ORIGINAL ARTICLE

Inter- and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools

Michaela Gstoettner · Katrin Sekyra · Nadja Walochnik · Peter Winter · Roland Wachter · Christian M. Bach

Abstract The objective of this study is to determine the intraobserver and interobserver reliability of end vertebra definition and Cobb angle measurement using printed and digital radiographs of 48 patients with scoliosis. The Cobb angle and the end vertebra were assessed by six observers in 48 patients with scoliosis using printed and digital radiographs. Definition of end vertebra and measurement of the Cobb angle was repeated three times with a 3 week interval. Intraclass correlation coefficients (ICC) were used to determine the interobserver and intraobserver reliabilities. 95% prediction limits for the errors in measurements are provided. For the Cobb angle a mean ICC of 0.97 was determined for intra- and interobserver reliability measurement of the printed radiographs. For the electronic radiographs a mean ICC value of 0.93 was determined for interobserver reliability and a mean ICC value of 0.96 for intraobserver reliability. Intraobserver ICC for definition of end vertebrae was 0.8 for both methods. Interobserver ICC was 0.83 for the manual and 0.74 in the digital method. One pitfall in angle measurement implies the Cobb method itself which measures in two dimensions. Until we develop a proper tri-dimensional measuring system an error is introduced. For the Cobb angle measurement the definition of end vertebrae introduces the main source of error. Digital radiography does not improve the measurement accuracy.

Keywords Intraobserver reliability · Interobserver reliability · Scoliosis · Cobb's angle · Radiography

M. Gstoettner · K. Sekyra · N. Walochnik ·
P. Winter · R. Wachter · C. M. Bach (✉)
Department of Orthopedic Surgery,
Innsbruck Medical University, Anichstrasse 35,
6020 Innsbruck, Austria
e-mail: christian.bach@i-med.ac.at

Introduction

Digital imaging substitutes are more and more for conventional X-ray films in the clinical routine. With the new possibilities of the digital radiography the expenses for developing the X-ray can be reduced as like for the storage of hardcopies. The new digital facilities allow a change of contrast and enlargement, a better vision of the vertebra morphology and to measure distances and angles easily.

The Cobb method is the standard method of quantization of the scoliosis angle [2]. It is an objective measure and is generally used to make decisions about the progression of a curve, as well as the need and success of treatment. Measurements of the Cobb angle bare an intra- and interobserver variability of approximately 4° to 8° [2, 4–6]. Sources of error are a wrong definition of end vertebra, an incorrect drawing of the lines through endplates or through the pedicles, drawing of perpendiculars or the measurement of the angle itself.

Our study was designed to determine the intra- and interobserver reliability of the Cobb angle measurement and for the assessment of the endvertebra using printed and digital radiographs. We investigated if tools used in digital radiography can improve the reliability of the Cobb method.

Methods

Postero-anterior radiographs of 48 patients with idiopathic and neuropathic scoliosis were used in this retrospective study. The image field in cranio-caudal direction ranged from the occiput to the acetabula. No patient was exposed to an additional study relating to radiation. All X-rays were stored digitally and printed out for manual measurements. The mean Cobb angle was 69.6° (range: 20° to 130°)(Table 1).

Table 1 Distribution of curve angles (Cobb method) in the 48 patients

Cobb's angle	Patients (%)
>20–40°	5
>40–60°	34.4
>60–80°	30.4
>80–100°	18.8
>100–120°	9.7
>120°	1.8
	100

Six examiners, all orthopedic surgeons and familiar with the measurement method of the Cobb angle, carried out the measurements independently for three times in each setting (manual measurement on films, digital measurement on the monitor), with a 3 week interval between each session. All observers were blinded to their prior measurements and to the other observers. There is a learning curve for digital measurement of the Cobb angle on the screen. However, because digital measurement is routinely used in the authors' hospital since 2002 all observers participating in the current study are already used to this technique.

In a first step end vertebrae were defined in the manual and the digital settings. In a second part of the study Cobb angles were measured with predefined end vertebrae: The cranial and caudal end vertebrae were marked by the senior orthopaedic spine surgeon on the same radiographs. Lines were drawn through the endplates of each vertebra, and the one with the largest angle to the horizontal was defined as an endvertebra.

For the manual set the printed radiographs were put on the light box, and the main angle was measured with pencil, ruler and protractor as described by Cobb [2]. No copies were used to avoid the loss of quality that can result from duplication. All films were blinded and numbered consecutively. The measured angle as well as the caudal and the cranial end vertebrae were recorded.

