Motricidade © Edições Desafio Singular
2016, vol. **12**, n. **2**, pp. **140-154** http://dx.doi.org/10.6063/motricidade.6470

Non-invasive postural assessment of the spine in the sagittal plane: a systematic review

Juliana Adami Sedrez[1*], Cláudia Tarragô Candotti[1], Tássia Silveira Furlanetto[1], Jefferson Fagundes Loss[1]

REVIEW ARTICLE

ABSTRACT

The objective of this review was to examine the scientific evidence regarding the aspects of validation in non-invasive methods of assessing the spine in the sagittal plane. A systematic search was conducted in following data bases Scopus, Science Direct, PubMed and Medline. To be included the papers must have: conducted a non-invasive assessment of thoracic kyphosis and/or lumbar lordosis; evaluated at least one aspect of validity; been written in English; and been published in the previously three decades. Papers that score less than three in the QUADAS scale were excluded. Initially, 70 articles were pre-selected. Of this, 52 were finally included as they met the quality criterion. Based on this review, the following techniques/instruments were found to present satisfactory results for all aspects of validity in the assessment of thoracic kyphosis: photogrammetry, flexible ruler, archometer, and DeBrunner'skyphometer. Similarly, photogrammetry, inclinometer, flexible ruler, archometer and kypholordometer were found to present satisfactory results in the assessment of lumbar lordosis. Therefore, it is suggested that these instruments be adopted as first choice for evaluating the spine in the sagittal plane, since they present adequate reproducibility and concurrent validity.

Keywords: evaluation; posture; methods; reproducibility of results; validity of tests

INTRODUCTION

The sagittal plain of the spine, in physiological conditions, is composed of a succession of opposed harmonious curves: lumbar lordosis, thoracic kyphosis, cervical lordosis. In several studies, the increase in thoracic curvature has been associated with back pain (Ensrud, Black, Harris, Ettinger, & Cummings, 1997), increased risk of fracture (Huang, Barrett-Connor, Greendale, & Kado, 2006) and falls (Kado, Huang, Nguyen, Barrett-Connor, & Greendale, 2007), as well as provoking reduction in quality of life (Imagama et al., 2011) and increased mortality (Kado et al., 2009). The reduction of lumbar lordosis also has been associated with the presence of back pain (Chaléat-Valayer et al., 2011), an increased risk of falls (Ishikawa, Miyakoshi, Kasukawa, Hongo, & Shimada, 2013), and reduced quality of life (Imagama et al., 2011).

Therefore, assessment of spinal curvature is an important factor in both the clinical and research environments. Clinically such an assessment can be used to select the appropriate treatment, since therapies are prescribed based on the degree of curvature and/or its progression. In the research environment, assessing spinal curvatures is essential to ensure that the results of treatments in intervention studies can be adequately reported.

Hence, there has been a growing interest in non-invasive quantitative methods of evaluating the spine in the sagittal plane, since anatomical and biomechanical assessment of the vertebral column frequently requires quantitative data. Such non-invasive methods provide several advantages such as low cost, reduced technical

---

Manuscript received at January 26th 2015; Accepted at January 5th 2016

[1] *Federal University of Rio Grande do Sul, Porto Alegre, Brasil.*
* *Corresponding Author: Escola Superior de Educação Física, Universidade Federal do Rio Grande do Sul, Rua Felizardo, 750, Jardim Botânico, CEP: 90690-200 - Porto Alegre, RS, Brasil. E-mail: julianasedrez@gmail.com*

complexity and absence of side effects. Moreover, the ideal instrument must be effective, precise, small in size, easy to use and affordable (D'Osualdo, Schierano, & Iannis, 1997). Recently, a systematic review of instruments for the evaluation thoracic kyphosis was published (Barrett, McCreesh, & Lewis, 2014), however, that study did not include all the instruments capable of evaluating thoracic curvature nor those for the evaluation of lumbar lordosis.

Therefore, the aim of this systematic review was to verify the scientific evidence regarding the validation of different non-invasive methods of evaluating the spine in the sagittal plane. This will help health professionals when choosing the most suitable instrument for use in different clinical situations or scientific research.

METHOD

In April 2013, a systematic search was conducted for scientific articles in the following databases: Scopus, Science Direct, Pubmed and Medline. The search terms used were: "*Noninvasive instrument*" OR "*Non-invasive Monitoring*" OR "*Measurement*" OR "*Measurements*" OR "*Postural Assessment*" OR "*Postural Evaluation Methods*" OR "*Non-radiological Measures*" AND "*Spine Curvatures*" OR "*Lumbar Curvatures*" OR "*Thoracic Curvatures*" OR "*Thoracic Curve*" OR "*Lordosis Curve*" OR "*Thoracic Kyphosis*" OR "*Lumbar Lordosis*" OR "*Kyphosis*" OR "*Lordosis*" OR "*Postural Assessment*" AND "*Validation*" OR "*Validity*" OR "*Repeatability*" OR "*Reproducibility*" OR "*Reliability*" OR "*Accurate*" OR "*Accuracy*". To be included in this systematic revision the articles found were required to meet the following inclusion criteria: (a) perform a non-invasive evaluation of the spinal curvatures; (b) perform an evaluation in the sagittal plane of thoracic kyphosis or lumbar lordosis; (c) evaluate some validation aspect; (d) to be written in English and (e) to have been published in the last three decades. All the search, selection, quality evaluation, reading and data extraction procedures were carried out by two independent evaluators. In the case of any divergence of opinion between the evaluators, a third evaluator was invited to analyze the article.

Firstly, the articles were selected based on the titles and abstracts. Those articles considered for inclusion in the review were read in full. After, only those articles that met all the above-mentioned inclusion criteria were included in this systematic review. Furthermore, the bibliography of each included article was checked with the aim of find any articles not found in the electronic search.

The QUADAS (*Quality Assessment of Diagnostic Accuracy Studies*) scale was used to evaluate the quality of the articles. This consists of a questionnaire with 14 items which were responded as "yes", "no" or "unclear". In the present study, 11 items were applicable to postural evaluation instruments (Whiting et al., 2004). A minimum of three points in the QUADAS scale was used as an exclusion criterion in this systematic review.

Moreover, with the aim of classifying the scientific evidence contained in the articles, the following rule was used based on QUADAS scale: (a) articles with three to five points were classified as presenting "poor evidence"; (b) articles with six to eight points were classified as presenting "moderate evidence"; and (c) articles with nine to eleven points were classified as presenting "strong evidence".

Given the variation in the terminology used in the studies, to facilitate comparison of their results, in this systematic review the terminology was standardized as follows: repeatability refers to the degree of agreement obtained between successive evaluations conducted by the same evaluator (short period of time); intra-evaluator reproducibility refers to the degree of agreement obtained between evaluations conducted by the same evaluator at different times (minimum interval - one day); inter-evaluator reproducibility refers to the degree of agreement obtained between evaluations conducted by different evaluators; and validity refers to the agreement between the measurements obtained using the instrument being tested and those obtained using the gold standard instrument (Joint Committee for Guides in Metrology, 2012).

RESULTS

Initially, 597 articles were found in a database search, of which 57 were included. When analyzing the bibliographical references in the selected articles, a further 19 articles were obtained, given a total of 76 articles that met the inclusion criteria of this systematic review (Figure 1).

