Accuracy and Reliability of X-ray Measurements in the Cervical Spine

Catarina Marques[1], Emma Granström[1], Anna MacDowall[1],
Nuno Canto Moreira[2], Martin Skeppholm[3], Claes Olerud[1]

[1]Department of Surgical Sciences, Uppsala University, Uppsala, Sweden
[2]Pediatric Radiology Section, Department of Clinical Neuroscience, K8, Karolinska Universitetssjukhuset Solna, Stockholm, Sweden
[3]Department for Learning, Informatics, Management and Ethics, Medical Management Center, Health Economics and Health Evaluation Research Group,
Karolinska Institutet, Stockholm, Sweden

Study Design: This study is a post hoc analysis of a multicenter prospective randomized controlled trial which compared artificial disc replacement and anterior cervical discectomy and fusion.

Purpose: Useful radiographic parameters for assessing cervical alignment include the Cobb angles, T1 slope (T1S), occipitocervical inclination (OCI), K-line tilt (KLT), and cervical sagittal vertical axis (cSVA). This study aimed to determine measurement accuracy and reliability for these parameters.

Overview of Literature: Various authors have assessed repeatability by comparing different methods of measurement, but knowledge of measurement error and minimal detectable change is scarce.

Methods: We evaluated 758 lateral cervical radiographs. One medical student and one spine surgeon (i.e., measured ×2 within 4 weeks) independently measured the parameters obtaining 5,850 values. Standard error of measurement (SEm) and minimum detectable change (MDC) were calculated for each parameter. The accuracy and reliability of the Cobb angle measurements were calculated for the different types of angles: cervical lordosis, prosthesis angle, segmental angle with two bone surfaces (SABB), and segmental angle with one bone and one metal surface. Reliability was determined with intraclass correlation coefficient (ICC).

Results: SEm was 1.8° and MDC was 5.0° for the Cobb angle, with an intraobserver/interobserver ICC of 0.958/0.886. All the different subtypes of Cobb angles had an ICC higher than 0.950, except SABB (intraobserver/interobserver ICC of 0.922/0.716). The most accurate and reliable measurement was for KLT.

Conclusions: This study provides normative data on SEm and MDC for Cobb angles, T1S, KLT, OCI, and cSVA in cervical lateral radiographs. Reliability was excellent for all parameters except SABB (e.g., good).

Keywords: Cervical spine; Alignment; Cobb angle; T1 slope; Cervical sagittal vertical axis

Introduction

Cervical malalignment is a major source of pain, disability, and poor surgical outcomes, and may lead to cervical my-elopathy, as cervical deformity may cause spinal cord tension and compression [1-3]. Assessing cervical alignment is necessary in several situations, such as degenerative diseases, fractures, and postoperative status, among others.

Several radiographic parameters are useful in clinical practice when assessing cervical alignment: Cobb angle [4,5], T1 slope (T1S) [5], the more recently described occipito-cervical inclination (OCI) [6], K-line tilt (KLT) [7], and the cervical sagittal vertical axis (cSVA) [5]. The accuracy and reliability of these instruments is paramount knowledge for correctly interpreting any change in the obtained values and judging its importance. For instance, a change in a certain parameter before and after surgery may be meaningful or may be merely due to the inherent error of measurement, and therefore devoid of importance.

Although these measurements are widely used in clinical practice and repeatability has been assessed comparing different methods of measurement [8-10], knowledge of measurement error and minimum detectable change (MDC) is scarce [11]. Furthermore, the accuracy and reliability of the Cobb angle is based mostly on coronal plane measurements in the context of thoracolumbar scoliosis [12,13], or on sagittal plane measurements in the thoracic or lumbar spine [14-16]. Measurements of the cervical lordosis (CL) angle [11] and T1S [17] have been validated; however, to the best of our knowledge, the accuracy and reliability of other types of Cobb angles in lateral cervical films have not been studied.