All digital images were stored in a Picture Archiving and Communication System (PACS) workstation. The radiographs were all blinded, numbered and viewed on the same monitor (Coronis® Color 3MP DL, Barco, Pres. Kennedypark 35, 8500 Kortrijk, Belgium). For the digital investigation the Icoview-software (icoserve information technologies GmbH, Austria) was used. Lines were drawn through the endplates of the upper and lower end vertebra of the curve. The program measured the Cobb angle automatically (Figs. 1–4).

Statistical analysis

The intraclass correlation coefficient (ICC) two-way mixed model on absolute agreement was used to analyse
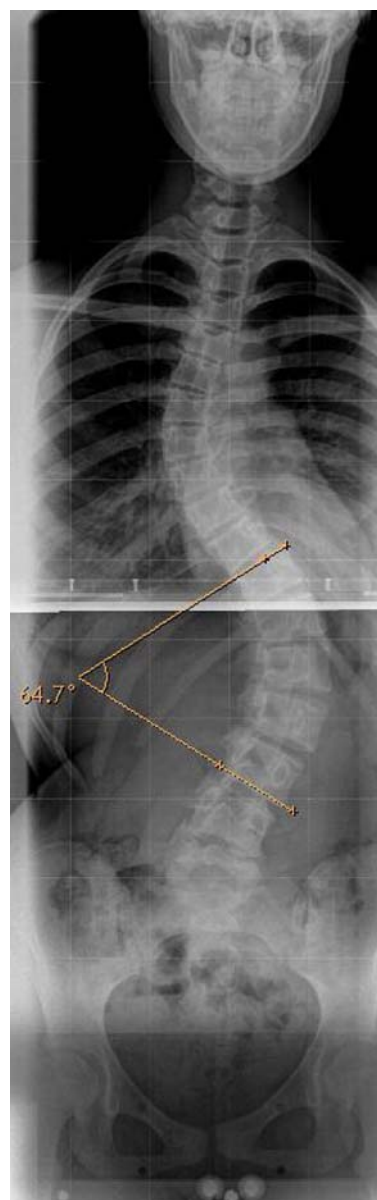

Fig. 1 Measuring the Cobb angle from the upper endplate of Th11 to the lower endplate of L3 results in an angle of 64.7°

measurement reliability [12]. It is essentially a ratio of the variance between subjects to the total variance (subjects, raters and error). The values of the ICC can range from 0 to 1, with a higher value indicating better reliability. ICC less than 0.40 was considered as poor; 0.40 to 0.59 as fair; 0.60 to 0.74 as good, and 0.75 to 1.00 as excellent [3]. The deviation of mean values was evaluated with the coefficient of variance (CV). Additionally, summary statistics from analyses-of-variance calculations were used to provide 95% prediction limits for the error in measurements. Statistical analyses were performed using SPSS 12.0 software (SPSS Inc., Chicago, IL, USA).
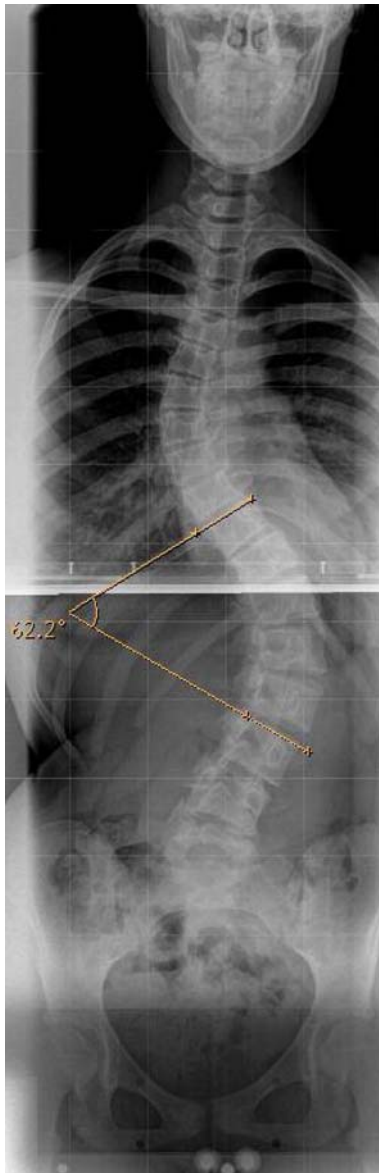
Fig. 2 Measuring the Cobb angle from the upper endplate of Th10 to the upper endplate of L3 results in an angle of 62.2°

Fig. 3 Measuring the Cobb angle from the upper endplate of Th11 to the lower endplate of L4 results in an angle of 101°

Results

The mean Cobb angle of all measured X-rays was 67.8° (range: 20°–123°) in the computer set and 71.1° (range: 20°–130°) in the manual set.