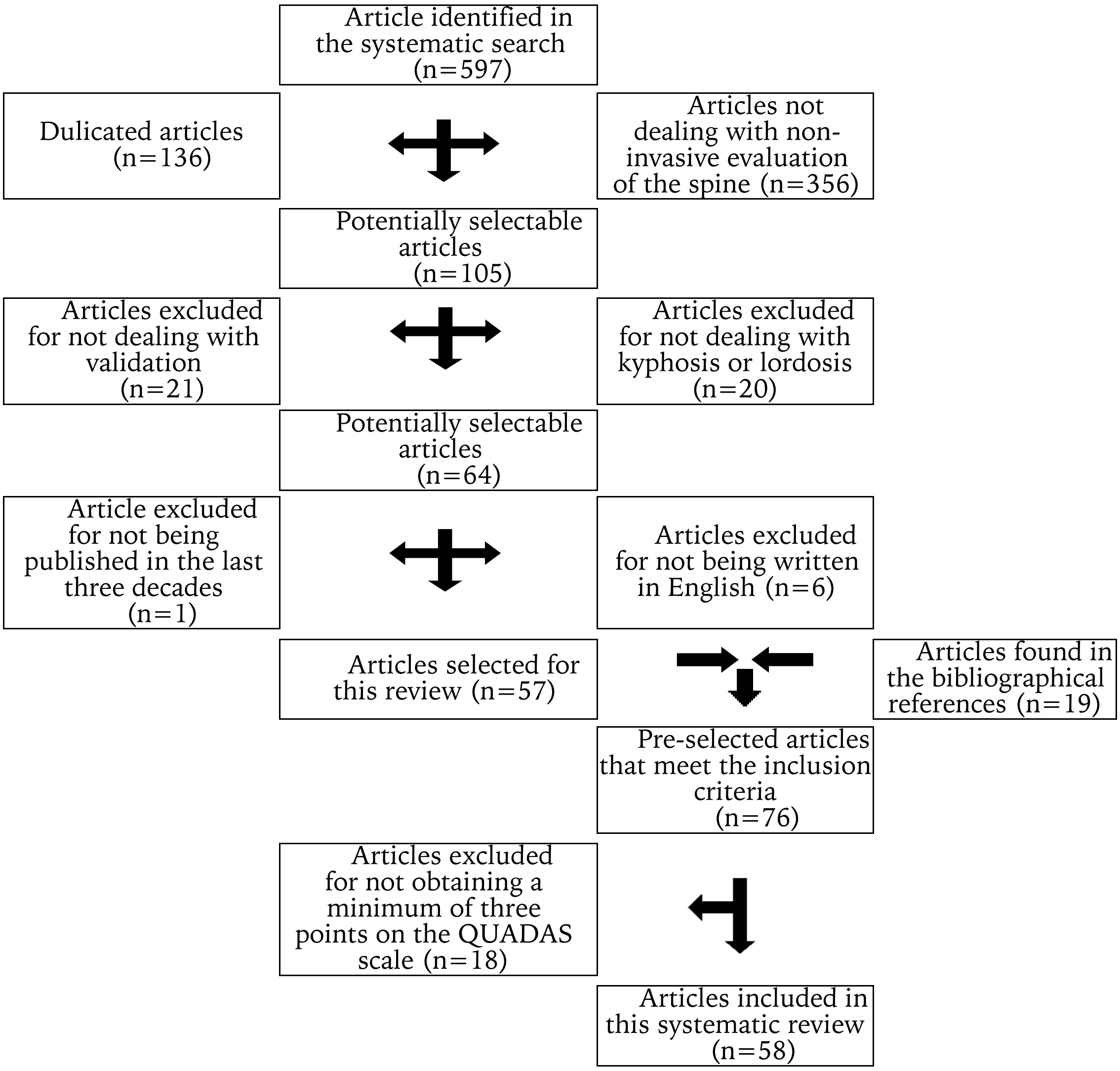

Figure 1. Flowchart of the article selection process

The pre-selected articles were evaluated regarding their methodological quality using the QUADAS scale (Table 1). Of the 76 evaluated articles, 18 were excluded because they did not obtain the stipulated minimum of three points on the scale, thus 58 articles were finally included in this systematic review. Regarding the quality of the scientific evidence in the articles, 29 presented poor scientific evidence, 17 moderate evidence and 12 strong evidence.

Table 2 presents the validation objective, the instrument used, the methodology and the results for the 58 articles included in this review. The validation objective shown is not necessarily the primary objective of the corresponding paper but instead that which deals with aspects of validation in relation to non-invasive instruments used to evaluate spinal posture in sagittal plane.

Table 1

Results of the evaluation of the article quality using the QUADAS scale

Articles 1st author (year)	QUADAS scale criteria											Total (n° of ✓)	Evidence classification
	1	2	3	4	5	6	8	9	10	11	14		
Finestone (2013)	✓	×	✓	?	✓	✓	✓	✓	✓	?	×	7	Moderate
Ranavolo (2013)	×	✓	✓	✓	✓	✓	✓	✓	?	?	×	7	Moderate
Schülein (2013)	✓	✓	—	—	×	—	✓	—	?	—	×	3	Poor
Celan (2012)	×	✓	—	—	×	—	✓	—	—	—	×	2	Excluded
Consmüller (2012)	✓	✓	—	—	×	—	✓	—	—	—	×	3	Poor
Czaprowski (2012)	✓	✓	—	—	×	—	✓	—	✓	—	×	4	Poor
Edmondston (2012)	✓	✓	✓	✓	✓	✓	✓	✓	✓	?	×	9	Strong
Fölsch (2012)	✓	✓	—	—	×	—	✓	—	?	—	✓	4	Poor
Fortin (2012)	?	✓	—	—	×	—	✓	—	?	—	×	2	Excluded
Gravina (2012)	✓	✓	✓	✓	✓	✓	?	×	?	?	×	6	Moderate
O’Sullivan (2012)	?	✓	✓	✓	✓	✓	✓	✓	?	?	✓	8	Moderate
Oliveira (2012)	✓	✓	✓	✓	✓	✓	✓	✓	✓	?	×	9	Strong
Rankine (2012)	×	×	—	—	×	—	×	—	×	—	×	0	Excluded
Saad (2012)	?	✓	—	—	×	—	✓	—	✓	—	✓	5	Poor
Van Blommestein (2012)	✓	✓	—	—	×	—	✓	—	✓	—	×	4	Poor
Williams (2012)	?	✓	—	—	×	—	✓	—	—	—	×	2	Excluded
Zaina (2012)	?	×	—	—	×	—	?	—	—	—	×	0	Excluded
Chaise (2011)	✓	✓	✓	✓	✓	✓	✓	✓	✓	?	×	10	Strong
Greendale (2011)	✓	✓	✓	✓	✓	✓	✓	✓	✓	?	✓	10	Strong
Letafatkar (2011)	✓	✓	✓	?	✓	✓	✓	✓	?	?	×	7	Moderate
MacIntyre (2011)	✓	✓	—	—	×	—	✓	—	✓	—	×	4	Poor
Dunleavy (2010)	×	✓	—	—	×	—	✓	—	✓	—	×	3	Poor
Ferreira (2010)	?	×	×	—	×	—	✓	—	✓	—	×	2	Excluded
Fortin (2010)	✓	✓	✓	×	✓	✓	✓	×	✓	✓	×	8	Strong
Lewis (2010)	✓	✓	—	—	×	—	✓	—	✓	—	×	4	Poor
Melvin (2010)	✓	✓	—	—	×	—	✓	—	—	—	×	3	Poor
O'Sullivan (2010)	✓	✓	—	—	×	—	✓	—	—	—	×	3	Poor
Perriman (2010)	✓	✓	✓	✓	✓	✓	✓	✓	?	?	×	8	Moderate
Sheeran (2010)	✓	✓	—	—	×	—	✓	—	?	—	×	3	Poor
Singh (2010)	✓	✓	—	—	×	—	✓	—	?	—	×	3	Poor
Williams (2010)	?	✓	✓	✓	✓	✓	✓	✓	?	?	×	7	Moderate
McAlpine (2009)	?	✓	—	—	×	—	✓	—	—	—	×	2	Excluded
Seidi (2009)	✓	×	✓	?	×	✓	✓	✓	?	?	×	5	Poor
Souza (2009)	✓	✓	✓	?	✓	✓	✓	✓	✓	✓	✓	10	Strong
Kellis (2008)	✓	✓	—	—	×	—	✓	—	✓	—	×	4	Poor
Perry (2008)	?	✓	—	—	×	—	✓	—	—	—	×	2	Excluded
Rajabi (2008)	?	✓	✓	?	✓	✓	✓	✓	?	?	×	6	Moderate
Van Niekerk (2008)	✓	✓	✓	✓	✓	✓	✓	✓	×	×	✓	9	Strong
Zubovic (2008)	✓	×	—	—	×	—	×	—	—	—	×	1	Excluded
Normand (2007)	?	×	—	—	×	—	✓	—	✓	—	×	2	Excluded
Teixeira (2007)	✓	✓	✓	✓	✓	✓	✓	×	?	✓	×	8	Moderate
Kado (2006)	✓	✓	✓	?	✓	✓	✓	✓	?	?	✓	8	Moderate
Campbell-Kyureghyan (2005)	✓	✓	✓	?	✓	×	✓	✓	?	?	✓	7	Moderate
Dunk (2005)	✓	✓	—	—	×	—	✓	—	—	—	×	3	Poor
Dunk (2004)	?	✓	—	—	×	—	✓	—	—	—	×	2	Excluded
Hinman (2004)	✓	✓	—	—	×	—	✓	—	?	—	×	3	Poor
Mannion (2004)	✓	×	—	—	×	—	✓	—	✓	—	×	3	Poor

