Gait & Posture 84 (2021) 357–367

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Radiography and photogrammetry-based methods of assessing cervical spine posture in the sagittal plane: A systematic review with meta-analysis

Luiza Rampi Pivotto [a,*], Isis Juliene Rodrigues Leite Navarro [a], Cláudia Tarragô Candotti [b]

[a] *Researcher of Grupo de Investigação da Mecânica do Movimento (BIOMEC) in Laboratório de pesquisa do exercício (LAPEX) at Universidade Federal do Rio Grande do Sul - UFRGS, Brazil*

[b] *Professor and researcher of Grupo de Investigação da Mecânica do Movimento (BIOMEC) in Laboratório de pesquisa do exercício (LAPEX) at Universidade Federal do Rio Grande do Sul - UFRGS, Brazil*

ARTICLE INFO

Keywords:
Neck
Posture
Evaluation
Radiography
Photogrammetry
Reproducibility of results

ABSTRACT

Background: There are numerous radiography and photogrammetry-based methods of assessing the cervical spine posture in the sagittal plane. The choice of instrument should be based on scientific parameters such as validity and reliability, thus avoiding restrictions to the applicability of the instrument.

Research question: What radiography and photogrammetry-based methods used to assess the cervical spine posture in the sagittal plane are valid and/or reliable?

Methods: Systematic searches were conducted following Meta-analysis of Observational Studies in Epidemiology guidelines. Methodological quality was assessed according to the Brink & Louw appraisal tool.

Results: Twenty-one studies were included in the qualitative analysis. Twenty different methods of calculating cervical spine posture in the sagittal plane were found. Two studies included validation measures, 16 studies assessed inter-rater reliability, and 17 studies assessed intra-rater reliability. Fourteen studies were included for the quantitative analysis. The meta-analysis shows that the cervical arrow and cervical lordosis photogrammetry-based methods present very high intra-rater reliability. In radiography, the meta-analysis also showed that the Cobb method (inferior C2 - inferior C7), Cobb method (middle C1 - inferior C7), absolute rotation angle, and Gore angle (C2-C7) present very high inter-rater reliability, and the Cobb method (inferior C2 - inferior C7) and absolute rotation angle present very high intra-rater reliability.

Significance: This systematic review presents an overview of the methods used to assess cervical spine posture and the respective information on validity and reliability. This panorama facilitates the choice of method when conducting radiography or photogrammetry-based assessment of the cervical spine in the sagittal plane. In addition, it shows the need for new studies that investigate the accuracy and precision of these methods for their possible use in larger studies.

1. Introduction

There are several ways to assess and identify changes in cervical curvature and they can be divided into two groups: invasive and noninvasive assessment techniques. Radiography is a type of invasive technique because it exposes the subject to radiation during the exam, despite which it is considered the gold standard exam to assess the spinal curvatures. However, there are a variety of radiograph-based methods that quantify the curvature in the sagittal plane of the cervical spine, such as the Cobb method [1], absolute rotation angle [2], and Ishihara index method [3,4], among others [3,5]. The Cobb method is considered the gold standard for quantifying the magnitude of spinal curvatures, despite having received criticism [2,6].

Noninvasive quantitative assessment techniques are intended to indirectly inform spinal posture. Photogrammetry combined with computerized techniques is recommended as an important tool for postural assessment [7–9]. Studies that use photogrammetry as an assessment tool tend to have/involve very similar data collection procedures, differing slightly according to the purpose of the study [10]. However, in relation to data analysis, the methods are very different.

There are numerous radiography and photogrammetry-based methods of assessing the cervical spine posture in the sagittal plane.

* Corresponding author at: Universidade Federal do Rio Grande do Sul - UFRGS, Escola de Educação Física, Fisioterapia e Dança (ESEFID), LAPEX/BIOMEC, Rua Felizardo, 750, Porto Alegre, RS, CEP 90690-200, Brazil.

E-mail addresses: luizarpivotto@gmail.com (L.R. Pivotto), isis.juliene@ufrgs.br (I.J.R.L. Navarro), claudia.candotti@ufrgs.br (C.T. Candotti).

https://doi.org/10.1016/j.gaitpost.2020.12.033
Received 2 April 2020; Received in revised form 29 December 2020; Accepted 31 December 2020
Available online 7 January 2021

The choice of instrument should be based on scientific parameters such as validity and reliability, thus avoiding restrictions to the applicability of the instrument [11,12]. Accordingly, the aim of the present systematic review is to identify which radiography and photogrammetry-based methods used to assess the cervical spine posture in the sagittal plane are valid and/or reliable.

2. Methods

2.1. Study design

The present study is a systematic review with meta-analysis, registered on PROSPERO under the code CRD42019123708 and following MOOSE (Meta-analysis of Observational Studies in Epidemiology) guidelines [13].

2.2. Search strategies

Systematic searches were conducted in March 2019 on PubMed, EMBASE, Scopus, Web of Science, Science Direct, Bireme, and SciELO. In August 2020 the systematic searches were updated.

The Medical Subject Headings (MeSH) terms and boolean operators used on the systematic search were: Neck[MeSH] OR “Cervical Vertebrae”[MeSH] AND Photogrammetry[MeSH] OR Radiograph*[MeSH] AND Posture[MeSH] AND “Validation Studies”[Publication Type] OR “Reproducibility of Results”[MeSH]. Entry terms for each MeSH term were also used in the search in the text words (tw) field and separated by the boolean operator OR. The search strategy used on PubMed is shown in Fig. 1. In addition, there was no restriction for language or publication date. Also, manual searches were performed on the references of the included studies.

2.3. Eligibility criteria

The eligible studies met the following criteria: (1) cervical posture evaluation; (2) standing position; (3) sagittal plane; (4) healthy adults and children of both sexes; (5) the use of photogrammetry or radiography; and (6) validity and/or intra- and/or inter-rater reliability study.

2.4. Study selection and data extraction

The bibliographic details of all the retrieved studies were stored in an EndNote file (version x7). Two independent reviewers selected potentially relevant studies, according to their titles and abstracts. The reviewers were master and doctoral students, both had experience with systematic reviews. When the study title and abstract did not provide sufficient information to confirm eligibility/exclusion, the text was read in full. When a study was not found in its entirety, the study authors were contacted or the Brazilian Bibliographic Commuting Program was engaged to help find the study.

In the next step, the reviewers read the full studies and selected them according to the eligibility criteria. Discordant cases were solved by consensus or by a third reviewer.

Only the included studies were submitted to data extraction and methodological quality assessment. Information was extracted to a standardized form that included: author, year of publication, sample size, sample age, type of assessment (radiography and/or photogrammetry), type of analysis (validity and/or reliability), method description (anatomical landmarks and calculation), description of reliability parameters (number of raters, interval between measurements, qualification of raters, correlation coefficient, and error measure), and description of validity (standard reference, interval between measurements, and results).

2.5. Quality assessment

The methodological quality was assessed using a critical appraisal tool proposed by Brink and Louw [14]. The scale consists of 13 items, of which five items relate to both validity and reliability studies, four items to validity studies only, and four items to reliability studies. This scale was used by the same independent reviewers. The included studies were considered of high methodological quality if they reached a score of ≥ 60 %, as proposed in a previous study [15].

2.6. Statistical analysis

Data were initially separated into subgroups according to type of assessment (photogrammetry or radiography), method, and type of analysis (validity, intra-rater or inter-rater reliability and statistical test conducted).

The data were meta-analyzed using RStudio software (version 1.3.1073). The studies were grouped according to the type of assessment (radiography or photogrammetry), type of analysis (intra-rater or inter-rater reliability), correlation coefficient (Intraclass Correlation Coefficient or Pearson’s correlation coefficient) and method used. The random-effect model was selected for the analysis.

The statistical data originated from Pearson’s correlation coefficient (r) or Intraclass Correlation Coefficient (ICC) were interpreted as follows: ≤ 0.25 very low correlation, 0.26 – 0.49 low correlation, 0.50 – 0.69 moderate correlation, 0.70 – 0.89 high correlation and ≥ 0.90 very high correlation.