This study aimed to provide normative data regarding the accuracy and reliability of measurements of the cervical lateral Cobb angle (and its subtypes) as well as T1S, OCI, KLT, and cSVA. It aimed to provide knowledge on measurement error (when measuring a radiographic parameter once), on MDC (when measuring the parameter twice, when comparing the same parameter in different positions such as flexion and extension, or in different moments, for instance before and after surgery), and to ascertain the agreement of repeated measurements (test-retest reliability).

Materials and Methods

1. Material and measurement

This study was a post hoc analysis of a prospective randomized controlled trial (RCT) involving 153 patients from three hospitals in Sweden who underwent surgery between 2007 and 2010 [18]. Patients with radiculopathy due to degenerative disc disease were randomized after exposure and decompression to either artificial disc replacement (ADR) (Discover; DePuy Spine, Johnson & Johnson, Raynham, MA, USA) or anterior cervical discectomy and fusion using autologous iliac crest graft. Inclusion and exclusion criteria as well as 2- and 5-year results have been previously published [18,19]. Plain flexion, extension, and neutral lateral cervical radiographs were acquired preoperatively and at 1, 2, and 5 years after surgery. Radiographs from the 126 patients present for radiological follow-up at 5 years postoperatively were evaluated. Because not all patients were present, and not all the incidences were performed at each follow-up, only 758 cervical lateral radiographs were available (Fig. 1).

Cobb angles, T1S, KLT, OCI, and cSVA were independently measured by two observers, which included a medical student and an experienced spine surgeon; the latter reassessed the measurements ×2 within 4 weeks
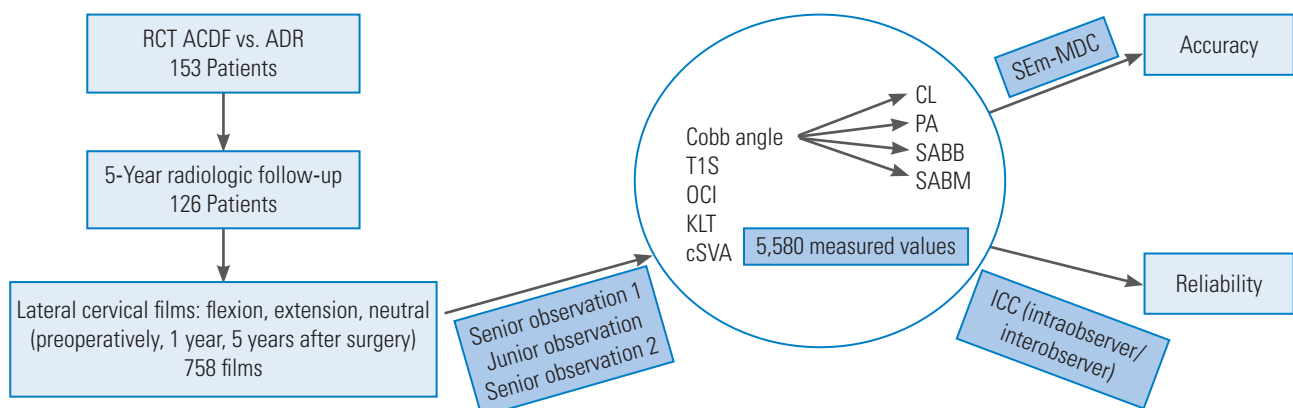
Fig. 1. Flowchart. RCT, randomized controlled trial; ACDF, anterior cervical discectomy and fusion; ADR, artificial disc replacement; T1S, T1 slope; OCI, occipitocervical inclination; KLT, K-line tilt; cSVA, cervical sagittal vertical axis; CL, cervical lordosis; PA, prosthesis angle; SABB, segmental angle with two bone surfaces; SABM, segmental angle with one bone and one metal surface; SEm, standard error of measurement; MDC, minimum detectable change; ICC, intraclass correlation coefficient.

(Fig. 1). The notation "NP" was used in cases where it was impossible to measure an angle or distance due to poor image quality and/or ill-defined anatomical landmarks. In total, 5,850 numerical values were obtained. Only pairs of numerical values could be included in the accuracy and reliability analysis. All measurements were done using OsiriX software (Pixmeo, Geneva, Switzerland).