Table 1 (continued)
Results of the evaluation of the article quality using the QUADAS scale

Articles 1st author (year)	QUADAS scale criteria											Total (n° of ✓)	Evidence classificatio n
	1	2	3	4	5	6	8	9	10	11	14		
D’Osualdo (2002)	×	?	—	—	×	—	✓	—	✓	—	×	2	—
Leroux (2002)	✓	✓	✓	✓	✓	✓	✓	✓	?	?	✓	9	Strong
Norton (2002)	✓	✓	—	—	×	—	✓	—	?	—	×	3	Poor
Korovessis (2001)	✓	✓	✓	?	✓	✓	✓	✓	✓	?	×	8	Moderate
Ng (2001)	✓	✓	—	—	×	—	✓	—	—	—	✓	4	Poor
Arnold (2000)	✓	✓	—	—	×	—	✓	—	✓	—	×	4	Poor
Leroux (2000)	✓	✓	✓	✓	✓	✓	✓	✓	?	?	✓	9	Strong
McGorry (2000)	?	✓	×	?	✓	✓	✓	×	?	?	×	4	Poor
Watson (2000)	?	✓	—	—	×	—	✓	—	—	—	×	2	—
Goh (1999)	✓	×	—	—	×	—	✓	—	—	—	×	2	—
Kovac (1999)	✓	×	✓	?	✓	✓	✓	✓	?	?	×	6	Moderate
Lundon (1998)	✓	✓	✓	?	✓	✓	✓	✓	✓	?	×	8	Moderate
D'Osualdo (1997)	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	×	9	Strong
Whittle (1997)	×	✓	—	—	×	—	✓	—	—	—	×	2	—
Caine (1996)	×	×	—	—	×	—	✓	—	—	—	×	1	—
Levine (1996)	✓	✓	—	—	×	—	✓	—	?	—	×	3	Poor
Walsh (1995)	×	×	✓	✓	✓	✓	×	✓	?	?	×	5	Poor
Youdas (1995)	?	✓	—	—	×	—	✓	—	✓	—	×	3	Poor
Raine (1994)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	11	Strong
Griegel-Morris (1992)	?	✓	—	—	×	—	✓	—	?	?	×	2	—
Bryan (1990)	✓	✓	✓	?	✓	✓	✓	✓	✓	?	×	8	Moderate
Bryan (1989)	✓	✓	✓	✓	✓	✓	✓	✓	?	?	✓	9	Strong
Lovell (1989)	?	✓	—	—	×	—	✓	—	✓	—	✓	4	Poor
Öhlén (1989)	?	×	—	—	×	—	✓	—	✓	—	×	2	—
Walker (1987)	?	✓	—	—	×	—	✓	—	✓	—	×	3	Poor
Adams (1986)	✓	✓	✓	✓	✓	✓	✓	✓	?	✓	×	8	Moderate
Burdett (1986)	✓	✓	✓	✓	✓	✓	✓	✓	?	?	×	8	Moderate
Hart (1986)	?	×	✓	?	×	✓	✓	✓	?	?	×	4	Poor
Mellin (1986)	✓	✓	—	—	×	—	✓	—	—	—	×	3	Poor

QUADAS scale criteria: 1) Was the spectrum of patients representative of the patients who will receive the test in practice? 2) Were selection criteria clearly described? 3) Is the reference standard likely to correctly classify the target condition? 4) Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? 5) Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis? 6) Did patients receive the same reference standard regardless of the index test result? 8) Was the execution of the index test described in sufficient detail to permit replication of the test? 9) Was the execution of the reference standard described in sufficient detail to permit its replication? 10) Were the index test results interpreted without knowledge of the results of the reference standard? 11) Were the reference standard results interpreted without knowledge of the results of the index test? 14) Were withdrawals from the study explained?
RESPONSES TO THE CRITERIA: ✓ = Yes; × = No; ? = Unclear; — = Not applicable.

DISCUSSION

The aim of this systematic review was to examine the scientific evidence related to the validation of alternative methods of non-invasive evaluation of the spine in the sagittal plane. Of the 58 articles included, only 12 were found to present strong scientific evidence according to the QUADAS scale, indicating the poor quality of the methodology applied in the reviewed validation articles. Given this situation, when conducting validation studies, it is necessary to pay attention to the criteria that determine their quality. Therefore it is important to evaluate both, the reproducibility and concurrent validity, thus avoiding any restriction to the applicability of the instrument for diagnostic use. This systematic review identified 22 different evaluation systems, which were categorized and analyzed below: (1) flexible ruler, (2) photogrammetry, (3) inclinometer, (4) spinal mouse, (5) goniometer and eletrogoniometer, (6) Debrunner’s Kyphometer, (7) surface topography, (8) archometer and (9) other instruments with only one validation study each.

Table 2

Summary of 58 studies included in this systematic review

1st author (year)	Validation Objective	Instrument	Methodology	Results
Finestone (2013)	Assess the validity and inter-evaluator reproducibility of the instrument to measure thoracic curvature.	SpineScan	n=28 Each individual was evaluated using X-ray and twice by two evaluators using the SpineScan.	Intra-evaluator reproducibility: variation of 3.2º(9.4º). Inter-evaluator reproducibility: variation between -16º and -24º There was a significant difference in the validity assessment (comparing the results from X-ray and SpineScan)
Ranavolo (2013)	Assess the validity of the instrument to measure thoracic and lumbar curvatures.	Optoelectronic system	n=10 Assessments conducted using optoelectronics recording and X-rays.	There was a significant difference in the validity assessment for both the thoracic and lumbar curvatures
Schülein (2013)	Assess the repeatability and inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Rasterstereographic	n=39 5 evaluators conducted 3 successive assessments.	Inter-evaluator reproducibility: thoracic (ICC between 0.95 and 0.96) and lumbar (ICC between 0.97 and 0.98) Repeatability: thoracic (r between .72 and .96) and lumbar (r between .90 and .98).
Consmüller (2012)	Assess the intra-evaluator reproducibility of the instrument to evaluate thoracic and lumbar curvatures.	Epionics SPINE	n=30 Assessments conducted in 3 days (5 day interval).	Intra-evaluator reproducibility: thoracic (ICC = 0.87) and lumbar (ICC = 0.85).
Czaprowski (2012)	Assess the intra and inter-evaluator reproducibility of the instrument to evaluate thoracic and lumbar curvatures.	Digital Inclinometer (Saunders)	n=30 3 assessments were conducted by 3 evaluators, one of whom repeated the assessment with a one-week interval	Intra-evaluator reproducibility: thoracic (Cronbach's alpha = 0.83) and lumbar (Cronbach's alpha = 0.87). There was a significant difference in the inter-evaluator reproducibility for lumbar curvature.
Edmondston (2012)	Assess the validity and repeatability of the instrument to measure thoracic curvature.	Digital photographs	5 assessments were conducted by 1 evaluator (n=4). Compared with X-rays (n=14)	Validity: r=.76 Repeatability: coefficient of variation of 4.8% and standard error of 0.5°
Fölsch (2012)	Assess the intra-evaluator reproducibility of the instrument to evaluate thoracic curvature.	Zebris CMS 20 ultrasound	n=28 Two 3D assessments were conducted by the same evaluator (24-hour interval).	Intra-evaluator reproducibility: ICC = 0.95
Gravina (2012)	Assess the validity of the instrument to measure the lumbar and thoracic curvatures.	Goniometer (IncliMed)	n=128 The assessments were conducted with goniometer and X-rays.	Validity: thoracic (b=0.89) and lumbar (b=0.52).
O'Sullivan (2012)	Assess the validity and repeatability of the instrument to measure the static and dynamic lumbar curvature.	Spinal posture monitoring device (Body-Guard^{TM})	n=12 Simultaneous assessments were conducted using CODA system and Body-Guard^{TM}. 2 consecutive assessments were conducted by 10 evaluators.	Repeatability: ICC>0.73. Validity: rho=0.88.
Oliveira (2012)	Assess the validity and intra and inter-evaluator reproducibility of the instrument to evaluate thoracic and lumbar curvatures.	Flexible ruler	The assessments were conducted by 3 evaluators (n=15), one of whom repeated the assessment with a one-week interval. Comparison with X-rays (n=47).	Validity: thoracic (b=0.89) and lumbar (b=0.52). Intra-evaluator reproducibility: thoracic (ICC = 0.82) and lumbar (ICC=0.78) Inter-evaluator reproducibility: thoracic (ICC = 0.94) and lumbar (ICC=0.83).