Both, the intrarater ($P = 0.031$) and interrater ($P = 0.036$) CV are significantly dependent on the method of measurement (manual vs. computer). The CV of intrarater ICC was 4.6% in the computer set and 5.1% in the manual set. The CV of interrater ICC was 7.2% in the computer set and 6.0% in the manual set.

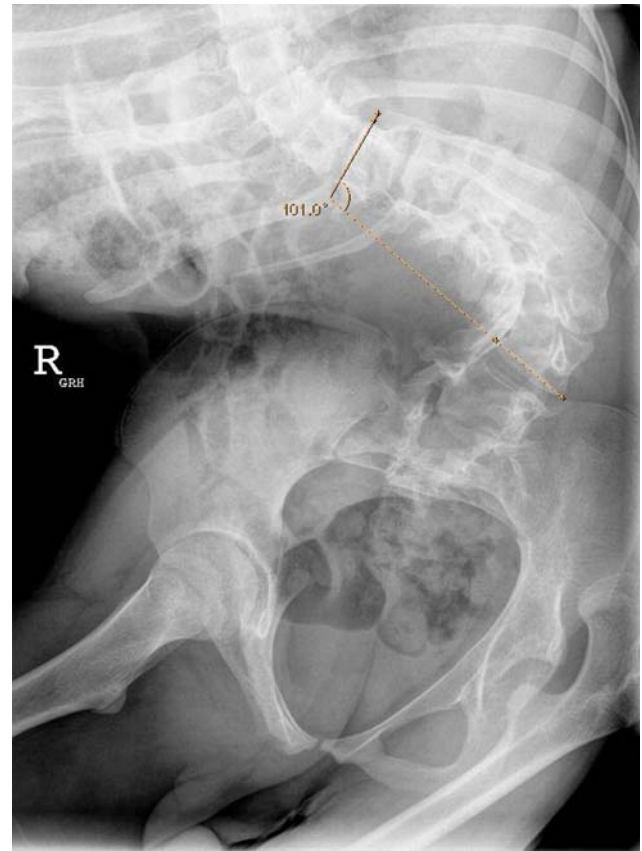For the definition of the end vertebrae the intraobserver ICC was excellent in the manual set, and good in the computer set. The interobserver ICC was excellent in the computer set and good in the manual set. Table 2 outlines the intra- and interobserver correlation of cobb angles.

The 95% prediction limit for each examiner indicated the difference between the six measurements of the Cobb angle and the definition of the end vetrebrae that would be exceeded approximately 5% of the time due to an error in measurement. Regarding the definition of end vertebrae, each rater varied between 1.91 and 3.03 vertebrae in the computer setting, whereas in the manual setting they varied between 1.6 and 2.8 (Table 3).

The 95% prediction limit regarding Cobb's angle was 9.038° (range: 4.87–14.53) in the computer setting and 7.68° (range: 6.41–8.60) in the manual setting. The interobserver prediction for the definition of end vertebrae and measurement of Cobb's angle was slightly better in the computer than in the manual setting.

Discussion

As treatment in patients with scoliosis is based on the degree of the curve, a precise measurement is crucial for the decisions of the treating physician [7, 13]. As modern
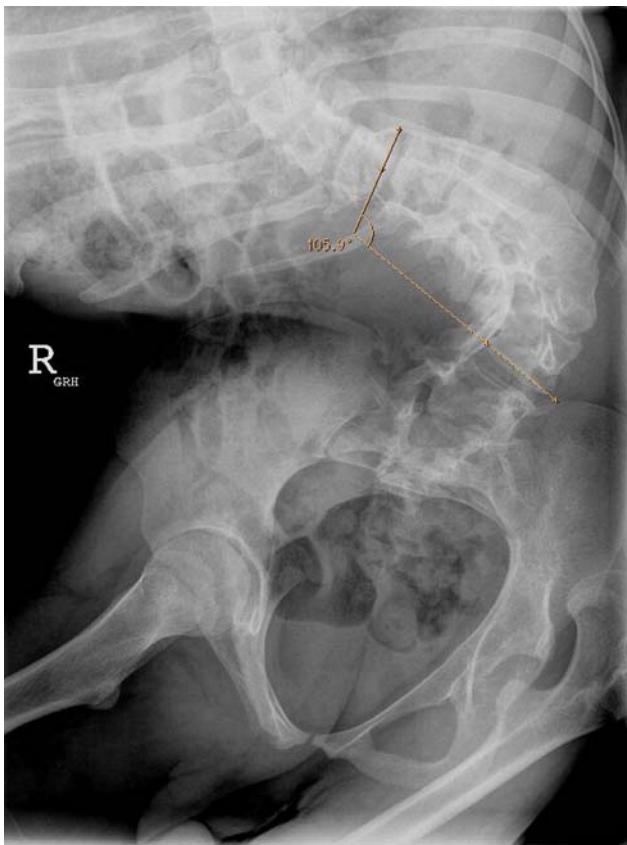
Fig. 4 Measuring the Cobb angle from the lower endplate of Th11 to the lower endplate of L4 results in an angle of 105.9°