Heterogeneity was checked using the Higgins Inconsistence test (I^2). Values over 50 % were considered of high heterogeneity [16].

3. Results

3.1. Qualitative analysis

Initially, 1382 studies were identified in the systematic searches, 581 studies were duplicates, 760 were excluded after the reading the titles and abstracts, with 41 remaining for the full-text reading. Based on the eligibility criteria 28 studies were excluded, and eight were included from the references lists, leaving a total of 21 studies for the qualitative analysis and 14 for the quantitative analysis. Fig. 2 shows the study selection flowchart. Table 1 describes the different methods used in the studies for cervical spine assessment. Tables 2 and 3 summarize the characteristics of the studies included in the qualitative analysis that assessed the reliability of the measurements obtained using the radiography and photogrammetry-based methods, respectively. Table 4 summarizes the characteristics of the study included in the qualitative analysis that evaluated the validity of the photogrammetry-based methods.

Only two studies included validation measures [17,18], 16 studies assessed inter-rater reliability [2,3,17,19–31], and 17 studies assessed

	Search (#1 AND #2 AND #3 AND #4)
#1	Neck[MeSH] OR Neck[tw] OR Necks[tw] OR “Cervical Vertebrae”[MeSH] OR “Cervical Vertebrae”[tw] OR “Vertebrae, Cervical”[tw] OR “Cervical Curvatures” OR “Cervical Curve” OR Cervical
#2	Photogrammetry[MeSH] OR Photogrammetry[tw] OR Photogrammetries[tw] OR Radiograph*[MeSH] OR Radiograph*[tw] OR “Diagnostic X-Ray”[tw] OR “Diagnostic X Ray”[tw] OR “Diagnostic X-Rays”[tw] OR “X-Rays, Diagnostic”[tw] OR “X-Ray Radiology, Diagnostic”[tw] OR “X Ray Radiology, Diagnostic”[tw] OR “Radiology, Diagnostic X-Ray”[tw] OR “Radiology, Diagnostic X Ray”[tw] OR “X-Ray, Diagnostic”[tw] OR “X Ray, Diagnostic”[tw] OR “Diagnostic X-Ray Radiology”[tw] OR “Diagnostic X Ray Radiology”[tw] OR “X-ray image”[tw] OR “X-ray diagnosis”[tw] OR X-ray[tw]
#3	Posture[MeSH] OR Postures[tw] OR Sagittal
#4	"Validation Studies"[Publication Type] OR “Reproducibility of Results"[MeSH] OR “Reproducibility of Results"[tw] "Reproducibility of Findings"[tw] OR "Reliability of Results"[tw] OR "Validity of Results"[tw] OR "Reliability and Validity"[tw] OR "Validity and Reliability"[tw] OR Reliability OR Reliabilities OR Validity OR Validities

Legend: MeSH = Medical Subject Headings; tw = Text words.

Fig. 1. Search strategy of PubMed.

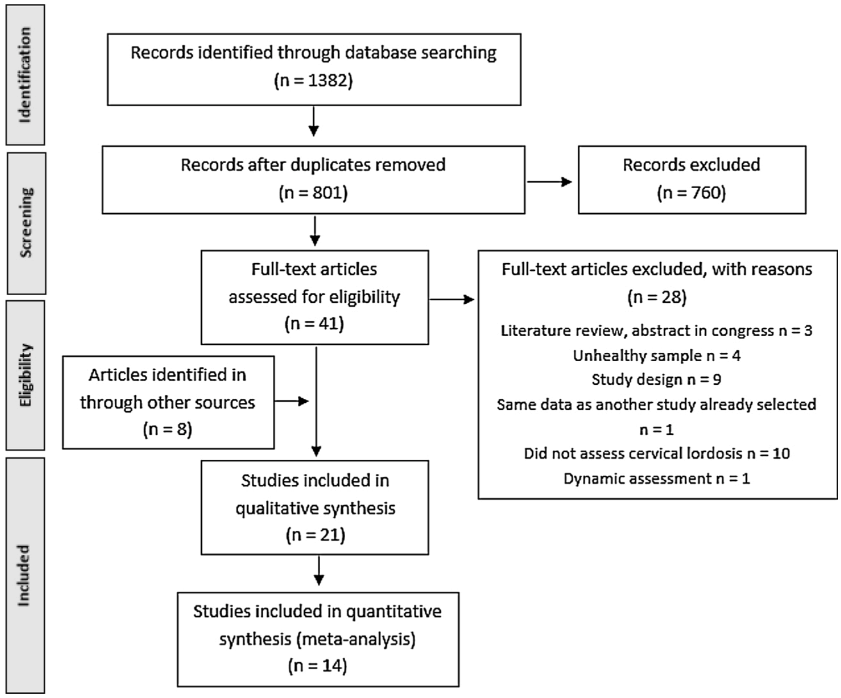


Fig. 2. Flowchart of the included studies.

intra-rater reliability [2,3,17–19,21–25,27,28,30,32–35].

Regarding the demographic characteristics, six studies selected a part of the sample to be used in the reliability calculations, rather than using the whole sample [18,27,28,32–34]. The number of subjects ranged from 5 to 218, and the median sample size was 30 (Q1 = 20; Q3 = 101). Seven studies did not report the mean age of the sample [2,18,22,26,28,32,33]. In the studies that reported this information, the mean age of the subjects was 31.3 (SD 15.7) years, ranging from 11 to 65.8.

In the 21 studies included in this review, 20 different methods of assessing cervical spine posture in the sagittal plane appeared; 16 involved calculating an angle to describe cervical spine posture and four were based on distances.

Regarding the methodological quality, 15 studies ranked high quality (score $\geq$ 60 %) [2,3,17–20,22,23,25–27,30,33–35]. The average score of the methodological quality appraisal was 71.8 %. The main areas of methodological weakness found were: randomization of evaluators or subjects (item 6), intra-rater blindness (item 5), and the time interval between repeated measures (item 8) (Table 5).

3.1.1. *Photogrammetry*

Six studies [17,18,23,27,34,35] investigated the reliability of photogrammetry-based methods of assessing cervical spine posture in the sagittal plane (Table 3), two of which also assessed the validity of the measurement [17,18]. In the six studies included, four methods of assessing cervical spine posture in the sagittal plane were found: cervical distance, cervical lordosis, cervical arrow and cervical angle. Table 1 shows the descriptions of the methods. The cervical arrow [17,35] and cervical lordosis [23,34] methods were reported in two studies each.

3.1.2. *Radiography*

There is a wide range of radiography-based methods of assessing cervical spine posture in the sagittal plane: Cobb method, cervical curvature ratio (CVT/EVT), lordotic curvature, cervical angle (CAr), spinous processes cervical angle (spCA), total cervical lordosis, centroid measurement, Ishihara index method, Gore angle (also called Jackson physiological stress lines or posterior tangent method), and absolute rotation angle (also called sum of posterior tangent method or Harrison's method). Table 1 shows the descriptions of the methods. Only one of the methods did not have a description or reference cited in the study [28]. Regarding the Cobb and the Gore angle methods, there were variations in the vertebrae used as reference.

Among the methods analyzed in the included studies, regarding the lower anatomical reference used to calculate cervical lordosis, C7 was the most used to calculate the angle of cervical curvature and C6 was used in two methods. Regarding the upper anatomical reference, C2 was the most used, appearing in nine methods, C1, odontoid process, and C3 were used in two methods, and the occipital was used in one method.

The Cobb method (inferior C2 - inferior C7) appeared most frequently, in seven studies [2,19,20,24–26,30]. The absolute rotation angle was used in five studies [2,3,27,29,30]; Gore angle (C2-C7) [20,30,31] and Cobb method (middle C1 - inferior C7) [2,21,26] were used in three studies, and the other methods were used in only one study each.