Table 2 (continued)

Summary of 58 studies included in this systematic review.

Saad (2012)	Assess the intra and inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Digital photographs (Corel Draw)	n=20 The assessments were conducted by 2 evaluators on the first day, and after 15 days, by 1 evaluator.	Intra-evaluator reproducibility: thoracic (ICC between 0.93 and 0,95) and lumbar (ICC between 0.85 and 0.90) Inter-evaluator reproducibility: thoracic (ICC = 0.97) and lumbar (ICC between 0.85 and 0.89)
Chaise (2011)	Assess the validity and intra and inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Adapted arcometer	The assessment was conducted by 2 evaluators (n=30), one of whom repeated the evaluation with a one-week interval (n=15). Comparison with X-rays (n=52).	Validity: thoracic (r=.94) and lumbar (r=.71). Intra-evaluator reproducibility: thoracic (ICC = 0.99) and lumbar (ICC=0.85) Intra-evaluator reproducibility: thoracic (ICC = 0.98) and lumbar (ICC=0.89)
Greendale (2011)	Assess the validity, the repeatability and inter-evaluator reproducibility of the instrument to measure thoracic curvature.	Debrunner’s kyphometer and the flexible ruler	The assessment was conducted by 2 evaluators (n=54), one of whom repeated 3 times (n=113). Comparison with X-rays (n=113).	Validity: kyphometer (r=.62) and flexible ruler (r=.68). Repeatability: kyphometer (ICC=0.98) and flexible ruler (ICC=0.96). Inter-evaluator reproducibility: kyphometer (ICC = 0.98) and flexible ruler (ICC=0.96).
Letafatkar (2011)	Assess the validity, the repeatability and inter-evaluator reproducibility of the instrument to measure lumbar curvature.	Digital photographs (AutoCAD software) and the flexible ruler	n=50 2 assessments were conducted by 2 evaluators on the same day using the AutoCAD software and the flexible ruler. Comparison with X-rays.	Repeatability: digital photographs (ICC=0.97 and 0.98) and flexible ruler (ICC=0.62 and 0.69). Inter-evaluator reproducibility: digital photographs (ICC = 0.97) and flexible ruler (ICC=0.54) Validity: digital photographs (ICC=0.94 and 0.96) and flexible ruler (ICC=0.50 and 0.52).
MacIntyre (2011)	Assess the inter-evaluator reproducibility of the instruments to measure thoracic and lumbar curvatures.	Digital inclinometer and the flexible ruler	n=9 3 assessments were conducted by 2 evaluators.	Inter-evaluator reproducibility of digital inclinometer : thoracic (ICC = 0.72) and lumbar (ICC=0.63) Inter-evaluator reproducibility of flexible ruler: lordosis index (ICC = 0.74) and kyphosis index (ICC=0.92)
Dunleavy (2010)	Assess the repeatability and inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Flexible ruler	n=22 3 assessments of length and width of the thoracic and lumbar curvatures were conducted by 2 evaluators on the same day	Repeatability: thoracic and lumbar length, thoracic and lumbar width (ICC between 0.61 and 0.80). Inter-evaluator reproducibility: thoracic and lumbar length, thoracic and lumbar width (ICC between 0.58 and 0.72) Significant differences were found between the lumbar lengths and widths as measured by the evaluators
Fortin (2010)	Assess the validity of photogrametry in comparison to topography; and assess the validity of the instruments/techniques with X-rays to measure the thoracic and lumbar curvatures.	Digital photographs and the surface topography (3D optical digitizers)	n= 70 The assessments were conducted with X-rays, 2D digital photographs and 3D surface topography .	Validity between photographs and topography: thoracic (r=.35) and lumbar (r=.30). Validity between photographs and X-rays: thoracic (r=-.77) and lumbar (r=-.48). Validity between topography and X-rays: high correlations.
Lewis (2010)	Assess the repeatability of the instrument to measure thoracic curvature.	Gravity dependent inclinometers	n= 90 3 consecutive assessments were conducted twice.	Repeatability: ICC between 0.88 and 0.97.

Table 2 (continued)

Summary of 58 studies included in this systematic review.

O'Sullivan (2010)	Assess the intra and inter-evaluator reproducibility of the CODA system to measure lumbar curvature.	CODA system	n=12 The assessments were conducted by 2 evaluators, one of whom repeated the assessment on 2 different days (3-14 day interval).	Intra-evaluator reproducibility: ICC between 0.70 and 0.97. Inter-evaluator reproducibility: ICC between 0.60 and 0.96.
Perriman (2010)	Assess the validity and intra-evaluator reproducibility of the instrument to measure thoracic curvature.	Flexible electrogoniometer	n=12 7 functional activities were performed (one-week interval) and X-rays were used.	Validity: r between .80 and .87. Intra-evaluator reproducibility: ICC between 0.89 and 0.95.
Sheeran (2010)	Assess the repeatability and intra and inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Spinal Wheel	n=17 3 assessments were conducted by 3 evaluators on 2 different days (one-week interval).	Repeatability: thoracic (ICC=0.98) and lumbar (ICC=0.96). Intra-evaluator reproducibility: thoracic (ICC=0.83) and lumbar (ICC=0.71). Inter-evaluator reproducibility: thoracic (ICC=0.98) and lumbar (ICC=0.95).
Singh (2010)	Assess the intra-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Electromagnetic tracking device (Fastrak[a]).	n=52 3 consecutive assessments were conducted by the same evaluator.	Intra-evaluator reproducibility: thoracic (ICC=0.93) and lumbar (ICC=0.98).
Williams (2010)	Assess the repeatability of the instrument to measure lumbar curvature.	Fibre-optic sensors	n=13 2 consecutive assessments were conducted during flexion and lifting. Comparison with VICON.	Validity: flexion (r=.95) and lifting (r=.94). Repeatability: flexion and lifting (average r=.97).
Seidi (2009)	Assess the validity, the repeatability and inter-evaluator reproducibility of the instrument to measure lumbar curvature.	Flexible ruler	2 consecutive assessments (n=20) were conducted by 2 evaluators. Comparison with X-rays (n=10).	Validity: ICC=0.91. Repeatability : ICC=0.89 and 0.92. Inter-evaluator reproducibility : ICC=0.82.
Souza (2009)	Assess the validity, the repeatability and inter-evaluator reproducibility of the instrument to measure lumbar curvature.	Kypholordometer	n=20 2 consecutive assessments were conducted by 3 evaluators. Comparison with X-rays.	Validity: r=.88. Repeatability : ICC between 0.97 and 0.99. Inter-evaluator reproducibility : ICC between 0.89 and 0.98.
Kellis (2008)	Assess the intra and inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Spinal Mouse	n=81 3 assessments were conducted by 3 evaluators on 2 different days (one-week interval).	Intra-evaluator reproducibility: thoracic (ICC between 0.81 and 0.87) and lumbar (ICC between 0.84 and 0.93). Inter-evaluator reproducibility: thoracic (ICC between 0.88 and 0.89) and lumbar (ICC between 0.87 and 0.94).
Rajabi (2008)	Assess the validity of the instrument to measure lumbar curvature.	Flexible ruler	n=10 Assessments were conducted with X-rays and the flexible ruler.	There was a significant difference between the assessment methods.
Van Niekerk (2008)	Assess the validity and the repeatability of the instrument to measure thoracic curvature in the seated position.	Digital photographs	n=39 5 consecutive assessments were conducted. Comparison with X-rays.	Validity: r=.92. Repeatability: ICC=0.96.
Teixeira (2007)	Assess the validity, the repeatability and inter-evaluator reproducibility of the instrument to measure thoracic curvature.	Flexible ruler	n=56 Assessments were conducted by 2 evaluators, one of whom repeated the assessment twice. Comparison with X-rays.	Validity: evaluator 1 (ICC=0.52), evaluator 2 (ICC=0.58) and average of two assessments (ICC=0.90). Repeatability: ICC=0.87 Inter-evaluator reproducibility: ICC=0.94.
Kado (2006)	Assess the validity of the instrument to measure thoracic curvature.	Debrunner ‘s kyphometer	n=120 Assessments were conducted with kyphometer and X-rays.	Validity: ICC=0.68.