digital imaging techniques have become widespread in most hospitals, we aimed to evaluate the reproducibility of Cobb angle measurement with digital and manual angle measurement tools.

High variabilities of the Cobb measurement may be due to wrong definition of the end vertebrae as well as defective angle measurement. Digital imaging has the advantages of rapid comparison between radiographs and cheap storage. Important parts of the spine can be enlarged and seen more clearly by changing the contrast, and the borders of the vertebrae can be enhanced by computerized options. After

drawing lines through the endplates of end vertebrae the software measures the angle automatically, which may reduce sources of error [11].

Despite this technical support, an inaccuracy in angle measurement remains due to a disadvantage of the Cobb method itself. We measure a two-dimensional picture of a three-dimensional structure. The endplate does not have a trajectory as a single clear line when the angle of its plane is other than perpendicular to the plane of the film. Even in the best conditions, one line an image of an endplate that appears fusiform on the screen or on the film cannot be reduced Figs. 1–4).

We aimed to compare the intra- and interobserver reliability in the definition of end vertebrae as well as measuring the Cobb's angle in radiographs of scoliotic spines in a manual and a computer setting. Our data shows that some observers get on well with the computer method, and some measured more reliably with the manual method. However, the main source of error revealed to be the definition of end vertebrae. When the variability of selection of the end vertebrae was eliminated, the quantum of actual error in the measurements among the examiners was relatively small. The largest range of variability was detected in the intrarater ICC of proximal and distal end vertebra definition (proximal: 0.79, distal: 0.77) and in the interrater ICC of proximal end vertebra definition (mean: 0.75). For the definition of end vertebrae the manual setting was slightly more reproducible between the different observers. The 95% prediction interval was 2 vertebrae in the manual setting and 2.3–3 vertebrae in the computer setting.

We found no significant difference in the intra- or interrater reliability between the conventional and the digital method in Cobb's angle measurement. The intraobserver reliability found in this study was nearly the same in both methods (mean ICC computer method: 0.96; mean ICC manual method: 0.97). These data suggest that the use of digital measurements does not improve measurement accuracy of the Cobb angle. These findings are in contrast to those reported by Shea et al. [11], who found a statistically better correlation in digital, compared to manual

Table 2 Intraclass correlation coefficients (ICC) in manual and digital measurement

		Manual set		Digital set	
		Mean	Range	Mean	Range
Proximal end vertebra	Intrarater ICC	0.8020	0.7423–0.8896	0.7903	0.6856–1
	Interrater ICC	0.8541	0.8029–0.9172	0.7507	0.6977–0.8210
Distal end vertebra	Intrarater ICC	0.76803	0.6519–0.8744	0.80165	0.6872–1
	Interrater ICC	0.8011	0.7580–0.8357	0.7261	0.7154–0.7414
Cobb angle	Intrarater ICC	0.9701	0.9488–0.9833	0.9647	0.9021–0.9890
	Interrater ICC	0.9692	0.9658–0.9722	0.9317	0.9207–0.9383

