DOI: 10.4025/jphyseduc.v29i1.2946

Review Article

ASSESSMENT OF THORACIC AND LUMBAR SPINE RANGE OF MOTION: SYSTEMATIC REVIEW WITH META-ANALYSIS

AVALIAÇÃO DA AMPLITUDE DE MOVIMENTO DA COLUNA TORÁCICA E LOMBAR: REVISÃO SISTEMÁTICA COM METANÁLISE

Marja Bochechin do Valle[1], Emanuelle Francine Detogni Schmit[1], Juliana Adami Sedrez[1] and **Cláudia Tarragô Candotti[1]**

[1]Universidade Federal do Rio Grande do Sul, Porto Alegre-RS, Brasil.

ABSTRACT

The spine presents distinct mobility and characteristics according to the anatomical region, and there are several instruments that allow it to be assessed. This systematic review aimed to identify methods and instruments used to assess the range of motion of the thoracic and lumbar spine in the sagittal plane, with confirmed validity and/or repeatability and/or reproducibility, evidencing their respective psychometric indexes. Searches were conducted on BIREME, EMBASE, PEDro, PubMed, Science Direct, SCOPUS and Web of Science databases, and there were manual searches as well. Two independent reviewers selected the studies, extracted data, evaluated methodological quality, risk of bias, and evidence (GRADE). A total of 46 studies were included in the qualitative analysis, seven of which only were included in the quantitative analysis. There is scientific evidence, confirmed by meta-analysis, on the inter-rater reproducibility of the measuring tape instrument in the modified Schöber's test for lumbar flexion, and the intra-rater reproducibility of the Flexicurve and video analysis system instruments for lumbar extension and flexion. Besides, based on GRADE criteria, there is still little scientific evidence on the validity, repeatability and reproducibility of the instruments and methods indicated for assessing the range of motion of the thoracic and lumbar spine in the sagittal plane.

Keywords: Range of motion. Spine. Review.

RESUMO

A coluna vertebral apresenta mobilidade e características distintas conforme a região anatômica, e, há diversos instrumentos que propiciam sua avaliação. Esta revisão sistemática objetivou identificar os métodos e instrumentos utilizados para avaliar a amplitude de movimento da coluna vertebral torácica e lombar no plano sagital que apresentam validade e/ou repetibilidade e/ou reprodutibilidade confirmados, evidenciando seus respectivos índices psicométricos. Foram realizadas buscas nas bases de dados BIREME, EMBASE, *PEDro*, *PubMed*, *Science Direct*, SCOPUS e *Web of Science*, além de buscas manuais. Dois revisores independentes realizaram a seleção dos estudos, extraíram os dados, avaliaram a qualidade metodológica, o risco de viés e a evidência (GRADE). Foram incluídos 46 estudos na análise qualitativa, e destes, apenas sete foram incluídos na análise quantitativa. Há evidência científica, confirmada por metanálise, acerca da reprodutibilidade interavaliador do instrumento fita métrica no teste de Schöber modificado para flexão lombar e da reprodutibilidade intra-avaliador dos instrumentos flexicurva e sistema de análise de vídeo para a extensão e flexão lombar. E, com base nos critérios do GRADE, ainda há baixa evidência científica sobre a validade, repetibilidade e reprodutibilidade dos instrumentos e métodos indicados para a avaliação da amplitude de movimento articular da coluna vertebral torácica e lombar no plano sagital.

Palavras-chave: Amplitude de movimento articular. Coluna vertebral. Revisão.

Introduction

The spine is a complex segment of the human body, whose mobility has different characteristics depending on the anatomical region, due to morphological differences related to the length and angle of spinal processes and to the volume of vertebral bodies[1]. Specifically, the thoracic and lumbar regions play a fundamental role in trunk movement and human locomotion; the balance between the musculoskeletal structures of the spine, by maintaining flexibility, avoids the onset of pathologies that may interfere with its autonomy and mobility[2]. In this sense, preserving the morphology and mobility of the spine is important for its functionality[3] and can reduce already high rates of back pain in the world population[4].