Table 2 (continued)

Summary of 58 studies included in this systematic review.

Campbell-Kyureghyan (2005)	Assess the validity of the instrument to measure lumbar curvature.	Electrogoniometer	n=39 Assessments were conducted with electrogoniometer, X-rays (n=15) and magnetic resonance (n=24).	Validity: r^2=.78. The predicted Cobb angle showed an error of 14.5%.
Dunk (2005)	Assess the intra-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Digital photographs	n=20 2 assessments were conducted (one-week interval)	Intra-evaluator reproducibility: thoracic (ICC between 0.63 and 0.71) and lumbar (ICC between 0.63 and 0.72).
Hinman (2004)	Assess the inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Flexible ruler	n=51 Assessments were conducted by 3 novice evaluators.	Inter-evaluator reproducibility: thoracic (ICC=0.94) and lumbar (ICC=0.60).
Mannion (2004)	Assess the intra and inter-evaluator reproducibility of the instrument to measure thoracic and lumbar curvatures.	Spinal Mouse	n=20 3 assessments were conducted by 2 evaluators on two different days	Intra-evaluator reproducibility: thoracic (ICC between 0.83 and 0.87) and lumbar (ICC between 0.87 and 0.93).
Leroux (2002)	Assess the validity of the instrument to measure thoracic and lumbar curvatures.	Video-based system (Motion Analysis Corp.)	n=97 Assessments were conducted using vídeo system and X-rays.	Validity: thoracic (b=0.84) and lumbar (b=0.86).
Norton (2002)	Assess the repeatability of the instruments to measure lumbar curvature.	Inclinometer and the eletrogoniometer (Metrecom) - tangent and the trigonometric method	n=30 2 consecutive assessments by the same evaluator (two-minute interval)	Metrecom tangent method – ICC=0.92 Metrocom trigonometric method – ICC=0.90 Inclinometer – ICC=0.92
Korovessis (2001)	Assess the validity and intra and inter-evaluator reproducibility of the instrument to measure thoracic curvature.	Debrunner’s Kyphometer	n=46 Assessments were conducted by 3 evaluators on 2 different days (two-week interval). Comparison with X-ray.	Validity: b=0.75. Intra-evaluator reproducibility : ICC=0.92 Inter-evaluator reproducibility : ICC=0.84
Ng (2001)	Assess the intra-evaluator reproducibility of the instrument to measure lumbar curvature.	Inclinometer	n=12 2 assessments were conducted by 1 evaluator (three-day interval).	Intra-evaluator reproducibility: ICC=0.95 and r=.95.
Arnold (2000)	Assess the repeatability and intra and inter-evaluator reproducibility of the instrument to measure thoracic curvature.	Flexible ruler	n=20 Assessments were conducted by 2 evaluators, one of whom repeated the assessment on 3 different days.	Repeatability: ICC=0.86 and 0.91. Intra-evaluator reproducibility : ICC=0.91 Inter-evaluator reproducibility : ICC=0.86
Leroux (2000)	Assess the validity of the instrument to measure thoracic and lumbar curvatures	Stereovideographic	n=124 Assessments were conducted with photographs and X-rays.	Validity: thoracic (ICC=0.94 and r=.89) and lumbar (ICC=0.91 and r=.84).
McGorry (2000)	Assess the validity of the instrument to measure lumbar curvature	Lordosimeter	n=6 Assessments were conducted with the lordosimeter and flexible ruler.	Validity: correlation=.95.
Kovac (1999)	Assess the validity of the instrument to measure thoracic curvature	Moiré topography	n=50 Assessment were conducted with topography and X-ray.	Validity: r=0.84.
Lundon (1998)	Assess the validity, repeatability and inter-evaluator reproducibility of the instrument to measure thoracic curvature.	Debrunner’s Kyphometer and the flexible ruler	n=26 3 assessments were conducted by 3 evaluators. Comparison with X-rays.	Validity: kyphometer (ICC=0.92). Repeatability: kyphometer (ICC between 0.89-0.99) and flexible ruler (ICC between 0.89-0.96).

The flexible ruler was the most frequently described instrument, having been tested in 16 studies, four of which evaluated the thoracic region, eight the lumbar region and four both regions. Based on the results (Table 2), the flexible ruler can be seen to be more reproducible and repeatable when used to evaluate the thoracic region by means of the angle (Greendale, Nili, Huang, Seeger, & Karlamangla, 2011; Lundon, Li, & Bibershtein, 1998; Oliveira et al., 2012; Teixeira & Carvalho, 2007) or the kyphosis index (Arnold, Beatty, Harrison, & Olszynski, 2000; Greendale et al., 2011; Macintyre, Bennett, Bonnyman, & Stratford, 2011; Perry, Smith, Straker, Coleman, & O'sullivan, 2008). On the other hand, the same instrument presented lower levels of reproducibility and repeatability for the variables thoracic 'length' and 'width' (Saad, Colombo, Ribeiro, & João, 2012). Regarding validity, the studies present moderate correlation with the radiological exams for the variable 'angle' (Greendale et al., 2011; Oliveira et al., 2012; Teixeira & Carvalho, 2007) and kyphosis index (Greendale et al., 2011). There are no results referring to the concurrent validity for the variables thoracic 'length' and 'width'.

The flexible ruler presented greater variability when used to evaluate the lumbar region, displaying excellent intra-evaluator reproducibility and repeatability in the evaluation of the angle of this region in most of the studies (Hart & Rose, 1986; Lovell, Rothstein, & Personius, 1989; Oliveira et al., 2012; Seidi, Rajabi, Ebrahimi, & Moussavi, 2009; Walker, Rothstein, Finucane, & Lamb, 1987), while there is disagreement in the literature regarding its inter-evaluator reproducibility (Letafatkar, Amirsasan, Abdolvahabi, & Hadadnezhad, 2011; Lovell et al., 1989; Oliveira et al., 2012; Seidi et al., 2009). In relation to the lordosis index, studies were only found to examine the inter-evaluator reproducibility, the results of which were moderate (Hinman, 2004; Macintyre et al., 2011). The validity of the flexible ruler presented a correlation varying from moderate to excellent (Letafatkar et al., 2011; Oliveira et al., 2012; Souza et al., 2009). Therefore, it is suggested that when using the flexible ruler to evaluate both thoracic kyphosis and lumbar lordosis, the angle calculation or index methodology should be used.

Another technique widely used in studies is photogrammetry, which was tested in eight studies, two in the thoracic region, one in the lumbar region and five in both regions. Regarding photogrammetry, it should be pointed out that data collection protocols used in studies tend to be very similar. However, regarding data analysis procedures, each of the methods found used a specific software or digital algorithm. Therefore, each proposed method should be submitted to validation procedures, which explains the large number of articles that validate photogrammetry found in this systematic review. Given this, the health care professional that decides to use any of these software or digital algorithms should ensure that all the steps in the validation procedure have been completed.

The inclinometer was tested in eight studies, one in the thoracic region, three in the lumbar region and four in both regions. When used to evaluate thoracic kyphosis, the studies demonstrated excellent levels of repeatability (Lewis & Valentine, 2010), intra (Czaprowski, Pawlowska, Gebicka, Sitarski, & Kotwicki, 2012; Mellin, 1986) and inter-evaluator reproducibility (Mellin, 1986). However, the concurrent validity of this instrument cannot be affirmed, which limits its use as a diagnostic tool. When used to evaluate lumbar lordosis, the studies demonstrated excellent levels of intra (Czaprowski et al., 2012; Mellin, 1986; Ng, Kippers, Richardson, & Parnianpour, 2001) and inter-evaluator reproducibility (Williams, Haq, & Lee, 2012; Mellin, 1986). Nevertheless, the concurrent validity of this instrument for the evaluation in the lumbar region has only been tested in the flexed position (Adams, Dolan, Marx, & Hutton, 1986) thus further investigations in other positions are necessary to permit its use in clinical practice.