3.1.3. *Intra-rater reliability*

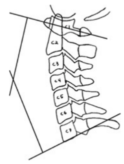
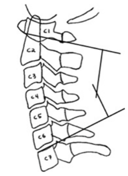
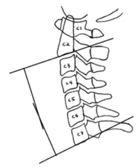
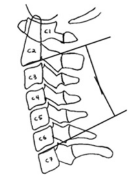
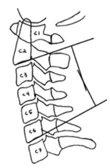
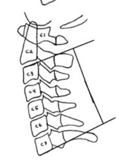
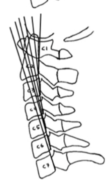
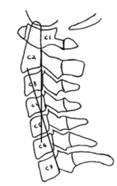
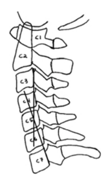
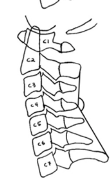
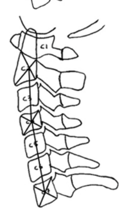
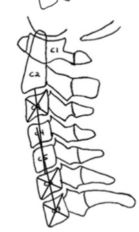
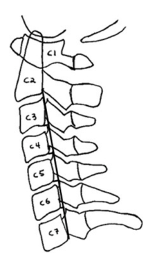
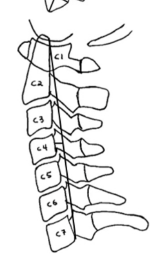
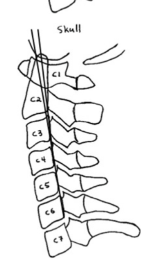
With respect to the 17 studies that assessed intra-rater reliability, 12 studies reported the interval between measurements [2,3,17–19,22,23,27,30,33–35]. The intervals ranged from 15 min to eight weeks. Of these 12 studies, five used a one-week interval [2,17,18,27,35], and two used a four-week interval [19,30]. Belli et al. [34] used an interval of 15 min, Silber et al. [22] used an interval of at least one week, Bernall et al. [33] used an interval of two weeks, Iunes et al. [23] used an interval of at least four weeks, and Ohara et al. [3] used an interval of eight weeks.

Thirteen studies used ICC as a measure of intra-rater reliability [2,17–19,23–25,27,30,32–35], four used Pearson's r [3,21,28,35], and one used another statistical analysis technique [22]. Furlanetto et al. [35] used both the ICC and Pearson's r.

In addition, the Standard Error of Measurement (SEM) was reported in four studies [17,25,27,30]. The Standard Error of Estimate (SEE) [21] and the Standard Error multiplied by two (2x SE) [23] were reported in one study each.

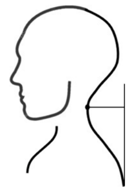
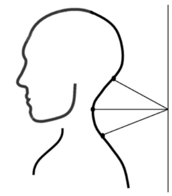
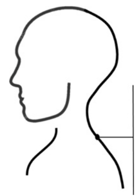
It is important to note that these error measures must be interpreted according to the magnitude of each measure, that is, individually within each method.

Table 1
Description and scheme of methods used in studies included.

Method	Method description	Method scheme	Method	Method description	Method scheme
Cobb method (middle C1 - inferior C7) [2,21,26]	This angle was formed by a line bisecting C1 and a line in the inferior endplate of C7.		Cobb method (inferior C1 - superior C7) [3]	This angle was formed by a line in the inferior endplate of C1 and the superior endplate of C7.	
Cobb method (inferior C2 - inferior C7) [2,19,20,24,25,26,30]	This angle was formed by a line in the inferior endplate of C2 and a line in the inferior endplate of C7.		Cobb method (inferior C2 -superior C7) [3]	This angle was formed by a line in the inferior endplate of C2 and a line in the superior endplate of C7.	
Cobb method (C2-C6) [32]	This angle was formed by a line in the inferior endplate of C2 and a line in the inferior endplate of C6.		Cobb method (C3-C7) [22]	This angle was formed by a line in the superior endplate of C3 and a line in the inferior endplate of C7.	
Absolute rotation angle or sum of posterior tangent method or Harrison's method (C2-C7) [2,3,27,29,30]	This angle was formed by drawing lines that are parallel to the posterior body margin from C2 to C7 and then summing the segmental angles.		Gore angle (C2-C7) or Jackson physiological stress lines or posterior tangent method [20,30,31]	This angle was formed by a line in the posterior body margin of C2 and a line in the posterior body margin of C7.	
Gore angle (C3-C7) [22]	This angle was formed by a line in the posterior body margin of C3 and a line in the posterior body margin of C7.		Spinous processes cervical angle (spCA) (C2-C4-C7) [27]	This angle was formed using the spinous processes of C2, C4 and C7.	
Cervical angle (CAr) (C2-C4-C7) [27]	This angle was formed using the centers of the vertebral body of C2, C4 and C7.		Centroid measurement (C2-C3-C6-C7) [3]	Points a, b, and c are the centroids of C3, C6, and C7, respectively. Point A is the midpoint of the inferior surface of C2. This angle was formed by a line between Aa and a line between bc.	
Ishihara index method [3]	The posterior inferior points of C2 and C7 are points C and D. The distance between the posterior inferior points of C3–C6 and the CD was called a3 to a6, respectively. The measurement is computed by this formula: (a3+a4+a5+a6)/CDx100.		Lordotic curvature (Odontoid-deepest vertebra-C7) [33]	The measurement is the line from the midpoint of the deepest vertebra to a line from the most superior posterior point of the odontoid to the most posterior inferior point of C7.	
Total cervical lordosis (Occipital-C7) [28]	No description	–	Cervical curvature ratio (CVT/EVT) (Odontoid-C4-C6) [18]	This angle was formed by the line that intersects the apex of the odontoid with the most posterior-inferior point of C4 (CVT) and the line that intersects the most posterior-inferior points of C4 and C6 (EVT).	
Cervical distance (Deepest region) [18]	The horizontal distance from a vertical line tangent, by the apex of the thoracic kyphosis, called as thoracic plan, and		Cervical lordosis (Occipital-C4-C7) [23,34]	This angle was formed by the straight lines between the occipital protuberance and C7 and that intersects the horizontal line	

(continued on next page)

Table 1 (*continued*)

Method	Method description	Method scheme	Method	Method description	Method scheme
	by the point of the apex of the cervical concavity.			between C4 and the true vertical line.	
Cervical arrow (C7) [17,35]	The horizontal distances between C7 and a vertical reference line originating from the S2.		Cervical angle (C2-C4-C7) [27]	This angle was formed by a line connecting C2 and C4 with a line connecting C4 and C7.	

3.1.4. *Inter-rater reliability*

With respect to the 16 studies that assessed inter-rater reliability, all the studies reported the number of raters who performed the measurements. Nine studies used two raters to check inter-rater reliability [3,19,20,24,25,27–29,31] and seven studies used three raters [2,17,21–23,26,30]. Regarding the raters' qualifications, 13 studies reported this information [2,3,19,20,22–30]. In only two studies the raters were physical therapists [23,27], in the other 11 studies the raters were physicians. Of these, in eight studies at least one of the raters was an orthopedic surgeon [3,19,20,22,25,26,28,30].

Twelve studies used the ICC [2,17,19,20,23–27,29–31], three used the Pearson's r as a measure of inter-rater reliability [3,21,28], and one used another statistical analysis technique [22].

In addition, the SEM was reported in five studies [17,25–27,30], the SEE [21] and 2x SE [23] was reported in one study each.

3.1.5. *Validity*

Only two validity studies were included in this systematic review [17,18]. These studies used radiography as a standard reference to validate a photogrammetry-based method (Table 4). These methods, namely cervical arrow and cervical distance, result in a distance measure.

3.2. *Quantitative analysis*

Fourteen studies were included in the meta-analysis [2,17,19,20,23–27,29–31,34,35]. It was possible to perform eight meta-analyses with the ICC values presented, four of inter-rater reliability (Cobb method (inferior C2 - inferior C7), Cobb method (middle C1 - inferior C7), absolute rotation angle, and Gore angle (C2-C7)), and four of intra-rater reliability (Cobb method (inferior C2 - inferior C7), absolute rotation angle, cervical lordosis, and cervical arrow).

In the inter- and intra-rater analyses of the Cobb method (inferior C2 - inferior C7) (Tables 6 and 7, respectively), inter-rater analysis of absolute rotation angle (Table 9), and inter-rater analysis of Gore angle (C2-C7) (Table 11), the ICCs ranged from high to very high, and the I^2 showed high heterogeneity (> 50 %).