Still regarding the biomechanical aspects of motor and postural control related to spinal structures, evidence points to a need to maintain the integrity of the active (musculotendinous), passive (osteoarticular and ligamentous) and neural subsystems[5]. It should be pointed out, in a conceptual way, that mobility, when related to functional range of motion, is associated with joint integrity, as well as the flexibility or extensibility of soft tissues that cross or surround the joints, qualities required for unrestricted and painless body movements during functional activities of daily living[6]. Therefore, mobility and flexibility are directly related as well as, and can be understood as complementary or synonymy.

In view of the above, assessing mobility and flexibility is an important requirement in physical and clinical assessment. There is evidence pointing to video systems, that is, cinemetry, as the gold standard for range of motion (ROM) assessments. Such systems provide accurate spatiotemporal information of the body as a whole or segmented[7,8], as well as linear and/or angular information of assessed segments such as position, speed and acceleration[7,9].

However, the high cost of these systems, along with the need for ample space for assessments, besides specialized people to perform them, makes the method clinically unviable[10], leaving it restricted to the research environment. Thus, alternative methods have been described to assess the ROM of the thoracic and lumbar spine and, given the wide range offered, it is appropriate to identify what methods with scientific reliability of evidence can be used in clinical practice. Therefore, this systematic review aimed to identify methods and instruments used to assess the ROM of the thoracic and lumbar spine in the sagittal plane that have confirmed validity and/or repeatability and/or reproducibility, evidencing their respective psychometric indexes. Conceptually, validity refers to the degree of veracity of measurements of a certain quantity, that is, how much the measures approach the true value[11]. Repeatability describes the degree of equality between obtained results, based on consecutive measurements performed by the same rater, using the same instrument and method[11]. Finally, reproducibility, which can be measured intra-rater and inter-rater, describes the degree of equality between results obtained in tests conducted by the same rater or by different raters, respectively, using the same instrument and method[11].

Study Type

The present study comprised a systematic literature review, being registered in PROSPERO under the code CRD42015026518 (http://www.crd.york.ac.uk/PROSPERO_REBRANDING/display_record.asp?ID=CRD42015025996).

Search Strategies

Systematic searches were done, as recommended by the Cochran Collaboration[13], from September 25 to October 1, 2015, on the following databases: BIREME, EMBASE, Physiotherapy Evidence Database (PEDro), PubMed, Science Direct, SCOPUS and Web of Science. The search terms used, with their respective Boolean operators, were Spine [AND] Evaluation [OR] Measurement [AND] Reproducibility of Results [OR] Reliability [OR] Validity [AND] Range of Motion, Articular [OR] Range of Motion [OR] Motion [OR] Pliability [OR] Flexibility. The search strategy used on PubMed can be seen in Figure 1. In addition, there were no restrictions as to language and date of publication, and studies were identified from the references of included studies.

#5	Search (#1 AND #2 AND #3 AND #4)
#4	Search ("Spine"[Mesh] OR "Spine" OR "Vertebral Column" OR "Column, Vertebral" OR "Columns, Vertebral" OR "Vertebral Columns" OR "Spinal Column" OR "Column, Spinal" OR "Columns, Spinal" OR "Spinal Columns" OR "Vertebra" OR "Vertebrae")
#3	Search ("Procedures" OR "Procedure" OR "Evaluation Studies as Topic"[Mesh] OR "Evaluation Studies as Topic" OR "Evaluation" OR "Evaluations" OR "Evaluation Indexes" OR "Indexes, Evaluation" OR "Measurement" OR "Measurements" OR "Instruments" OR "Evaluation Methods" OR "Assess" OR "Assessment")
#2	Search ("Range of Motion, Articular"[Mesh] OR "Motion"[Mesh] OR "Motion" OR "Movement" OR "Movement"[Mesh] OR "Range of Motion, Articular" OR "Range of Motion" OR "Movements" OR "Pliability" OR "Pliability"[Mesh] OR "Flexibility")
#1	Search ("Validation Studies" [Publication Type] OR "Reproducibility of Results"[Mesh] OR "Reproducibility of Results" OR "Reproducibility of Findings" OR "Reliability" OR "Reliabilities" OR "Validity" OR "Validities" OR "Validity of Results" OR "Reliability and Validity" OR "Validity and Reliability" OR "Reliability of Results")