Another instrument, which uses a similar mechanism to the inclinometer, is the Spinal mouse. It has been described in two studies that evaluated both the thoracic and lumbar regions. The Spinal mouse was shown to have excellent levels of intra and inter-evaluator reproducibility, with results referring to evaluations carried out in both adults (Mannion et al., 2004) and children (Kellis, Adamou, Tzilios, & Emmanouilidou, 2008). However, no study was found to demonstrate the concurrent validity of the method and only one study was found for each population, which limits the generalization of the data obtained.

The goniometer was evaluated in two articles, one study evaluated the lumbar region and the other both regions. For the thoracic kyphosis only the concurrent validity, which was found to be excellent, was presented (Gravina, Ferraro, Frizziero, Ferraro, & Masiero, 2012). For the lumbar lordosis excellent inter-evaluator reproducibility was obtained (Burdett, Brown, & Fall, 1986) without presenting adequate validity (Burdett et al., 1986; Gravina et al., 2012). The flexible electrogoniometer was evaluated in four studies, three evaluated the lumbar and one the thoracic region. Among the studies, the validity for the lumbar region presented divergent results (Campbell-Kyureghyan, Jorgensen, Burr, & Marras, 2005; Walsh & Breen, 1995), while the intra-evaluator reproducibility presented excellent levels (Norton, Hensler, & Zou, 2002; Walsh & Breen, 1995). For the thoracic region excellent reproducibility and validity were obtained (Perriman et al., 2010). Notably no studies were found that evaluated all validation aspects of either the goniometer or the electrogoniometer. Thus, their use in clinical situations and scientific research is limited as new studies that evaluate the remaining aspects of validity and in different populations are required.

Debrunner's kyphometer was developed to evaluate thoracic kyphosis. Its validity has been tested in four studies with moderate to excellent correlations (Greendale et al., 2011; Kado et al., 2006; Korovessis, Petsinis, Papazisis, & Baikousis, 2001; Lundon et al., 1998). Moreover, in relation to the aspects of reproducibility excellent results were obtained (Greendale et al., 2011; Korovessis et al., 2001; Lundon et al., 1998). It should be noted that three of the studies evaluated elderly populations and only one evaluated adolescents (Korovessis et al., 2001), therefore, more studies are necessary in adolescent populations, as well as in children, young adults and the obese.

Four studies were found that assessed surface topography, with excellent concurrent validity reported in the evaluation of kyphosis and lordosis (Fortin, Feldman, Cheriet, & Labelle, 2010; Kovac & Pecina, 1999), as well as excellent reproducibility for both curvatures (Melvin et al., 2010; Schülein, Mendoza, Malzkorn, Harms, & Skwara, 2013). However, these studies cannot be directly compared because they refer to different systems (*InSpeck* 3D Digitalizer System, Moiré Topography and Jenoptik Formetric). The presented results are not sufficient to test the validity of the system since each one only presents some validations aspects.

The archometer has been described in two studies, one evaluated only the thoracic region and the other both regions. Based on these studies the archometer was found to provide valid and reproducible measurements (Chaise et al., 2011; D'Osualdo et al., 1997). However, the model from Chaise et al. (2011) has the advantage that it can be used to evaluate the lumbar and thoracic region, although all the indices obtained in the evaluation of the lumbar region were lower than those for the thoracic region. Nevertheless, further studies are necessary to evaluate the use of the archometer in the lumbar region in different populations in order to allow its wide scale use.

Each of the other instruments identified in this systematic review (Spinal Wheel, ultrasound, fastrack video system, VICON, optoelectronic system, BodyGuard Monitor, Spine Epions System, SpineScan, photograph-based visual evaluation, lordosimeter, fiber optics system, kypholordometer and the CODA system) only found to be evaluated in one study.

This systematic review shows that most of the instruments have been submitted to validity tests in only a few or in many cases only one study. In such cases, the validity of the instruments is dependent on the quality of the methodology applied in the study. Moreover, in some studies only the internal validity of the instrument was verified, hence the validity is limited to a specific population with well controlled characteristics.

When the same instrument has been evaluated by different studies the external validity is increased, hence, there is greater possibility of using the instrument in various populations.

Furthermore, the important methodological differences found between the studies hamper any attempt to compare the data obtained, as, for example: different gold standards used, different statistical analysis techniques employed and the lack of standardization of terminology, among others. Another important aspect was the methodological quality of the validation studies, as there is a lack of studies in the literature with strong scientific evidence with regard to the validation of non-invasive instruments for evaluating the spine in the sagittal plane.

CONCLUSION

In the literature there is a wide range of non-invasive instruments for evaluating the spine in the sagittal plane, however, of the 58 studies included in this review only 12 presented strong scientific evidence. Moreover, only four instruments were evaluated with regard all the aspects of validity for thoracic kyphosis, namely photogrametry, the flexicurve, the archometer and DeBrunner's kyphometer. Similarly, for the evaluation of lumbar lordosis, five instruments were evaluated with regard all the aspects of validity, namely photogrametry, the inclinometer, the flexicurve, the archometer and the kypholordometer. Therefore, it is suggested that this instruments are adopted as first choice for conducting evaluation of the spine in the sagittal plane, since they present adequate reproducibility and concurrent validity. While the instruments that present satisfactory results in relation to the aspects of reproducibility can be used in clinical follow-up, it is necessary to note the region the instrument is capable of evaluating and whether it can be used by the same or distinct evaluators. It is particularly important to pay attention to the population for which the instrument was validated, since its use in populations with distinct characteristics may lead to inconsistent results, thus it is suggested that instruments be used only for those populations for which the aspects of validity have been evaluated.

Acknowledgments:
Nothing to Declare

Conflict of interest:
Nothing to Declare

Funding:
Nothing to Declare

REFERÊNCIAS

Adams, M. A., Dolan, P., Marx, C., & Hutton, W. C. (1986). An electronic inclinometer technique for measuring lumbar curvature. *Clinical Biomechanics*, *1*(3), 130–134. http://doi.org/10.1016/0268-0033(86)90002-1

Arnold, C. M., Beatty, B., Harrison, E. L., & Olszynski, W. (2000). The reliability of five clinical postural alignment measures for women with osteoporosis. *Physiotherapy Canada*, *52*(4), 286–294.

Barrett, E., McCreesh, K., & Lewis, J. (2014). Reliability and validity of non-radiographic methods of thoracic kyphosis measurement: a systematic review. *Manual Therapy*, *19*(1), 10–17. http://doi.org/10.1016/j.math.2013.09.003

Burdett, R. G., Brown, K. E., & Fall, M. P. (1986). Reliability and validity of four instruments for measuring lumbar spine and pelvic positions. *Physical Therapy*, *66*(5), 677–684.

Campbell-Kyureghyan, N., Jorgensen, M., Burr, D., & Marras, W. (2005). The prediction of lumbar spine geometry: method development and validation. *Clinical Biomechanics (Bristol, Avon)*, *20*(5), 455–464. http://doi.org/10.1016/j.clinbiomech.2005.01.006

Chaise, F. O., Candotti, C. T., Torre, M. L., Furlanetto, T. S., Pelinson, P. P. T., & Loss, J. F. (2011). Validation, repeatability and reproducibility of a noninvasive instrument for measuring thoracic and lumbar curvature of the spine in the sagittal plane. *Revista Brasileira De Fisioterapia (São Carlos (São Paulo, Brazil))*, *15*(6), 511–517.