In the inter-rater analysis of Cobb method (middle C1 - inferior C7) (Table 8), the ICCs ranged from high to very high, and the I^2 showed low heterogeneity (0%).

In the intra-rater analysis of absolute rotation angle (Table 10), and cervical arrow (Table 12), the ICCs ranges were very high for both the lower and upper limits, and the I^2 showed high heterogeneity (67 %) and low heterogeneity (0%), respectively.

In the intra-rater analysis of cervical lordosis (Table 13), the ICC ranged from moderate to very high, and the I^2 showed heterogeneity of 50 %.

4. Discussion

Increasingly, researchers and professionals are encouraged to choose instruments to measure a variable of interest based on scientific parameters that guarantee the accuracy and precision of the measurements, such as validity and reliability.

It is important to note that there is a lack of standardization and definition of terms used in validity and reliability studies. In the present review, validity refers to "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" [36], that is, the degree of accuracy of measurements of a given magnitude. The concept of validity is associated with the concept of accuracy. Reliability refers to the consistency of a measure. According to Bartko [37], reliability is "the degree to which multiple assessments of a subject agree (reproducibility)". The concept of reliability is associated with the concept of precision.

Considering that radiography is the gold standard for assessing cervical spine posture in the sagittal plane, it was expected to find validation studies of photogrammetry-based methods that used radiography as reference. However, only two studies did this [17,18]. Albuquerque et al. [38] also found few validity and reliability studies that used photogrammetry to measure the angle of cervical lordosis. It is speculated that the scarcity of studies on the subject is due to the difficulty involved in assessing the region using photography, which uses the skin surface of the cervical region as a reference.

Refshauge, Goodsell and Lee [39] investigated the degree to which surface measurements of cervical alignment reflect the underlying vertebral body alignment. The results showed Pearson's correlation coefficients of 0.32 to 0.82 (p < 0001), poor to moderate coefficients according to the reference used by the authors [40] and low to high coefficients according to the reference used in the present review. The authors argue that the findings can be explained by a combination of factors: the difference in length of the spinous processes in the cervical region, the depth of overlying soft tissues and the individual variability of the bone dimensions and overlying soft tissue. Therefore, methods that assess cervical spine in the sagittal plane based on the skin surface need to present validity measures to be accurate.

Another reason that may explain the small amount of studies on the subject is the confusion between the concepts of cervical spine posture and head posture. A widely used parameter is the angle formed between a line connecting the C7 spinous process and the tragus of the ear and a horizontal line drawn from C7 [41,42]. However, this parameter is mistakenly called the cervical angle or neck angle in some studies [43,44], since it was proposed to assess head posture. In this review, studies that evaluated head posture were excluded. Still, the opposite may also have happened, and studies that evaluate the cervical posture but call it the head posture parameter may not have been found in the systematic

Table 2
Characteristics of the studies, reliability results and standard error of measurement of radiographic measurements.

Author	Sample (mean ± SD age in years)	Method (unit of measure)	Inter-rater reliability	Intra-rater reliability
Abelin-Genevois et al. [25]	150 asymptomatic children (13.8 ± 1.7)	Cobb method (inferior C2 - inferior C7) (°)	ICC = 0.95; SEM < 2°	ICC = 0.98; SEM < 1°
Armijo-Olivo et al. [32]	10 of 68 radiographs (14.3 ± 9.1)	Cobb method (C2-C6) (°)	–	ICC = 0.69
Bernal et al. [33]	15 of 107 children (8.56 ± 1.5)	Lordotic curvature (mm)	–	ICC = 0.945 – 0.996
Côté et al. [26]	30 radiographs	Cobb method (middle C1 - inferior C7) (°)	ICC = 0.94; SEM = 9.1°	–
		Cobb method (inferior C2 - inferior C7) (°)	ICC = 0.96; SEM = 8.3°	–
Gadotti et al. [27]	22 (28 ± 4.37) of 39 women	Absolute rotation angle (°)	ICC = 0.90; SEM = 3.78°	ICC = 0.99; SEM = 0.99° ICC = 0.94; SEM = 0.5°
		Cervical angle (°)	ICC = 0.99; SEM = 0.01°	ICC = 0.99; SEM = 0.15° ICC = 0.97; SEM = 0.81°
		Spinous processes cervical angle (°)	ICC = 0.79; SEM = 6.89°	ICC = 0.96; SEM = 0.33° ICC = 0.94; SEM = 4.03°
Hardacker et al. [28]	30 of 50 radiographs (38.4 ± 9.4 and 38.6 ± 9.2)	Total cervical lordosis (°)	r = 0.85	r = 0.95
Harrison et al. [2]	30 radiographs	Cobb method (middle C1 - inferior C7) (°)	ICC = 0.91	ICC = 0.94
		Cobb method (inferior C2 - inferior C7) (°)	ICC = 0.92	ICC = 0.95
		Absolute rotation angle (°)	ICC = 0.94	ICC = 0.97
Iyer et al. [29]	115 asymptomatic adults (50.1)	Harrison's method (C2-C7) (°)	ICC = 0.98	–
Janusz et al. [30]	44 radiographs (15.8 ± 3.7)	Cobb method (inferior C2 - inferior C7) (°)	ICC = 0.92; SEM = 2.71°	ICC = 0.96; SEM = 2.06°
		Posterior tangent method / Gore angle / Jackson physiological stress lines (C2-C7) (°)	ICC = 0.94; SEM = 2.62°	ICC = 0.96; SEM = 1.99°
		Sum of posterior tangent method / Harrison's method (°)	ICC = 0.93 SEM = 2.78°	ICC = 0.96; SEM = 1.98°
Lee et al. [31]	181 asymptomatic children (11.7 ± 4.4)	Posterior tangent method / Gore angle (C2-C7) (°)	ICC = 0.862–0.922	–
Ohara et al. [3]	120 radiographs equally divided into 3 groups according to cervical alignment (lordosis, straight or sigmoid and kyphosis) (45.0 ± 18.0)	Cobb method (inferior C1 - superior C7) (°)	r = 0.954 r = 0.980 r = 0.985	r = 0.918 r = 0.989 r = 0.989
		Cobb method (inferior C2 - superior C7) (°)	r = 0.945 r = 0.988 r = 0.979	r = 0.898 r = 0.984 r = 0.984
		Centroid measurement (°)	r = 0.976 r = 0.974 r = 0.919	r = 0.972 r = 0.965 r = 0.965
		Absolute rotation angle (°)	r = 0.966 r = 0.996 r = 0.975	r = 0.929 r = 0.989 r = 0.989
		Ishihara index method	r = 0.996 r = 0.831 r = 0.978	r = 0.996 r = 0.991 r = 0.991
Park et al. [20]	101 volunteer adults (29.1)	Gore angle (C2-C7) (°)	ICC = 0.99 ICC = 0.97	–
		Cobb method (inferior C2 - inferior C7) (°)	ICC = 0.98 ICC = 0.97	–
Park et al. [19]	100 asymptomatic adults (23.4 and 65.8)	Cobb method (inferior C2 - inferior C7) (°)	ICC = 0.672	ICC = 0.777

Plaugher; Cremata; Phillips. [21]	98 radiographs from 49 patients (44.0 ± 20.0)	Cobb method (middle C1 - inferior C7) (°)	r = 0.89;
			SEE = 4.98°
			r = 0.96;
			SEE = 3.45°
			r = 0.94;
			SEE = 4.04°
				r = 0.97;
				SEE = 2.88°

Silber et al. [22]	20 radiographs of nonspondylotic spines	Gore angle (C3-C7) (°)	95 % CI = 7	95 % CI = 4
		Cobb method (C3-C7) (°)	95 % CI = 9	95 % CI = 5
Weber et al. [18]	20 of 80 women (28.3 ± 3.65)	Cervical curvature ratio (°)	–	ICC = 0.979
Zhou et al. [24]	218 healthy volunteers (48.4 ± 16.9)	Cobb method (inferior C2 - inferior C7) (°)	ICC = 0.96	ICC = 0.86

Legend: ICC = Intraclass Correlation Coefficient; r = Pearson's Coefficient; SEM = Standard Error of Measurement; SEE = Standard Error of Estimate; CI = Confidence Interval.