2. Radiographic parameters
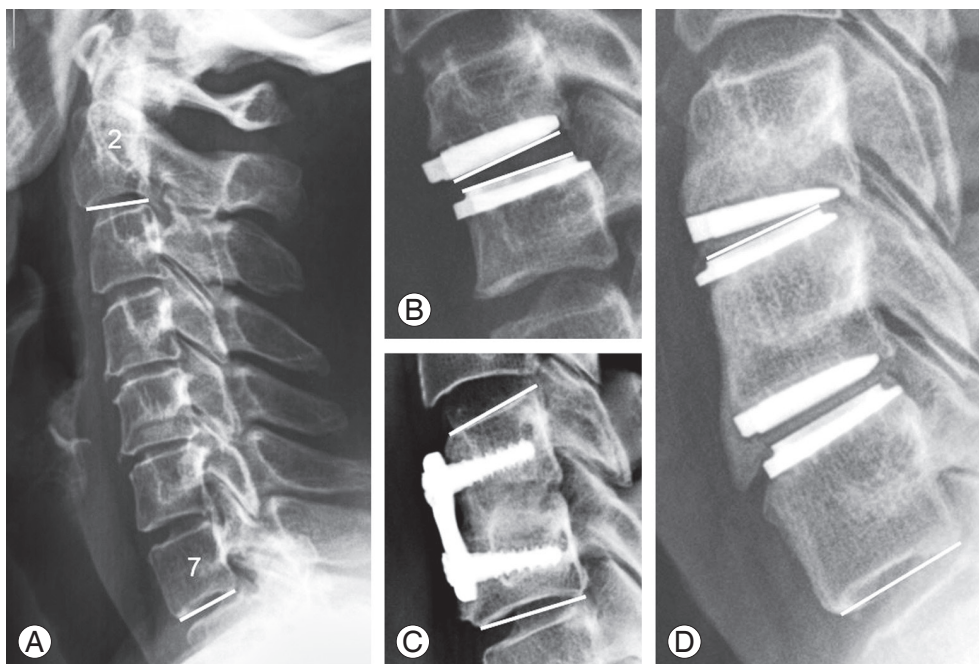
1) Cobb angles (Fig. 2)
CL, C2–C7 lordosis, or C2–C7 angle is the angle of intersection between two lines, drawn parallel to the inferior endplates of C2 and C7 (Fig. 2A). Prosthesis angle (PA) is the angle of intersection between two lines, drawn parallel to the superior and the inferior articulating surfaces of an artificial disc device (Fig. 2B). Segmental angle (SA) is the angle of intersection between two lines, drawn parallel to the superior endplate of the upper vertebra and to the inferior endplate of the lower vertebra in a specific spine segment (functional spine unit). In cases where two contiguous intervertebral discs had been replaced by arthroplasties there were two functional spine units; thus, two SAs were measured, one for each operated level. In those cases, one of the lines was drawn on the vertebral endplate

and a second line was drawn on the (metal) surface of the artificial disc device. Therefore, two different types of SAs were studied: (1) SA with two bone surfaces (SABB)—in fusions (regardless of the number of fused levels) and in one-level ADR (Fig. 2C); (2) SA with one bone and one metal surface (SABM)—in two-level ADR (Fig. 2D).

2) Other parameters (Fig. 3)
T1S is the angle of intersection between a line drawn at the superior endplate of the first thoracic vertebra and a horizontal line (Fig. 3A). OCI is the anteroinferior angle at the intersection between the McGregor's line and a line drawn parallel to the posterior border of C4 (Fig. 3B). KLT is the angle of intersection between the K-line (a line connecting the middle of the spinal canal in C2 and C7) and a vertical line (Fig. 3C). cSVA, or C2–C7 SVA, is the horizontal distance between the posterior aspect of the superior endplate of C7 and a plumbline taken from the center of the vertebral body of C2 (Fig. 3D).