Chaléat-Valayer, E., Mac-Thiong, J.-M., Paquet, J., Berthonnaud, E., Siani, F., & Roussouly, P. (2011). Sagittal spino-pelvic alignment in chronic low back pain. *European Spine Journal: Official Publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, *20 Suppl 5*, 634–640. http://doi.org/10.1007/s00586-011-1931-2

Consmüller, T., Rohlmann, A., Weinland, D., Druschel, C., Duda, G. N., & Taylor, W. R. (2012). Comparative evaluation of a novel measurement tool to assess lumbar spine posture and range of motion. *European Spine Journal: Official Publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, *21*(11), 2170–2180. http://doi.org/10.1007/s00586-012-2312-1

Czaprowski, D., Pawłowska, P., Gębicka, A., Sitarski, D., & Kotwicki, T. (2012). Intra- and interobserver repeatability and reliability of the assessment of the anterio-posterior curvatures of the spine. *Ortopedia Traumatologia Rehabilitacja*, *14*(2), 145–154. http://doi.org/10.5604/15093492.992283

D'Osualdo, F., Schierano, S., & Iannis, M. (1997). Validation of clinical measurement of kyphosis with a simple instrument, the arcometer. *Spine*, *22*(4), 408–413.

Dunk, N. M., Lalonde, J., & Callaghan, J. P. (2005). Implications for the use of postural analysis as a clinical diagnostic tool: reliability of quantifying upright standing spinal postures from photographic images. *Journal of Manipulative and Physiological Therapeutics*, *28*(6), 386–392. http://doi.org/10.1016/j.jmpt.2005.06.006

Dunleavy, K., Mariano, H., Wiater, T., & Goldberg, A. (2010). Reliability and minimal detectable change of spinal length and width measurements using the Flexicurve for usual standing posture in healthy young adults. *Journal of Back and Musculoskeletal Rehabilitation*, *23*(4), 209–214. http://doi.org/10.3233/BMR-2010-0269

Edmondston, S. J., Christensen, M. M., Keller, S., Steigen, L. B., & Barclay, L. (2012). Functional radiographic analysis of thoracic spine extension motion in asymptomatic men. *Journal of Manipulative and Physiological Therapeutics*, *35*(3), 203–208. http://doi.org/10.1016/j.jmpt.2012.01.008

Ensrud, K. E., Black, D. M., Harris, F., Ettinger, B., & Cummings, S. R. (1997). Correlates of kyphosis in older women. The Fracture Intervention Trial Research Group. *Journal of the American Geriatrics Society*, *45*(6), 682–687.

Finestone, A. S., Marcus, G., Anekstein, Y., Mirovsky, Y., & Agar, G. (2013). Assessing kyphosis with SpineScan: another attempt to reduce our dependence on radiography. *The Spine Journal: Official Journal of the North American Spine Society*, *13*(8), 926–931. http://doi.org/10.1016/j.spinee.2013.03.044

Fölsch, C., Schlögel, S., Lakemeier, S., Wolf, U., Timmesfeld, N., & Skwara, A. (2012). Test-retest reliability of 3D ultrasound measurements of the thoracic spine. *PM & R: The Journal of Injury, Function, and Rehabilitation*, *4*(5), 335–341. http://doi.org/10.1016/j.pmrj.2012.01.009

Fortin, C., Feldman, D. E., Cheriet, F., & Labelle, H. (2010). Validity of a quantitative clinical measurement tool of trunk posture in idiopathic scoliosis. *Spine*, *35*(19), E988-994. http://doi.org/10.1097/BRS.0b013e3181cd2cd2

Gravina, A. R., Ferraro, C., Frizziero, A., Ferraro, M., & Masiero, S. (2012). Goniometer evaluation of thoracic kyphosis and lumbar lordosis in subjects during growth age: a validity study. *Studies in Health Technology and Informatics*, *176*, 247–251.

Greendale, G. A., Nili, N. S., Huang, M.-H., Seeger, L., & Karlamangla, A. S. (2011). The reliability and validity of three non-radiological measures of thoracic kyphosis and their relations to the standing radiological Cobb angle. *Osteoporosis International: A Journal Established as Result of Cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA*, *22*(6), 1897–1905. http://doi.org/10.1007/s00198-010-1422-z

Hart, D. L., & Rose, S. J. (1986). Reliability of a noninvasive method for measuring the lumbar curve*. *The Journal of Orthopaedic and Sports Physical Therapy*, *8*(4), 180–184.

Hinman, M. R. (2004). Interrater reliability of flexicurve postural measures among novice users. *Journal of Back and Musculoskeletal Rehabilitation*, *17*(1), 33–36. http://doi.org/10.3233/BMR-2004-17107

Huang, M.-H., Barrett-Connor, E., Greendale, G. A., & Kado, D. M. (2006). Hyperkyphotic posture and risk of future osteoporotic fractures: the Rancho Bernardo study. *Journal of Bone and Mineral Research: The Official Journal of the American Society for Bone and Mineral Research*, *21*(3), 419–423. http://doi.org/10.1359/JBMR.051201

Imagama, S., Hasegawa, Y., Matsuyama, Y., Sakai, Y., Ito, Z., Hamajima, N., & Ishiguro, N. (2011). Influence of sagittal balance and physical ability associated with exercise on quality of life in middle-aged and elderly people. *Archives of Osteoporosis*, *6*, 13–20. http://doi.org/10.1007/s11657-011-0052-1

Ishikawa, Y., Miyakoshi, N., Kasukawa, Y., Hongo, M., & Shimada, Y. (2013). Spinal sagittal contour affecting falls: cut-off value of the lumbar spine for falls. *Gait & Posture*, *38*(2), 260–263. http://doi.org/10.1016/j.gaitpost.2012.11.024

Joint Committee for Guides in Metrology. (2012). *International Vocabulary of Metrology – Basic and general concepts and associated terms* (3.ª ed.). Joint Committee for Guides in Metrology.

Kado, D. M., Christianson, L., Palermo, L., Smith-Bindman, R., Cummings, S. R., & Greendale, G. A. (2006). Comparing a supine radiologic versus standing clinical measurement of kyphosis in older women: the Fracture Intervention Trial. *Spine*, *31*(4), 463–467. http://doi.org/10.1097/01.brs.0000200131.01313.a9

Kado, D. M., Huang, M.-H., Nguyen, C. B., Barrett-Connor, E., & Greendale, G. A. (2007). Hyperkyphotic posture and risk of injurious falls in older persons: the Rancho Bernardo Study. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *62*(6), 652–657.

Kado, D. M., Lui, L.-Y., Ensrud, K. E., Fink, H. A., Karlamangla, A. S., Cummings, S. R., & Study of Osteoporotic Fractures. (2009). Hyperkyphosis predicts mortality independent of vertebral osteoporosis in older women. *Annals of Internal Medicine*, *150*(10), 681–687.

Kellis, E., Adamou, G., Tzilios, G., & Emmanouilidou, M. (2008). Reliability of spinal range of motion in healthy boys using a skin-surface device. *Journal of Manipulative and Physiological Therapeutics*, *31*(8), 570–576. http://doi.org/10.1016/j.jmpt.2008.09.001

Korovessis, P., Petsinis, G., Papazisis, Z., & Baikousis, A. (2001). Prediction of thoracic kyphosis using the Debrunner kyphometer. *Journal of Spinal Disorders*, *14*(1), 67–72.

Kovac, V., & Pećina, M. (1999). Moiré topography in measurement of the sagittal curvatures of the spine. *Collegium Antropologicum*, *23*(1), 153–158.

Leroux, M. A., Zabjek, K., Simard, G., Badeaux, J., Coillard, C., & Rivard, C. H. (2000). A noninvasive anthropometric technique for measuring kyphosis and lordosis: an application for idiopathic scoliosis. *Spine*, *25*(13), 1689–1694.

Leroux, M. A., Zabjek, K., Simard, G., Coillard, C., & Rivard, C. H. (2002). Estimated kyphosis and lordosis changes at follow-up in patients with idiopathic scoliosis. *Journal of Pediatric Orthopedics*, *22*(1), 73–79.

Letafatkar, A., Amirsasan, R., Abdolvahabi, Z., & Hadadnezhad, M. (2011). Reliability and validity of the AutoCAD software method in lumbar lordosis measurement. *Journal of Chiropractic Medicine*, *10*(4), 240–247. http://doi.org/10.1016/j.jcm.2011.02.003

Lewis, J. S., & Valentine, R. E. (2010). Clinical measurement of the thoracic kyphosis. A study of the intra-rater reliability in subjects with and without shoulder pain. *BMC Musculoskeletal Disorders*, *11*, 39. http://doi.org/10.1186/1471-2474-11-39

Lovell, F. W., Rothstein, J. M., & Personius, W. J. (1989). Reliability of clinical measurements of lumbar lordosis taken with a flexible rule. *Physical Therapy*, *69*(2), 96–105.