Table 3
Characteristics of the study, reliability results and standard error of measurement of photogrammetric measurements.

Author	Sample (mean ± SD age in years)	Method (Analysis software)	Inter-rater reliability	Intra-rater reliability
Belli et al. [34]	5 (11 ± 1.41) of 30 non asthmatic children	Cervical lordosis (ALCimage software)	–	ICC > 0.75
Furlanetto et al. [35]	15 university students (24.7 ± 4.0)	Cervical arrow (DIPA software)	–	ICC = 0.958 r = 0.920
Furlanetto et al. [17]	16 individuals (23.7 ± 3.6)	Cervical arrow (DIPA software)	ICC = 0.936; SEM = 0.6 cm; MDC = 1.2 cm	ICC = 0.948; SEM = 0.6 cm; MDC = 1.1 cm
Gadotti et al. [27]	22 (28 ± 4.37) of 39 women	Cervical angle (ALCimage software)	ICC = 0.91; SEM = 7.06°	ICC = 0.98; SEM = 0.37° ICC = 0.98; SEM = 0.22°
Iunes et al. [23]	21 university students (24.19 ± 1.3)	Cervical lordosis (ALCimage software)	ICC = 0.748; 2x SE = 2.94	ICC = 0.966; 2x SE = 2.65
Weber et al. [18]	20 of 80 women (28.3 ± 3.65)	Cervical distance (SAPO®)	–	ICC = 0.974

Legend: ICC = Intraclass Correlation Coefficient; r = Pearson's Correlation Coefficient; SEM = Standard Error of Measurement; MDC = Minimal Detectable Change; SE = Standard Error.

Table 4
Characteristics of the studies and validity results of photogrammetric. measurements.

Author	Sample (mean ± SD age in years)	Method (Analysis software)	Standard reference	Validity
Furlanetto et al. [17]	54 individuals (45.4 ± 18.1)	Cervical arrow (DIPA software)	Cervical arrow (Radiograph)	r = 083; p < 0.001 t = -1485; p = 0.145 P = 008; p = 0.999
Weber et al. [18]	80 women (28.3 ± 3.65)	Cervical distance (SAPO®)	Cervical curvature ratio (Radiograph)	r = 007; p = 0.52

Legend: r = Pearson's Correlation Coefficient; t = independent *t*-test; P = Morgan-Pitman's Coefficient.
t-test: $p < 0.05$ indicates difference between the means.
Morgan-Pitman's Coefficient: $p < 0.05$ indicates heterogeneity of variances.

search or excluded after reading the titles and abstracts.

Both validity studies included in the present review used Pearson's r as a statistical analysis tool [17,18]. However, Pearson's r alone is insufficient to define the degree of agreement between two measures [38,45]. Furlanetto et al. [17] also included the independent *t* test and the Morgan–Pitman's test in the statistical analysis. But to prove that a new technique agrees well enough with an already established one, Bland and Altman [45] suggest a statistical analysis involving graphical techniques and simple calculations. One critical conclusion made by these authors is that reliability studies require a measure of reliability in the scale of the original measurements, such as the limits of agreement. This alternative analysis can be used to assess the agreement between two measures that were obtained either through different instruments or different raters [45].

All the studies included in this review evaluated some measure of reliability, intra-rater or inter-rater. It was possible to meta-analyze the intra-rater reliability values of the photogrammetry-based cervical arrow and cervical lordosis methods. Only two studies were included in each analysis, all of which presented methodological quality ≥ 60 %. Still, both methods showed a very high correlation and low heterogeneity (0% and 50 %, respectively). The cervical arrow method presented both validity and reliability measures. It can be assessed by the same rater on different days or by different raters. However, there is a lack of information on how the measurements obtained from the photographs agree with the information from the respective X-rays.

In relation to radiography, the present review found there to be a large number of methods of assessing the curvature of the cervical spine in the sagittal plane. Furthermore, some methods are used in different ways in different studies. The Cobb method, for example, was used in six different ways, varying the anatomical references used. Despite this, it was possible to meta-analyze the inter-rater and intra-rater reliability values of the Cobb method (inferior C2 - inferior C7). Both results showed a very high correlation and no study presented low methodological quality. However, the results showed a high heterogeneity (95 % and 97 %, respectively). It was also possible to meta-analyze the inter-rater reliability values of the Cobb method (middle C1 - inferior C7). Only two studies were included in this analysis, both with methodological quality ≥ 60 %, and the results showed a very high correlation and low heterogeneity (0%).

Although the Cobb method is considered the gold standard for quantifying the magnitude of spinal curvature, several other methods serve the same purpose. It was possible to meta-analyze the intra-rater and inter-rater reliability values of the absolute rotation angle and the inter-rater reliability values of the Gore angle (C2-C7). The results showed a very high correlation and high heterogeneity (67 %, 86 % and 98 %, respectively).

All the studies included in the meta-analysis of intra-rater reliability values of the absolute rotation angle showed methodological quality ≥ 60 %. Of the four studies included in the meta-analysis of the inter-rater reliability values of the absolute rotation angle and of the three included in the meta-analysis of the intra-rater reliability of the Gore angle (C2-C7), one study showed methodological quality ≤ 60 % in each analysis.

Regarding these two methods, the absolute rotation angle and Gore angle (C2-C7), there is confusion in the nomenclature used by the studies. From the description of the methods and references cited in the studies, two alternative labels for the absolute rotation angle, namely the sum of posterior tangent method and Harrison's method (C2-C7) and two alternative labels for the Gore angle (C2-C7), namely the Jackson physiological stress lines and posterior tangent method, were found (Table 1). The methods are very similar, the difference is that the first takes into account parallel lines on the posterior margins of the bodies from the second to the seventh cervical vertebra and then adds the segmented angles formed. While the second method only takes into account the angle formed between the parallel lines drawn on the posterior margins of the bodies of the second and seventh cervical vertebra. According to Harrison et al. [2], the absolute rotation angle more accurately denotes the state of curvature from the interpretation of the segmented angles calculated to obtain the final angle.

Some studies used the Pearson product moment correlation coefficient, which is an interclass correlation coefficient, as a measure of reliability for continuous data. However, the use of Pearson's r has been discouraged for this type of analysis, and ICC, which is an intraclass correlation coefficient [12,37,46–48], is the most appropriate. According to Weir et al. [47], Pearson's r does not detect systematic errors. According to Baumgartner [46], there are four reasons why the ICC is the correct coefficient to estimate reliability: (1) ICC permits more than two scores per person, whereas r is limited to two scores per person; (2) ICC is sensitive to more sources of error than r; (3) ICC is affected by change in the mean and standard deviation from one set of scores to the next but r is not affected by changes in these statistics; and (4) ICC is the proper coefficient because it is designed for repeated measures of a test as in reliability, whereas r is designed to determine the relation between two sets of scores.

Table 5
Quality assessment through the critical evaluation tool proposed by Brink & Louw.