Accuracy and reliability were calculated for each variable (Cobb angles, T1S, KLT, OCI, and cSVA). In addition, the accuracy and reliability of the Cobb angle measurements were calculated for the different types of Cobb angles (CL, PA, SABB, and SABM) (Fig. 1).

Fig. 2. Cobb angles. **(A)** Cervical lordosis. **(B)** Prosthesis angle. **(C)** Segmental angle with two bone surfaces (in anterior cervical discectomy and fusion and one-level ADR). **(D)** Segmental angle with one bone and one metal surface (in two-levels ADR). ADR, artificial disc replacement.

3. Statistical analysis

Mean difference was calculated to screen for systematic errors (whether one observer systematically measured larger or smaller values than the other). Accuracy was determined in terms of standard error of measurement (SEm) and MDC. SEm, the typical measurement error, was calculated with a random effects model, with the patient as a random effect. MDC ($1.96 \times \sqrt{2} \times$ SEm) is the smallest change in a true value which can be detected with 95% statistical confidence, taking measurement error into account.

Intraobserver and interobserver reliability were evaluated by the intraclass correlation coefficient (ICC). In addition, Bland-Altman plots were obtained. For SEm, MDC, and ICC, 95% confidence intervals were computed using bootstrap with 10,000 bootstrap replicates [20]. All statistical analyses were performed in R ver. 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria) [21].

4. Trial oversight

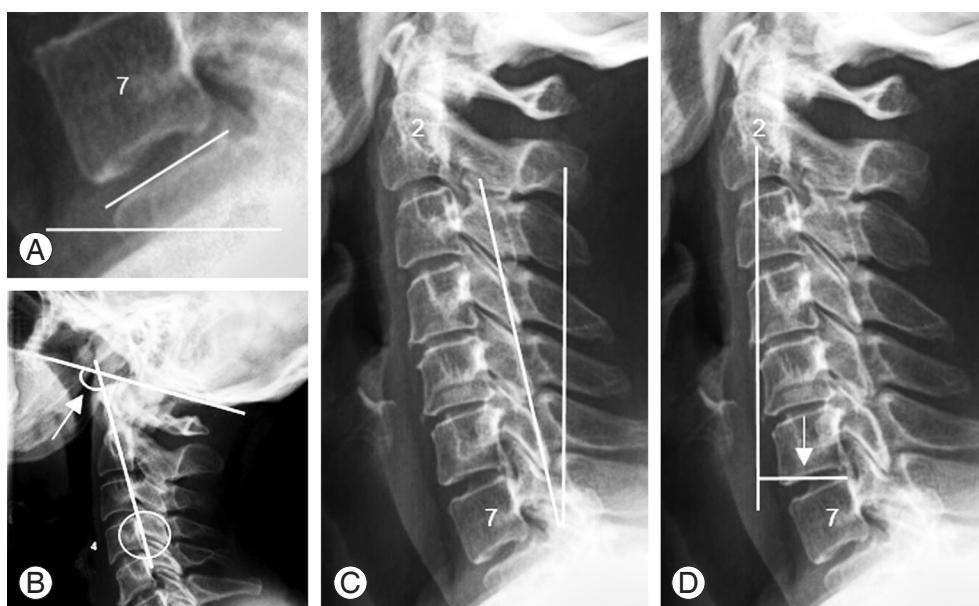The original trial was approved by the local Swedish eth-

Fig. 3. (A) T1 slope. **(B)** Occipitocervical inclination. **(C)** K-line tilt. **(D)** Cervical sagittal vertical axis.