Lundon, K. M., Li, A. M., & Bibershtein, S. (1998). Interrater and intrarater reliability in the measurement of kyphosis in postmenopausal women with osteoporosis. *Spine*, *23*(18), 1978–1985.

Macintyre, N. J., Bennett, L., Bonnyman, A. M., & Stratford, P. W. (2011). Optimizing reliability of digital inclinometer and flexicurve ruler measures of spine curvatures in postmenopausal women with osteoporosis of the spine: an illustration of the use of generalizability theory. *ISRN Rheumatology*, *2011*, 571698. http://doi.org/10.5402/2011/571698

Mannion, A. F., Knecht, K., Balaban, G., Dvorak, J., & Grob, D. (2004). A new skin-surface device for measuring the curvature and global and segmental ranges of motion of the spine: reliability of measurements and comparison with data reviewed from the literature. *European Spine Journal: Official Publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*, *13*(2), 122–136. http://doi.org/10.1007/s00586-003-0618-8

McGorry, R. W., & Hsiang, S. M. (2000). A method for dynamic measurement of lumbar lordosis. *Journal of Spinal Disorders*, *13*(2), 118–123.

Mellin, G. (1986). Measurement of thoracolumbar posture and mobility with a Myrin inclinometer. *Spine*, *11*(7), 759–762.

Melvin, M., Mohokum, M., Sylvia, M., Mendoza, S., Udo, W., Sitter, H., … Skwara, A. (2010). Reproducibility of rasterstereography for kyphotic and lordotic angles, trunk length, and trunk inclination: a reliability study. *Spine*, *35*(14), 1353–1358. http://doi.org/10.1097/BRS.0b013e3181cbc157

Ng, J. K., Kippers, V., Richardson, C. A., & Parnianpour, M. (2001). Range of motion and lordosis of the lumbar spine: reliability of measurement and normative values. *Spine*, *26*(1), 53–60.

Norton, B. J., Hensler, K., & Zou, D. (2002). Comparisons among noninvasive methods for measuring lumbar curvature in standing. *The Journal of Orthopaedic and Sports Physical Therapy*, *32*(8), 405–414. http://doi.org/10.2519/jospt.2002.32.8.405

Oliveira, T. S., Candotti, C. T., La Torre, M., Pelinson, P. P. T., Furlanetto, T. S., Kutchak, F. M., & Loss, J. F. (2012). Validity and reproducibility of the measurements obtained using the flexicurve instrument to evaluate the angles of thoracic and lumbar curvatures of the spine in the sagittal plane. *Rehabilitation Research and Practice*, *2012*, 186156. http://doi.org/10.1155/2012/186156

O'Sullivan, K., Clifford, A., & Hughes, L. (2010). The reliability of the CODA motion analysis system for lumbar spine analysis: a pilot study. *Physiotherapy Practice and Research*, *31*(1), 16–22. http://doi.org/10.3233/PPR-2010-31104

O'Sullivan, K., O'Sullivan, L., Campbell, A., O'Sullivan, P., & Dankaerts, W. (2012). Towards monitoring lumbo-pelvic posture in real-life situations: concurrent validity of a novel posture monitor and a traditional laboratory-based motion analysis system. *Manual Therapy*, *17*(1), 77–83. http://doi.org/10.1016/j.math.2011.09.006

Perriman, D. M., Scarvell, J. M., Hughes, A. R., Ashman, B., Lueck, C. J., & Smith, P. N. (2010). Validation of the flexible electrogoniometer for measuring thoracic kyphosis. *Spine*, *35*(14), E633-640. http://doi.org/10.1097/BRS.0b013e3181d13039

Perry, M., Smith, A., Straker, L., Coleman, J., & O'Sullivan, P. (2008). Reliability of sagittal photographic spinal posture assessment in adolescents. *Advances in Physiotherapy, 10*(2), 66–75. http://doi.org/10.1080/14038190701728251

Rajabi, R., Seidi, F., & Mohamadi, F. (2008). Which method is accurate when using the flexible ruler to measure the lumbar curvature angle? Deep pint or mid point of arch? *World Applied Sciences Journal, 4*(6), 849–852.

Ranavolo, A., Don, R., Draicchio, F., Bartolo, M., Serrao, M., Padua, L., … Sandrini, G. (2013). Modelling the spine as a deformable body: Feasibility of reconstruction using an optoelectronic system. *Applied Ergonomics, 44*(2), 192–199. http://doi.org/10.1016/j.apergo.2012.07.004

Saad, K. R., Colombo, A. S., Ribeiro, A. P., & João, S. M. A. (2012). Reliability of photogrammetry in the evaluation of the postural aspects of individuals with structural scoliosis. *Journal of Bodywork and Movement Therapies, 16*(2), 210–216. http://doi.org/10.1016/j.jbmt.2011.03.005

Schülein, S., Mendoza, S., Malzkorn, R., Harms, J., & Skwara, A. (2013). Rasterstereographic evaluation of interobserver and intraobserver reliability in postsurgical adolescent idiopathic scoliosis patients. *Journal of Spinal Disorders & Techniques, 26*(4), E143-149. http://doi.org/10.1097/BSD.0b013e318281608c

Seidi, F., Rajabi, R., Ebrahimi, T. I., Tavanai, A. R., & Moussavi, S. J. (2009). The Iranian flexible ruler reliability and validity in lumbar lordosis measurements. *World Journal of Sport Sciences, 2*(2), 95–99.

Sheeran, L., Sparkes, V., Busse, M., & van Deursen, R. (2010). Preliminary study: reliability of the spinal wheel. A novel device to measure spinal postures applied to sitting and standing. *European Spine Journal: Official Publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society, 19*(6), 995–1003. http://doi.org/10.1007/s00586-009-1241-0

Singh, D. K., Bailey, M., & Lee, R. (2010). Biplanar measurement of thoracolumbar curvature in older adults using an electromagnetic tracking device. *Archives of Physical Medicine and Rehabilitation, 91*(1), 137–142. http://doi.org/10.1016/j.apmr.2009.08.145

Souza, F. R., Ferreira, F., Narciso, F. V., Makhoul, C. M. B., Canto, R. S. T., & Barauna, M. A. (2009). Evaluation of lumbar concavity using a radiographic method and kypholordometry. *Brazilian Journal of Physical Therapy, 13*(2), 103–109. http://doi.org/10.1590/S1413-35552009005000016

Teixeira, F. A., & Carvalho, G. A. (2007). Reliability and validity of thoracic kyphosis measurements using the flexicurve method. *Brazilian Journal of Physical Therapy, 11*(3), 199–204. http://doi.org/10.1590/S1413-35552007000300005

van Niekerk, S.-M., Louw, Q., Vaughan, C., Grimmer-Somers, K., & Schreve, K. (2008). Photographic measurement of upper-body sitting posture of high school students: a reliability and validity study. *BMC Musculoskeletal Disorders, 9*, 113. http://doi.org/10.1186/1471-2474-9-113

Walker, M. L., Rothstein, J. M., Finucane, S. D., & Lamb, R. L. (1987). Relationships between lumbar lordosis, pelvic tilt, and abdominal muscle performance. *Physical Therapy, 67*(4), 512–516.

Walsh, M., & Breen, A. C. (1995). Reliability and validity of the Metrecom Skeletal Analysis System in the assessment of sagittal plane lumbar angles. *Clinical Biomechanics (Bristol, Avon), 10*(4), 222–223.

Whiting, P., Rutjes, A. W. S., Dinnes, J., Reitsma, J., Bossuyt, P. M. M., & Kleijnen, J. (2004). Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technology Assessment (Winchester, England), 8*(25), iii, 1-234.

Williams, J. M., Haq, I., & Lee, R. Y. (2010). Dynamic measurement of lumbar curvature using fibre-optic sensors. *Medical Engineering & Physics, 32*(9), 1043–1049. http://doi.org/10.1016/j.medengphy.2010.07.005

Williams, J. M., Haq, I., & Lee, R. Y. (2012). Dynamic lumbar curvature measurement in acute and chronic low back pain sufferers. *Archives of Physical Medicine and Rehabilitation, 93*(11), 2094–2099. http://doi.org/10.1016/j.apmr.2012.06.012
