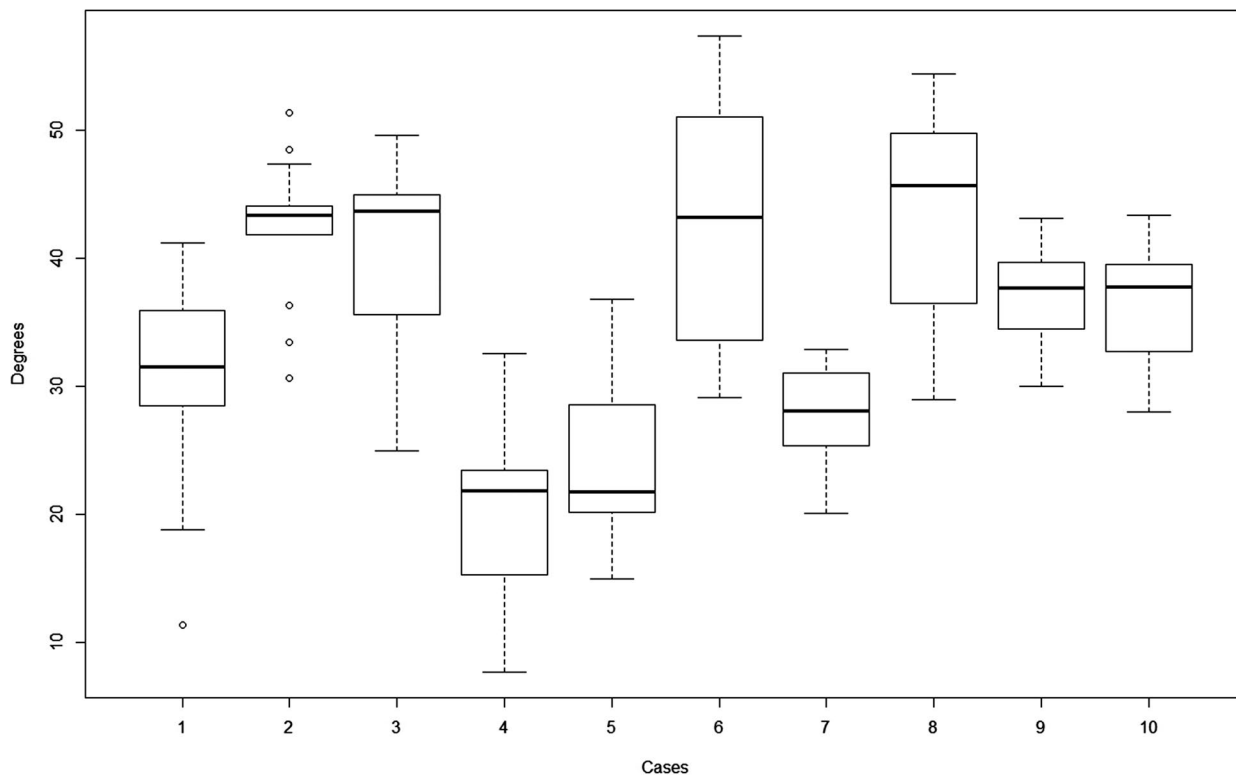


Figure 1.
Boxplot with Whisker Plots of Cobb Measurements

Figure 1. Demonstrates the Cobb Measurements with boxplot and whisker plots. For each box, the lower border is the 25th percentile and the top border is the 75th percentile. The dark line in the middle of the box is the 50th percentile (the median). The whiskers extend to the furthest data point which is within 1.5 times the interquartile range (from the 25th to the 75th). Data points beyond the whiskers are considered outliers and indicated as circles.