Table 1. Accuracy and reliability of X-ray measurements in the cervical spine

Variable	SEm (95% CI) pv	MDC (95% CI) pv	Intraobserver ICC (95% CI) pv	Interobserver ICC (95% CI) pv	NP (%)
Cobb angle	1.8 (1.7–1.9) 1,171	5.0 (4.7–5.4) 1,171	0.958 (0.949–0.966) 1,171	0.886 (0.858–0.907) 1,190	15
CL	2.2 (1.9–2.6) 143	6.2 (5.1–7.3) 143	0.966 (0.948–0.978) 143	0.933 (0.903–0.953) 143	26
PA	1.5 (1.3–1.7) 422	4.2 (3.6–4.7) 422	0.951 (0.932–0.965) 422	0.974 (0.965–0.980) 429	4
SABB	2.0 (1.8–2.2) 417	5.5 (5.0–6.1) 417	0.922 (0.895–0.941) 417	0.716 (0.632–0.782) 425	20
SABM	1.6 (1.4–1.8) 183	4.5 (4.0–5.0) 183	0.974 (0.961–0.982) 183	0.922 (0.871–0.950) 187	18
T1S	2.0 (1.5–2.5) 95	5.6 (4.2–6.9) 95	0.929 (0.871–0.963) 95	0.838 (0.749–0.892) 95	54
OCI	2.9 (1.3–4.7) 68	8.2 (3.5–13.0) 68	0.925 (0.790–0.986) 68	0.909 (0.762–0.977) 54	43
KLT	0.7 (0.6–0.8) 248	1.8 (1.6–2.1) 248	0.988 (0.984–0.992) 248	0.961 (0.945–0.973) 263	12
cSVA	1.3 (0.7–2.0) 276	3.7 (2.0–5.4) 276	0.987 (0.970–0.996) 276	0.978 (0.963–0.988) 276	7

SEm and MDC values for the Cobb angles, T1S, OCI, and KLT in degrees; SEm and MDC values for cSVA in millimetres.
SEm, standard error of measurement; CI, confidence interval; pv, pairs of measured values; MDC, minimum detectable change; ICC, intraclass correlation coeficient; NP, impossible to measure due to poor quality X-ray and/or ill-defined anatomical landmarks; CL, cervical lordosis; PA, prosthesis angle; SABB, segmental angle with two bone surfaces; SABM, segmental angle with one bone and one metal surface; T1S, T1 slope; OCI, occipitocervical inclination; KLT, K-line tilt; cSVA, cervical sagittal vertical axis.

ics review board and all participants provided oral and written informed consent. The study was registered at IS-RCTN (registration no., 44347115).

Results

Screening for systematic errors was negative (negligible mean differences) for all parameters, except T1S and SABB, where the junior observer systematically obtained larger values by 3° for T1S, and a nearly 4° smaller value for SABB than the senior observer. Results obtained for all parameters are displayed in detail in Table 1. SEm was 1.8° and MDC was 5.0° for the Cobb angle group, with an intraobserver ICC (IaCC) of 0.958 and an interobserver ICC (IeCC) of 0.886 (Table 1). Regarding Cobb angle subtypes, the most accurate was PA (SEm=1.5°, MDC=4.2°) and the least accurate was CL (SEm=2.2°, MDC=6.2°). All different subtypes of Cobb angles had an IaCC higher than 0.950, except for SABB (IaCC=0.922, IeCC=0.716) (Table 1). Overall, the most accurate measurement was for KLT, with a SEm of 0.7° and MDC of 1.8°, and the least accurate measurement was for OCI, with a SEm of 2.9° and MDC of 8.2° (Table 1). The most reliable measurements were for KLT (IaCC=0.988) and cSVA (IeCC=0.978), whereas the least reliable were for the aforementioned SABB, followed by OCI (IaCC=0.925) and T1S (IeCC=0.838) (Table 1). More than half of the attempts to measure T1S failed due to poor image quality and/or ill-defined anatomical landmarks, as did over one third of attempts to measure OCI and one quarter of attempts to measure CL. The success-

ful measuring rate was 80% or higher for the remaining parameters (Table 1). Fig. 4 shows the Bland-Altman plots for the Cobb angle group.

Discussion

Exact knowledge of the accuracy and reliability of any measurement is paramount in order to interpret the obtained values. This validation study is the first to address measurement accuracy and reliability for the different subtypes of Cobb angles in the cervical spine, as well as T1S, OCI, KLT, and cSVA, in terms of both The standard error of the measurement (SEm) and MCD.