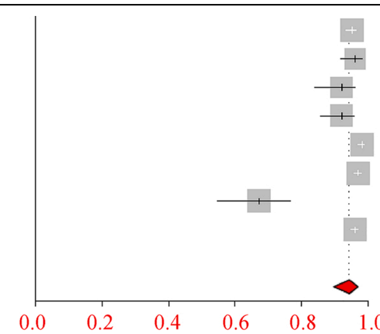
Study	1	2	3	4	5	6	7	8	9	10	11	12	13	%
Abelin-Genevois et al. [25]	y	y	n/a	n	n	n/a	n/a	n	n/a	y	n/a	y	y	62.5
Armijo-Olivo et al. [32]	y	n	n/a	n/a	n	n/a	n/a	n	n/a	y	n/a	y	y	57.1
Belli et al. [34]	y	n	n/a	n/a	y	n	n/a	n	n/a	y	n/a	y	y	62.5
Bernall et al. [33]	y	n	n/a	n/a	n	n/a	n/a	y	n/a	y	n/a	y	y	71.4
Côté et al. [26]	n	y	n/a	y	n/a	n/a	n/a	n	n/a	y	n/a	y	y	71.4
Furlanetto et al. [35]	y	n	n/a	n/a	n	n	n/a	y	n/a	y	n/a	y	y	62.5
Furlanetto et al. [17]	y	n	y	y	y	y	y	y	y	y	y	y	y	92.3
Gadotti et al. [27]	y	y	n/a	y	y	n/a	n/a	y	n/a	y	n/a	y	y	100
Hardacker et al. [28]	y	y	n/a	y	n	n/a	n/a	n	n/a	n	n/a	y	n	50
Harrison et al. [2]	n	y	n/a	y	y	n/a	n/a	y	n/a	y	n/a	y	y	87.5
Iunes et al. [23]	y	y	n/a	y	n	n	n/a	y	n/a	y	n/a	y	y	77.8
Iyer et al. [29]	y	y	n/a	n	n/a	n/a	n/a	n	n/a	n	n/a	y	y	57.1
Janusz et al. [30]	y	y	n/a	y	y	n/a	n/a	y	n/a	y	n/a	y	y	100
Lee et al. [31]	y	n	n/a	n	n/a	n/a	n/a	n	n/a	y	n/a	y	y	57.1
Ohara et al. [3]	y	y	n/a	y	n	n/a	n/a	n	n/a	y	n/a	y	n	71.4
Park et al. [20]	y	y	n/a	y	n/a	n/a	n/a	n	n/a	y	n/a	y	y	85.7
Park et al. [19]	y	y	n/a	n	y	n/a	n/a	y	n/a	y	n/a	y	y	87.5
Plaugher; Cremata; Phillips [21]	n	n	n/a	y	n	n/a	n/a	n	n/a	y	n/a	y	n	37.5
Silber et al. [22]	n	y	n/a	y	y	n/a	n/a	y	n/a	y	n/a	y	n	75
Weber et al. [18]	y	n	n/a	n/a	n	y	n/a	y	n/a	y	n/a	y	y	85.7
Zhou et al. [24]	y	y	n/a	n	n	n	n/a	n	n/a	y	n/a	y	y	55.6
% yes answers	81	61.9	100	68.8	41.2	33.3	100	4.76	100	90.5	100	100	81	

1.Description of the sample; 2. Raters characterization; 3. Explanation of the reference standard; 4. Inter-rater blindness; 5. Intra-rater blindness; 6. Randomization of evaluators or subjects; 7. Period of time between the test collection; 8. Time interval between repeated measures; 9. The studied test is not part of the gold standard; 10. Description of the collection procedures from experimental test; 11. Description of the gold standard collection procedures; 12. Description of cases of sample loss; 13. Adequacy of the statistical method. y = yes; n = no; n/a = not applicable; % = final score reached by the study.

Table 6
Meta-analysis of inter-rater reliability of Cobb method (inferior C2 - inferior C7) in radiography.

Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Abelin-Genevois et al. [25]		0.950	0.932	0.964	150	13.1	
Côté et al. [26]		0.960	0.917	0.981	30	11.4	
Harrison et al. [2]		0.920	0.837	0.962	30	11.4	
Janusz et al. [30]		0.920	0.857	0.956	44	12.1	
Park et al. [20]		0.980	0.970	0.986	101	12.9	
Park et al. [20]		0.970	0.956	0.980	101	12.9	
Park et al. [19]		0.672	0.548	0.767	100	12.9	
Zhou et al. [24]		0.960	0.948	0.969	218	13.3	
Total (random effect)		0.943	0.896	0.969	774	100.0	

Heterogeneity: $I^2 = 95\%$

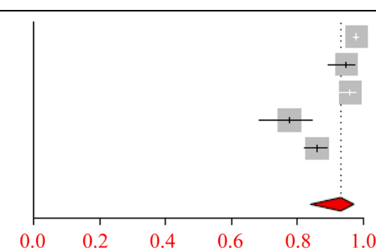


Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

Table 7
Meta-analysis of intra-rater reliability of Cobb method (inferior C2 – inferior C7) in radiography.

Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Abelin-Genevois et al. [25]		0.980	0.972	0.985	150	20.7	
Harrison et al. [2]		0.950	0.897	0.976	30	18.5	
Janusz et al. [30]		0.960	0.927	0.978	44	19.4	
Park et al. [19]		0.777	0.685	0.845	100	20.4	
Zhou et al. [24]		0.860	0.821	0.891	218	20.9	
Total (random effect)		0.932	0.841	0.972	542	100.0	

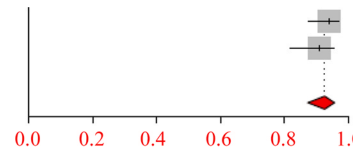
Heterogeneity: $I^2 = 97\%$



Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

Table 8
Meta-analysis of inter-rater reliability of Cobb method (middle C1 - inferior C7) in radiography.

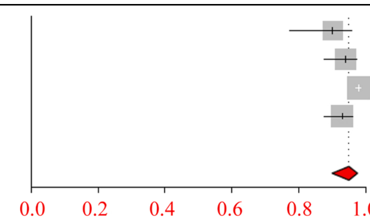
Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Côté et al. [26]		0.940	0.877	0.971	30	50.0	
Harrison et al. [2]		0.910	0.818	0.957	30	50.0	
Total (random effect)		0.926	0.874	0.958	60	100.0	
Heterogeneity: $I^2 = 0\%$							



Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

Table 9
Meta-analysis of inter-rater reliability of absolute rotation angle in radiography.

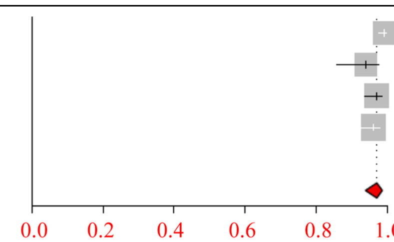
Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Gadotti et al. [27]		0.900	0.771	0.958	22	21.2	
Harrison et al. [2]		0.940	0.877	0.971	30	23.5	
Iyer et al. [29]		0.980	0.971	0.986	115	29.4	
Janusz et al. [30]		0.930	0.875	0.961	44	25.9	
Total (random effect)		0.949	0.899	0.975	211	100.0	
Heterogeneity: $I^2 = 86\%$							



Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

Table 10
Meta-analysis of intra-rater reliability of absolute rotation angle in radiography.

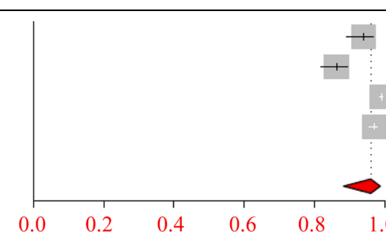
Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Gadotti et al. [27]		0.990	0.976	0.996	22	23.1	
Gadotti et al. [27]		0.940	0.859	0.975	22	23.1	
Harrison et al. [2]		0.970	0.937	0.986	30	25.6	
Janusz et al. [30]		0.960	0.927	0.978	44	28.2	
Total (random effect)		0.970	0.938	0.986	118	100.0	
Heterogeneity: $I^2 = 67\%$							



Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

Table 11
Meta-analysis of inter-rater reliability of Gore angle (C2-C7) in radiography.

Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Janusz et al. [30]		0.940	0.892	0.967	44	24.1	
Lee et al. [31]		0.862	0.819	0.895	181	25.5	
Park et al. [20]		0.990	0.985	0.993	101	25.2	
Park et al. [20]		0.970	0.956	0.980	101	25.2	
Total (random effect)		0.960	0.882	0.987	427	100.0	
Heterogeneity: $I^2 = 98\%$							



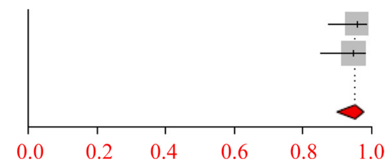
Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

However, ICC does not provide an estimate of the precision of measurement. The standard error of measurement (SEM) provides such an estimate and is independent of the population from it was determined [47,48]. According to Denegar and Ball [48], it is important to provide both values, because a high ICC may not reflect an acceptable measurement if the SEM suggests that the precision of the measurement is not acceptable for the intended purpose. Only five studies included in this review reported SEM values.