Beekman and Hall⁸ assessed variability in scoliosis measurement by two physicians, using ten radiographs. Carman et al. tested four orthopedic surgeons and one physical therapist, measuring eight separate scoliosis images.¹ Gstoettner et al. tested inter- and intra-observer reliability of six orthopedic surgeons.⁵ Despite the fact that Cobb Measurement reliability has been studied extensively in many licensed health care professionals, it has yet to be examined in interns in training who are still making important contributions to patient management decisions. Carmen et al. examined the clinical importance of observer error in an effort to determine acceptable limits of measurement and subsequent application of changes in patient management.¹ It was proposed that when five degrees or

less of measurement difference between radiographs is used to identify curve progression, approximately 30% of patients will meet this criterion because of observer error alone.

The review by Malfair et al.⁹ found that the major sources of error leading to variability are a product of radiographic quality, technique, and measurement error. The use of PACS to measure digital radiographs is purported to be equivalent in proficiency to manual measurements on analog conventional radiographs.⁹ In an error analysis of scoliosis measurement, it was established that Cobb Measurement error is also not a result of curve magnitude.¹⁰ In this study, Case 6 showed the largest interquartile range (Figure 1) and also the largest Cobb Measurements recorded which

indicates it was the curve with the largest magnitude. Case 4 however, recorded far lower Cobb Measurements but also had a large interquartile range. These findings are consistent with the above assertion that error in measurement is not a result of curve magnitude.

Selection of the incorrect end vertebrae has been identified by Gstoettner et al.⁵, Morrissy et al.², Shea et al.¹¹, and others as the most significant variable contributing to measurement error. There remains some debate about whether to include selection of end vertebrae in such reliability studies. Some researchers such as De Carvalho et al.¹², Morrissy et al.², and Shea et al.¹¹ elected to eliminate the selection of end vertebrae by the observers by having the researchers pre-select the end vertebrae prior to measuring. For the purpose of this study, we elected to not pre-select the end vertebrae. As such, the observers in our study (interns) were instructed to select the end vertebrae that they believed was most appropriate. While this added another potentially significant variable, we reasoned that this approach is more realistic and that it more accurately reflects the demands of real-life practice.

Intra-observer scores of six orthopedic surgeons using the digital mode of assessment as reported by Gstoettner et al.⁵ found that the mean ICC for proximal end vertebra to be 'good' (0.79), for the distal end vertebra to be 'good' (0.80), and the Cobb Measurement ICC to be 'excellent' (0.96). Inter-observer scores using the digital mode of assessment by Gstoettner et al. found that the mean ICC for proximal end vertebra to be 'good' (0.75), for the distal end vertebra to be 'poor' (0.73), and the Cobb Measurement ICC to be 'excellent' (0.93).⁵ These findings offer insight into the reliability findings when examining the same variables as proposed by this study. The main difference represented by the fact that all six Gstoettner et al.⁵ observer participants were experienced orthopedic surgeons proficient in Cobb Measurement (as opposed to inexperienced chiropractic interns).

There was little inter-observer agreement on cephalic and caudal end vertebra selection. Inter-observer agreement was only 52% in round 1, and 57% round 2. The combined percentage agreement of round 1 and round 2 was 54%. There was no case in round 1 or round 2 where inter-observer agreement was 100% on either or both of the same end vertebrae (cephalic and caudal) in the same case. There was also no case in round 1 or round 2 where inter-observer agreement was 100% on either the cephalic

or caudal vertebra in the same case. The combined range of inter-observer agreement in round 1 and round 2 on the cephalic end vertebra is 0.36 – 0.71, and 0.36 – 0.79 on the caudal end vertebra. As a result, it was concluded that inter-observer agreement on end vertebra is not strong. There is not a scale reported in the literature to report this, however the wide range of vertebral level selection and the low percentage agreement in most cases suggest that the observers' ability to identify the correct end vertebra was not strong.

Intra-observer reliability of the Cobb Measurement was estimated using the Pearson *r* correlation coefficient (see Table 2). The average intra-observer reliability was 0.79 (excellent) following round 1 and round 2 evaluations. These values imply that there is 'excellent' reliability of the assessment of the Cobb Measurement on an intra-observer level. Consequently, it may be concluded that chiropractic interns were effective and proficient in Cobb Measurements when each intern performs multiple Cobb Measurements on the same subject.