Because SEm and MDC are the most relevant in clinical practice and clinical research, accuracy was ascertained using them. The literature is somewhat confusing regarding measurement errors, because it is often not clear if the error the articles refer to is SEm or MDC, which are quite different entities, to be considered in different situations. It is therefore important to understand the difference between those two entities. SEm represents by how many degrees or millimeters one *single* measurement can be wrong. MDC is the smallest change in true value that can be detected with 95% statistical confidence, taking measurement error into account. Thus, MDC is the minimum change that can be detected when *comparing* two images. When an angle, inclination, or distance is measured twice there are two errors to consider, one error for each time the measurement is taken. Therefore, MDC contains both errors and is larger than the error of a single measure-
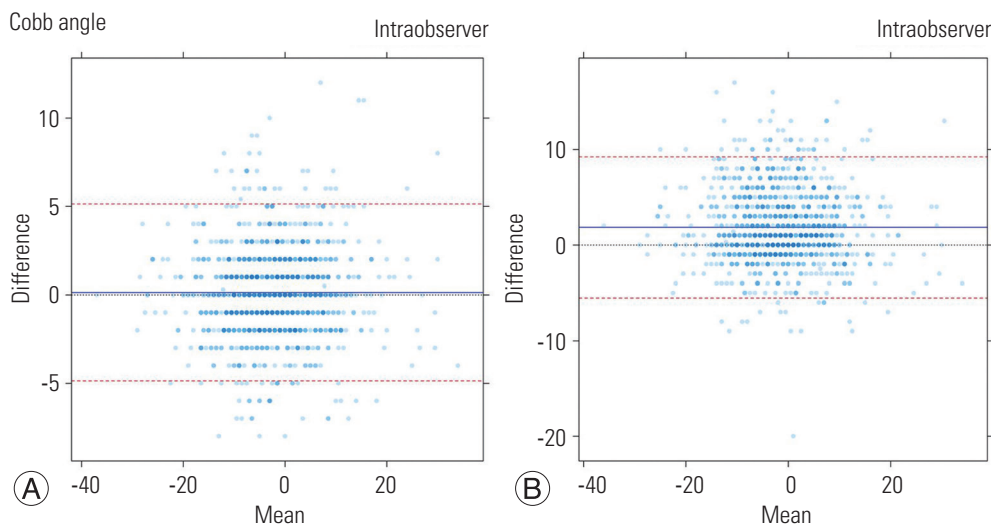

Fig. 4. (A, B) Bland-Altman plot for the Cobb angle (°). Dotted line: perfect agreement; solid line: mean difference; dashed lines: limits of the reference region containing 95% of the points.

ment. Hence, MDC should be considered when assessing progress through time (before and after treatment, or when assessing the natural progress of a deformity), or when checking for movement or movement-related dislocation (comparing flexion-extension images).

With that in mind, it becomes clear that the most interesting error value to consider for OCI, is SEm, as the parameter is measured only once in the operating room; e.g., when positioning a patient for craniocervical fixation; whereas, for a segmental Cobb angle, when checking for movement in a prosthetic disc or fusion, MDC is the error value to be considered. The error value to consider depends on what exactly is being assessed. Both SEm and MDC are important and should be well differentiated from one another, and appropriately used.

In this study, the SEm and MDC values were calculated for intraobserver comparisons. That decision, based on the assumption that in clinical and research practice the values are generally measured by the same observer (when examining a control X-ray one should always check the previous X-ray; and the flexion-extension comparisons are done by one and the same observer), was supported by the excellent interobserver reliability (ICC) obtained.

One of the strengths of this study is the fact that it was based on a very large number of measurements, nearly 6,000—over 3,500 for the Cobb angle alone—and, even when broken into the different parameters studied separately, there still remained a large number to be analyzed for each parameter. The patient group was homogeneous, all patients had the same diagnosis, and the radiographs were obtained under similar conditions for every patient.