Four limitations of the study need to be emphasized. Firstly, the meta-analysis was performed using values from different forms of ICC,

Table 12
Meta-analysis of intra-rater reliability of cervical arrow in photogrammetry.

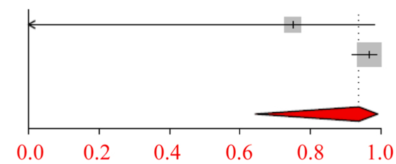
Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Furlanetto et al. [35]		0.958	0.875	0.986	15	48.0	
Furlanetto et al. [17]		0.948	0.853	0.982	16	52.0	
Total (random effect)		0.953	0.900	0.978	31	100.0	
Heterogeneity: $I^2 = 0\%$							



Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

Table 13
Meta-analysis of intra-rater reliability of cervical lordosis in photogrammetry.

Author	Model	Correlation (ICC)	Lower limit	Upper limit	n	Weight (%)	Correlation and 95 %CI
Belli et al. [34]		0.750	−0.391	0.982	5	30.0	
Iunes et al. [23]		0.966	0.917	0.986	21	70.0	
Total (random effect)		0.937	0.643	0.990	26	100.0	
Heterogeneity: $I^2 = 50\%$							



Legend: ICC = Intraclass Correlation Coefficient; n = sample size; I^2 = Higgins Inconsistence test; CI = Confidence Interval.

because it is uncommon to find all necessary information about the ICC reported in the studies. There are several forms of ICC, six according to Shrout and Fleiss [49], and ten according to McGraw and Wong [50]. The choice of which one to use is not always obvious [47]. Each form of ICC varies according to the reliability analysis and the randomness of the raters (model selection), the number of raters/measurements (type selection), and the relevance of variability between the raters (definition selection) [51]. Koo and Li [51] suggest reporting the following items: software information, “model,” “type,” and “definition” selections; in addition, both ICC estimates and their 95 % confidence intervals. The authors also suggest interpreting the results with caution when ICC information is missing.

The second limitation is related to the fact that the ICC values are dependent on the population from which it was determined. A meta-analysis of absolute measurements of reliability would be more informative. However, in general, studies do not provide this information.

The third limitation concerns the fact that reliability assessment is not often a study objective, but rather a part of the methodology. For this reason, some studies that evaluated the reliability of measurements of cervical posture may not have been found in the systematic search because this analysis was not mentioned in the title or abstract. Furthermore, the fourth limitation was the large number of studies included from other resources. This indicates the search key used may not have been the most appropriate to find all the studies available on the subject. Thus, the results of this systematic review should be interpreted with caution, since there may be more studies on the subject that have not been included.

5. Conclusion

Many methods of assessing cervical spine posture in the sagittal plane were found. Sixteen based on radiography and four on photogrammetry. The meta-analysis showed the photogrammetry-based cervical arrow and cervical lordosis methods present very high intra-rater reliability. For radiography, the meta-analysis also showed the Cobb method (inferior C2 - inferior C7), Cobb method (middle C1 - inferior C7), absolute rotation angle, and Gore angle (C2-C7) present very high inter-rater reliability, and the Cobb method (inferior C2 - inferior C7) and absolute rotation angle present very high intra-rater reliability. However, the results showed high heterogeneity and/or few studies were included in the analyzes. Regarding validity, it is suggested that further studies be conducted to investigate the degree of accuracy of the photogrammetry-based methods.

Disclosure statement

All authors have approved the final article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors have no conflict of interest to disclose.

Acknowledgments

No acknowledgments to declare.

References

[1] J. Cobb, Outlines for the Study of Scoliosis, Ann Arbor: MI, 1948.
[2] D. Harrison, D. Harrison, R. Cailliet, S. Troyanovich, T. Janik, B. Holland, Cobb method or harrison posterior tangent method: which to choose for lateral cervical radiographic analysis, Spine (Phila. Pa. 1976) 25 (2000) 2072–2078, https://doi.org/10.1097/00007632-200008150-00011.
[3] A. Ohara, K. Miyamoto, T. Naganawa, K. Matsumoto, K. Shimizu, Reliabilities of and correlations among five standard methods of assessing the sagittal alignment of the cervical spine, Spine (Phila. Pa. 1976) 31 (2006) 2585–2591, https://doi.org/10.1097/01.brs.0000240656.79060.18.
[4] A. Ishihara, Roentgenographic studies on the normal pattern of the cervical curvature, Seikeigeka Gakkai Zasshi 42 (1968) 1033–1044.
[5] D. Gore, S. Sepic, G. Gardner, Roentgenographic finding of the cervical spine in asymptomatic people, Spine (Phila. Pa. 1976) 11 (1986) 521–524.
[6] D. Polly, F. Kilkelly, K. McHale, L. Asplund, M. Mulligan, A. Chang, Measurement of lumbar lordosis: evaluation of intraobserver, interobserver, and technique variability, Spine (Phila. Pa. 1976) 21 (1996) 1530–1536, https://doi.org/10.1097/00007632-199607010-00008.
[7] A. Watson, Procedure for the production of high quality photographs suitable for the recording and evaluation of posture, Rev. Fisioter. 5 (1998) 20–26.

[8] N. Dunk, J. Lalonde, J. Callaghan, Implications for the use of postural analysis as a clinical diagnostic tool: reliability of quantifying upright standing spinal postures from photographic images, J. Manipulative Physiol. Ther. 28 (2005) 386–392, https://doi.org/10.1016/j.jmpt.2005.06.006.

[9] M. Leroux, K. Zabjek, G. Simard, J. Badeaux, C. Coillard, C. Rivard, A noninvasive anthropometric technique for measuring kyphosis and lordosis, Spine (Phila. Pa. 1976) 25 (2003) 1689–1694, https://doi.org/10.1097/00007632-200007010-00012.

[10] T. Furlanetto, J. Sedrez, C. Candotti, J. Loss, Photogrammetry as a tool for the postural evaluation of the spine: a systematic review, World J. Orthop. 7 (2016) 136, https://doi.org/10.5312/wjo.v7.i2.136.

[11] J. Sedrez, C. Candotti, T. Furlanetto, J. Loss, Non-invasive postural assessment of the spine in the sagittal plane: a systematic review, Motricidade 12 (2016) 140–154. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=119227807&amp%0Alang=es&site=eds-live.

[12] M. Bédard, N. Martin, P. Krueger, K. Brazil, Assessing reproducibility of data obtained with instruments based on continuous measurements, Exp. Aging Res. 26 (2000) 353–365, https://doi.org/10.1080/036107300750015741.

[13] D. Stroup, J. Berlin, S. Morton, I. Olkin, G. Williamson, D. Rennie, D. Moher, B. Becker, T. Sipe, S. Thacker, Meta-analysis of observational studies in epidemiology, JAMA. 283 (2000).

[14] Y. Brink, Q. Louw, Clinical instruments: reliability and validity critical appraisal, J. Eval. Clin. Pract. 18 (2012) 1126–1132, https://doi.org/10.1111/j.1365-2753.2011.01707.x.

[15] L. Cohen, S. Kobayashi, M. Simic, S. Dennis, K. Refshauge, E. Pappas, Non-radiographic methods of measuring global sagittal balance: a systematic review, Scoliosis Spinal Disord. 12 (2017) 1–12, https://doi.org/10.1186/s13013-017-0135-x.

[16] J. Higgins, S. Thompson, J. Deeks, D. Altman, Measuring inconsistency in meta-analyses, BMJ Online. 327 (2003) 557–560, https://doi.org/10.1136/bmj.327.7414.557.

[17] T. Furlanetto, C. Candotti, J. Sedrez, M. Noll, J. Loss, Evaluation of the precision and accuracy of the DIPA software postural assessment protocol, Eur. J. Physiother. 19 (2017) 179–184, https://doi.org/10.1080/21679169.2017.1312516.

[18] P. Weber, E. Corrêa, J. Milanesi, J. Soares, M. Trevisan, Craniocervical posture: cephalometric and biophotogrammetric analysis, Brazilian J. Oral Sci. 11 (2012) 416–421, https://doi.org/10.20396/bjos.v11i3.8641399.