Many patients choose a chiropractic clinic for management of problems related to the spine. As a result, there is a need for chiropractors to become especially proficient in radiologic spinal measurement and assessments such as the Cobb Measurement. Therefore, there is a need to place further emphasis on the Cobb angle measurement as well as to assign more practice opportunities for chiropractic students and interns during the course of their education. This will better develop their proficiency and thereby better prepare interns for the challenges of treating patients professionally.

Limitations

This study was limited in that it only includes chiropractic interns enrolled at the same chiropractic college in the Northeastern United States. There were 17 total interns at the time of this study attending this chiropractic college, and 14 completed the study. The study could have been improved by increasing the sample sizes of images and of students. The study does, however, meet and at times exceed previous studies.

Conclusion

It was concluded that inter-observer reliability of the Cobb Measurement between chiropractic interns was 'good' to 'excellent'. If the premise is accepted that a 95%

confidence interval is acceptable in regard to the Cobb Measurement reliability, then the observers in this study were accurate and thus are unlikely to make incorrect management decisions based on poor radiographic analysis. However, it is likely that larger degrees of error will occur in chiropractic interns than in other more experienced health care professionals described in the literature such as orthopaedic surgeons.

There is a need for further research on the reliability of the Cobb Measurement in both chiropractic interns, and graduate chiropractors. This study was specific to chiropractic interns who attended the same chiropractic school, and have received the same chiropractic and radiologic education. It is suggested that this study be expanded to include a wider range of chiropractic interns with a broader representation of chiropractic schools. Such a study will provide a better understanding of the larger population of chiropractic interns and their proficiency in the Cobb Measurement.

Acknowledgement

The authors thank Dr. Ian McLean and Dr. Siri Leech at Palmer Chiropractic College for providing case material.

References

1. Carmen D, Browne R, Birch J. Measurement of scoliosis and kyphosis radiographs. *J Bone Joint Surg Am.* 1990;72(3):284-287.
2. Morrissy M, Goldsmith G, Hall E, Kehl D, Cowie G. Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error. *J Bone Joint Surg.* 1999;72:320-327.
3. Oda M, Rauh S, Gregory P, Silverman F, Bleck E. The significance of roentgenographic measurement of scoliosis. *J Pediatr Orthop B.* 1982;2(4):378-382.
4. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice* 3rd ed. New Jersey: Prentice Hall; 2009.
5. Gstoettner M, Sekyra K, Walochnik N, Winter P, Wachter R, Bach C. Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools. *Eur Spine J.* 2007;16(10):1587-1592.
6. Reville W. *Procedures for Personality and Psychological Research.* 2013. Northwestern University, Evanston, Illinois, USA, <http://CRAN.R-project.org/package=psych> Version = 1.4.5.
7. National Board of Chiropractic Examiners, Practice Analysis of Chiropractic 2010, <https://www.nbce.org/links/publications/practiceanalysis>.
8. Beekman C, Hall V. Variability of scoliosis measurement from spinal roentgenograms. *Phys Ther.* 1979;59(6):764-765.
9. Malfair D, Flemming A, Dvorak M, Munk P, Vertinsky A, Heran M, et al. Radiographic evaluation of scoliosis: review. *Am J Roentgenol.* 2010;194(3):S8-S22.
10. Gross C, Gross M, Kuschner S. Error analysis of scoliosis curvature measurement. *Bull Hosp Jt Dis Orthop Inst.* 1983;43(2):171-177.
11. Shea K, Stevens P, Nelson M, Smith J, Masters K, Yandow S. A comparison of manual versus computer-assisted radiographic measurement. *Spine.* 1998;23(5):551-555.
12. De Carvalho A, Vialle R, Thomsen L, Amzallag J, Cluzel G, Le Pointe HD, et al. Reliability analysis for manual measurement of coronal plane deformity in adolescent scoliosis. *Eur Spine J.* 2007;16:1615-1620.