When measuring Cobb angles in the context of scoliosis, each observer is usually entitled to decide which vertebrae to include in the segment. That accounts for a large variability in the error among different studies, depending on how much freedom is allowed in placing the lines [12]. One strong feature of this study is the fact that when measuring any parameter, it was pre-determined and clear where the lines were to be drawn, which leads to robust results.

As for limitations, it can be argued that there are too few observers in this study. However, repeatability was excellent even though one of the observers is an experienced spine surgeon, whereas the other is a medical student with null experience.

Image quality in some radiographs was suboptimal and/or the anatomical landmarks were ill-defined, leading to

impossibility to measure (and therefore smaller samples) or, when the observers deemed it possible to measure, presumable larger errors and lower agreement. That was particularly true for T1S and OCI. The upper endplate of T1 (T1S) was often not seen and the same was true for the hard palate (OCI). The latter is explained by the fact that OCI assessment was not a part of the original RCT, so the radiographs were not meant to include that structure. Therefore, we expect the error and minimum change for the OCI to be lower in real-life operating room situations (which is where the OCI is useful). Future research should address that issue.

Difficulty in finding landmarks in the lower cervical spine also explains the lower accuracy of CL measurement as compared with the other Cobb angles. Poor definition of the anatomical landmarks, credibly due to degeneration at adjacent levels, also accounts for the fact that the error (SEm and MDC) is larger and agreement is worse, the less metal and more bone structures are involved in the measurements. Therefore, PA was the most, and SABB was the least accurate and reliable among the SAs; therefore, the existence of a systematic error for SABB. When a systematic error occurs, SEm and MDC are not very meaningful, since they are based on the assumption that there are no systematic differences. However, ICC is still meaningful.

The most important finding of this study along with the exact knowledge of the errors, minimum changes, and reliability for the parameters commonly used when assessing cervical alignment, is that the magnitude of the error is clearly different for different types of Cobb angles. Thus, Cobb angles should not be considered as a whole in terms of error. Each Cobb angle has its own SEm and MDC, and that should be taken into account when evaluating a change in one given angle, be it CL, PA, or any other angle.

This study was conducted with patients who have degenerative disc disease, but the measured parameters are commonly used in several other pathologies such as trauma, tumors, or deformity. We have no reason to believe that the measurement errors would be different in those cases. The results are therefore probably extrapolatable to those pathologies as well.

Conclusions

This validation study provides normative data on SEm and MDC for Cobb angles (CL, PA, SABB, and SAMB), T1S,

OCI, KLT, and cSVA in lateral cervical X-rays. Reliability is excellent for all parameters, except SABB, for which it is merely good. We therefore suggest that this valid and reliable information on accuracy should be used when assessing and interpreting a change in cervical alignment in the context of degenerative disc disease and, with reasonable certainty, any other conditions where such parameters are used.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

ORCID

Catarina Marques: https://orcid.org/0000-0002-5530-2909
Emma Granström: https://orcid.org/0000-0002-6507-1948
Anna MacDowall: https://orcid.org/0000-0002-9231-7608
Nuno Canto Moreira: https://orcid.org/0000-0001-5503-1793
Martin Skeppholm: https://orcid.org/0000-0003-2228-2478

Author Contributions

Conception and design of work: CM, AMD, CO; acquisition, analysis, and interpretation of data: CM, EG; drafting the manuscript: CM; critical revision of the work for important intellectual content: CM, AMD, NCM, MS, CO; and final approval of the version to be published: CM, EG, AMD, NCM, MS, CO.

References

1. Scheer JK, Tang JA, Smith JS, et al. Cervical spine alignment, sagittal deformity, and clinical implications: a review. J Neurosurg Spine 2013;19:141-59.

2. Ames CP, Blondel B, Scheer JK, et al. Cervical radiographical alignment: comprehensive assessment techniques and potential importance in cervical myelopathy. Spine (Phila Pa 1976) 2013;38(22 Suppl 1):S149-60.