[19] M. Park, S. Moon, H. Lee, S. Kim, T. Kim, S. Lee, K. Riew, The effect of age on cervical sagittal alignment, Spine (Phila. Pa. 1976) 38 (2013) 458–463, https://doi.org/10.1097/BRS.0b013e31828802c2.

[20] S. Park, K. Song, S. Park, H. Kang, K. Riew, Does whole-spine lateral radiograph with clavicle positioning reflect the correct cervical sagittal alignment? Eur. Spine J. 24 (2015) 57–62, https://doi.org/10.1007/s00586-014-3525-2.

[21] G. Plaugher, E. Cremata, R. Phillips, A retrospective consecutive case analysis of pretreatment and comparative static radiological parameters following chiropractic adjustments, J. Manipulat. Physiol. Ther. 13 (1990) 498–506.

[22] J. Silber, J. Lipetz, V. Hayes, B. Lonner, Measurement variability in the assessment of sagittal alignment of the cervical spine, J. Spinal Disord. Tech. 17 (2004) 301–305.

[23] D. Iunes, F. Castro, H. Salgado, I. Moura, A. Oliveira, D. Grossi, Confiabilidade intra e interexaminadores e repetibilidade da avaliação postural pela fotogrametria, Rev. Bras. Fisioter. 9 (2005) 327–334.

[24] S. Zhou, F. Xu, W. Wang, D. Zou, Z. Sun, W. Li, Age-based normal sagittal alignment in Chinese asymptomatic adults: establishment of the relationships between pelvic incidence and other parameters, Eur. Spine J. 29 (2020) 396–404, https://doi.org/10.1007/s00586-019-06178-9.

[25] K. Abelin-Genevois, A. Idjerouidene, P. Roussouly, J.M. Vital, C. Garin, Cervical spine alignment in the pediatric population: a radiographic normative study of 150 asymptomatic patients, Eur. Spine J. 23 (2014) 1442–1448, https://doi.org/10.1007/s00586-013-3150-5.

[26] P. Côté, J.D. Cassidy, K. Yong-Hing, J. Sibley, J. Loewy, Apophysial joint degeneration, disc degeneration, and sagittal curve of the cervical spine. Can they be measured reliably on radiographs? Spine (Phila Pa 1976) 22 (1997) 859–864. http://pesquisa.bvsalud.org/portal/resource/pt/mdl-9127918.

[27] I. Gadotti, S. Armijo-Olivo, A. Silveira, D. Magee, Reliability of the craniocervical posture assessment: visual and angular measurements using photographs and radiographs, J. Manipulative Physiol. Ther. 36 (2013) 619–625, https://doi.org/10.1016/j.jmpt.2013.09.002.

[28] J. Hardacker, R. Shuford, P. Capicotto, P. Pryor, Radiographic standing cervical segmental alignment in adult volunteers without neck symptoms, Spine (Phila. Pa. 1976) 22 (1997) 1472–1480, https://doi.org/10.1097/00007632-199707010-00009.

[29] S. Iyer, L. Lenke, V. Nemani, M. Fu, G. Shifflett, T. Albert, B. Sides, L. Metz, M. Cunningham, H. Kim, Variations in occipitocervical and cervicothoracic alignment parameters based on age: a prospective study of asymptomatic volunteers using full-body radiographs, Spine (Phila. Pa.1976) (2016), https://doi.org/10.1097/BRS.0000000000001644.

[30] P. Janusz, M. Tyrakowski, H. Yu, K. Siemionow, Reliability of cervical lordosis measurement techniques on long-cassette radiographs, Eur. Spine J. (2015), https://doi.org/10.1007/s00586-015-4345-8.

[31] C. Lee, H. Noh, D. Lee, C. Hwang, H. Kim, S. Cho, Analysis of sagittal spinal alignment in 181 asymptomatic children, J. Spinal Disord. Tech. 25 (2012) 259–263.

[32] S. Armijo-Olivo, X. Jara, N. Castillo, L. Alfonso, A. Schilling, E. Valenzuela, R. Frugone, D. Magee, A comparison of the head and cervical posture between the self-balanced position and the Frankfurt method, J. Oral Rehabil. 33 (2006) 194–201, https://doi.org/10.1111/j.1365-2842.2005.01554.x.

[33] L. Bernal, M. Harold, C. Herrera, C. Montoya, Y. Herrera, Craniocervical posture in children with class I, II and III skeletal relationships, Pesqui. Bras. Odontopediatria Clin. Integr. 17 (2017), https://doi.org/10.4034/PBOCI.2017.171.07.

[34] J. Belli, T. Chaves, A. De Oliveira, D. Grossi, Analysis of body posture in children with mild to moderate asthma, Eur. J. Pediatr. 168 (2009) 1207–1216, https://doi.org/10.1007/s00431-008-0911-y.

[35] T. Furlanetto, F. Chaise, C. Candotti, J. Loss, Fidedignidade de um protocolo de avaliação postural, Rev. Da Educ. Fis. 22 (2011) 411–419, https://doi.org/10.4025/reveducfis.v22i3.10124.

[36] AERA, Standards for Educational and Psychological Testing, American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Washington, DC, 1999.

[37] J.J. Bartko, Measurement and reliability: statistical thinking considerations, Schizophr. Bull. 17 (1991) 483–489.

[38] P.M.N.M. de Albuquerque, G.G. de Alencar, D.A. de Oliveira, G.R. de Siqueira, Concordance and reliability of photogrammetric protocols for measuring the cervical lordosis angle: a systematic review of the literature, J. Manipulative Physiol. Ther. 41 (2018) 71–80, https://doi.org/10.1016/j.jmpt.2017.08.004.

[39] K. Refshauge, M. Goodsell, M. Lee, The relationship between surface contour and vertebral body measures of upper spine curvature, Spine (Phila. Pa. 1976) 19 (1994) 2180–2185.

[40] D. Hinkle, W. Wiersma, S. Jurs, Applied Statistics for the Behavioural Sciences, Chicago, 1979.

[41] D. Watson, P. Trott, Cervical Headache: an Investigation of Natural Head Posture and Upper Cervical Flexor Muscle Performance, Cephalalgia, 1993.

[42] S. Raine, L. Twomey, Posture of the head, shoulders and thoracic spine in comfortable erect standing, Aust. J. Physiother. 40 (1994) 25–32, https://doi.org/10.1016/S0004-9514(14)60451-7.

[43] R. Ruivo, P. Pezarat-Correia, A. Carita, Intrarater and interrater reliability of photographic measurement of upper-body standing posture of adolescents, J. Manipulative Physiol. Ther. 38 (2015) 74–80, https://doi.org/10.1016/j.jmpt.2014.10.009.

[44] S. Van Niekerk, Q. Louw, C. Vaughan, K. Grimmer-Somers, K. Schreve, Photographic measurement of upper-body sitting posture of high school students: a reliability and validity study, BMC Musculoskelet. Disord. 9 (2008) 113, https://doi.org/10.1186/1471-2474-9-113.

[45] J.M. Bland, D.G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet. 1 (1986) 307–310.

[46] T.A. Baumgartner, Measurement in physical education and exercise science estimating the stability reliability of a score, Meas. Phys. Educ. Exerc. Sci. 7841 (2009) 37–41, https://doi.org/10.1207/S15327841Mpee0403.

[47] J.P. Weir, Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM, J. Strength Cond. Res. 19 (2005) 231–240, https://doi.org/10.1519/15184.1.

[48] C.R. Denegar, D.W. Ball, Assessing reliability and precision of measurement: an introduction to intraclass correlation and standard error of measurement, J. Sport Rehabil. 2 (1993) 35–42, https://doi.org/10.1123/jsr.2.1.35.

[49] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, Psychol. Bull. 86 (1979) 420–428. http://www.ncbi.nlm.nih.gov/pubmed/18839484.

[50] K.O. McGraw, S.P. Wong, Forming inferences about some intraclass correlation coefficients, Psychol. Methods 1 (1996) 30–46, https://doi.org/10.1037/1082-989X.1.1.30.

[51] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, J. Chiropr. Med. 15 (2016) 155–163, https://doi.org/10.1016/j.jcm.2016.02.012.