ICC Intraclass correlation coefficients

Table 3 95% prediction limits in manual and digital measurement

	95% prediction limit	Manual set		Digital set	
		Mean	Range	Mean	Range
Proximal end vertebra	Intrarater	2.06	1.6–2.8	2.64	2.29–3.03
	Interrater	2.203	2.17–2.25	2.18	2.01–2.43
Distal end vertebra	Intrarater	1.93	1.41–2.51	2.26	1.91–2.89
	Interrater	1.93	1.88–1.97	1.75	1.61–1.95
Cobb angle	Intrarater	7.68°	6.41–8.60	9.038	4.87–14.53
	Interrater	6.82	6.38–7.16	6.34	4.69–7.97

measurement of curves. However the intra- and interobserver reliability results of this study are comparable with those from previous studies [2, 6, 8] (Table 4). Analysing these previous studies it can be shown that most authors choose scoliotic curves smaller than 40° which may not reflect the true every day clinical situation [4] (Table 1). Sixty percent of our patients had curve angles of more than 60°, whereas most of the former studies investigated curves between 20° and 40° [6, 7]. Therefore the relative error in the current measurements is small in relation to the low grade of curves of former studies. Nevertheless the current interobserver ICC are comparable to those of Morrissy et al. [6], who measured curves between 20° and 40°. Both studies revealed an interobserver 95% confidence interval of 6°–7°.

The correlation between the six observers and between the different times of measurement was good to excellent.

There was no improvement in measurement accuracy by using digital radiography.

Conclusions

One pitfall in angle measurement implies the Cobb method itself. Until we develop a proper tri-dimensional measuring system, no matter how good the antero-posterior and lateral imaging results are, it is still only a two-dimensional picture. For the Cobb angle measurement the definition of end vertebrae introduces the main source of error. Digital radiography does not improve the measurement accuracy.

Table 4 Comparison of results with former studies

Author	Range of Cobb angle	Intraobserver 95% prediction limit (mean value)	Interobserver 95% prediction limit (mean value)
Shea et al. [11]	20°–40°	2.6[oa, c]	NC
		3.3[ob, c]	NC
Sevastikoglou et al. [10]	NC	NC	3.12°
Morrissy et al. [8]	20°–40°	4.9°	7.2°
		3.8[oc]	6.3[oc]
Gross et al. [6]	4.5°–106.5°	2°–4°	2.3°
Beekmann and Hall [1]	5°–25°	NC	4.2°
Oda et al. [9]	NC	10°–30.2°	20.18°
Current study	20°–130°	9.038[oa]	6.34[oa]
		7.68[ob]	6.82[ob]

NC no comment

[a] Computer set

[b] Manual set

[c] Predefined end vertebrae

References

1. Beekmann CE, Hall V (1979) Variability of scoliosis measurement from spinal roentgenograms. Phys Ther 59:764–765
2. Carman DL, Browne RH, Birch JG (1990) Measurement of scoliosis and kyphosis radiographs. Intraobserver and interobser variation J Bone Joint Surg Am 72(3):328–33
3. Fleiss JL (1986) Reliability of measurement. In: Fleiss JL (ed) The design and analysis of clinical experiments. Wiley, Toronto, pp 1–32
4. Geijer H, Beckman K, Jonsson B, Andersson T, Persliden J (2001) Digital radiography of scoliosis with a scanning method: initial evaluation. Radiology 218(2):402–410
5. Geijer H, Verdonck B, Beckman KW, Andersson T, Persliden J (2003) Digital radiography of scoliosis with a scanning method: radiation dose optimization. Eur Radiol 13(3):543–551 (Epub14 June 2002)
6. Gross C, Gross M, Kuschner S (1983) Error analysis of scoliosis curvature measurement. Bull Hosp Jt Dis Orthop Inst 43(2):171–177
7. Lonstein JE, Carlson JM (1984) The prediction of curve progression in untreated idiopathic scoliosis during growth. J Bone Joint Surg Am 66(7):1061–1071
8. Morrissy RT, Goldsmith GS, Hall EC, Kehl D, Cowie GH (1990) Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error J Bone Joint Surg Am 72(3):320–327
9. Oda M, Rauh S, Gregory PB, Silverman FN, Bleck EE (1982) The significance of roentgenographic measurement in scoliosis. J Pediat Orthop 2:378–382

10. Sevastikoglou JA, Bergquist E (1969) Evaluation of the reliability of radiological methods for registration of scoliosis. Acta Orthop Scand 40:608–613

11. Shea KG, Stevens PM, Nelson M, Smith JT, Masters KS, Yandow SA (1998) Comparison of manual versus computer-assisted radiographic measurement. Intraobserver measurement variability for Cobb angles. Spine 23(5):551–555

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

13. Weinstein SL, Ponseti IV (1983) Curve progression in idiopathic scoliosis. J Bone Joint Surg Am 65(4):447–455