3. Smith JS, Lafage V, Ryan DJ, et al. Association of myelopathy scores with cervical sagittal balance and normalized spinal cord volume: analysis of 56 preoperative cases from the AOSpine North America Myelopathy study. Spine (Phila Pa 1976) 2013;38(22 Suppl 1):S161-70.

4. Cobb JR. Outline for the study of scoliosis. Instr Course Lect 1948;5:261-75.

5. Ames CP, Smith JS, Eastlack R, et al. Reliability assessment of a novel cervical spine deformity classification system. J Neurosurg Spine 2015;23:673-83

6. Yoon SD, Lee CH, Lee J, Choi JY, Min WK. Occipitocervical inclination: new radiographic parameter of neutral occipitocervical position. Eur Spine J 2017;26:2297-302.

7. Sakai K, Yoshii T, Hirai T, Arai Y, Matsukura Y, Okawa A. The K-line tilt, a novel radiographic parameter of cervical sagittal balance, is a predictor of postoperative kyphotic deformity after laminoplasty for cervical myelopathy caused by ossification of the posterior longitudinal ligament. Proceedings of the 44th Annual Meeting of the Cervical Spine Research Society; 2016 Dec 1-3; Toronto, Canada. Milwaukee (WI): Cervical Spine Research Society; 2017.

8. Janusz P, Tyrakowski M, Yu H, Siemionow K. Reliability of cervical lordosis measurement techniques on long-cassette radiographs. Eur Spine J 2016;25:3596-601.

9. Singhatanadgige W, Kang DG, Luksanapruksa P, Peters C, Riew KD. Correlation and reliability of cervical sagittal alignment parameters between lateral cervical radiograph and lateral whole-body EOS stereoradiograph. Global Spine J 2016;6:548-54.

10. Lee HD, Jeon CH, Chung NS, Kwon HJ. Comparative analysis of three imaging modalities for evaluation of cervical sagittal alignment parameters: a validity and reliability study. Spine (Phila Pa 1976) 2017;42:1901-7.

11. Cote P, Cassidy JD, Yong-Hing K, Sibley J, Loewy J. Apophysial joint degeneration, disc degeneration, and sagittal curve of the cervical spine: can they be measured reliably on radiographs? Spine (Phila Pa 1976) 1997;22:859-64.

12. 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.

13. 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.

14. Ulmar B, Guhring M, Schmalzle T, Weise K, Badke A, Brunner A. Inter- and intra-observer reliability of the Cobb angle in the measurement of vertebral, local and segmental kyphosis of traumatic lumbar spine fractures in the lateral X-ray. Arch Orthop Trauma Surg 2010;130:1533-8.

15. Lee SW, Hong JT, Son BC, Sung JH, Kim IS, Park CK. Analysis of accuracy of kyphotic angle measurement for vertebral osteoporotic compression fractures. J Clin Neurosci 2007;14:961-5.

16. Street J, Lenehan B, Albietz J, et al. Intraobserver and interobserver reliabilty of measures of kyphosis in thoracolumbar fractures. Spine J 2009;9:464-9.

17. Janusz P, Tyrakowski M, Glowka P, Offoha R, Siemi-onow K. Influence of cervical spine position on the radiographic parameters of the thoracic inlet alignment. Eur Spine J 2015;24:2880-4.

18. Skeppholm M, Lindgren L, Henriques T, Vavruch L, Lofgren H, Olerud C. The Discover artificial disc replacement versus fusion in cervical radiculopathy: a randomized controlled outcome trial with 2-year follow-up. Spine J 2015;15:1284-94.

19. MacDowall A, Skeppholm M, Lindhagen L, et al. Artificial disc replacement versus fusion in patients with cervical degenerative disc disease with radiculopathy: 5-year outcomes from the National Swedish Spine Register. J Neurosurg Spine 2018;30:159-67.

20. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what?: a practical guide for medical statisticians. Stat Med 2000;19:1141-64.

21. R Core Team. R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing; 2014 [cited 2018 Jan 19]. Available from: http://www.R-project.org/.