Figure 1. Search strategy on PubMed
Source: The authors

Study Selection

Two raters, independently, selected potentially relevant studies by reading titles and abstracts. When the latter did not provide enough information to exclude the study, the full text was verified. Afterwards, the same raters independently evaluated the full studies and made a selection according to the eligibility criteria, which were: (1) assessment of the thoracic or lumbar regions, or both; (2) assessment of flexibility/ROM/mobility; (3) assessment of a non-exclusive sample of children and patients with pathologies; (4) not being a systematic review; (5) validation or repeatability study (*measurements repeated on the same day by the same rater*)[11], or inter-rater reproducibility (measurements performed by different raters)[11] or intra-rater reproducibility (measurements performed by the same rater on different days)[11], with positive results that confirmed psychometric indexes; (6) text in Brazilian Portuguese, Spanish or English. Discrepant cases were resolved by consensus or by a third rater[14].

Data Extraction, Analysis of Quality and Risk of Bias

Only included studies were subjected to data extraction, analysis of quality and risk of bias. Information was extracted through a standardized form and included: name of the first author, year of publication, participants (total number and per group, age), assessment protocol and results of interest (Table 1). Quality and risk of bias were evaluated using the critical evaluation scale for reproducibility and validity studies[15] by the same two raters, independently. In case of disagreement, consensus was intermediated with a third rater. This scale consists of a 13-item checklist[15]. Although this scale[15] does not provide a cut-off point, in the present systematic review the studies were considered of high methodological quality when they reached scores ≥ 60% in the applied items, according to the proposition of previous studies[16].

Statistical Analysis

Data were initially analyzed by means of descriptive statistics, separated into subgroups according to instrument and assessed movement, as well as to used methodology used and type of analysis (validity, repeatability and reproducibility – intra- or inter-rater; statistical test conducted). Meta-analysis was carried out on the Medal software, version 11.0 (MedCalc Software, Mariakerke, Belgium), based on sampling size (total n of the study) and correlation (r value) information, by means of inferential statistics with Higgins's Inconsistency test (I^2) to verify inter-study homogeneity, considering low heterogeneity if I^2<50%, and moderate/high if I^2≥50%[13].

Quality of Evidence

In order to summarize the quality of the evidence, the GRADE (Grading of Recommendations Assessment, Development, and Evaluation)[17] system was used, which takes into account the following criteria: design and methodological limitations of included studies; inconsistency (homogeneity of studies); whether the studies present direct evidence; accuracy of results presented in included studies; and whether the systematic review presents a publication bias, not including the totality of published studies about the research problem. Based on these criteria, the pieces of evidence were classified into the four levels presented by the GRADE system: high quality – it is very unlikely that additional research will change the results presented by the systematic review; moderate quality – further research is likely to have a major impact and may change the results presented by the systematic review; low quality – it is more likely that further research will have a significant impact and change the results presented by the systematic review; and very low quality – any estimation of results presented by the systematic review is rather uncertain, generating the need to develop new studies.

Results

A total of 4,027 studies were initially identified from the systematic searches, of which 1,682 were duplicates and 2,257 were excluded after the reading of titles and abstracts, leaving 88 for detailed analysis. Based on the eligibility criteria, 42 studies were excluded, leaving 46 articles for qualitative analysis. Figure 2 shows the flowchart of included studies, and Table 1 summarizes the characteristics of these studies.

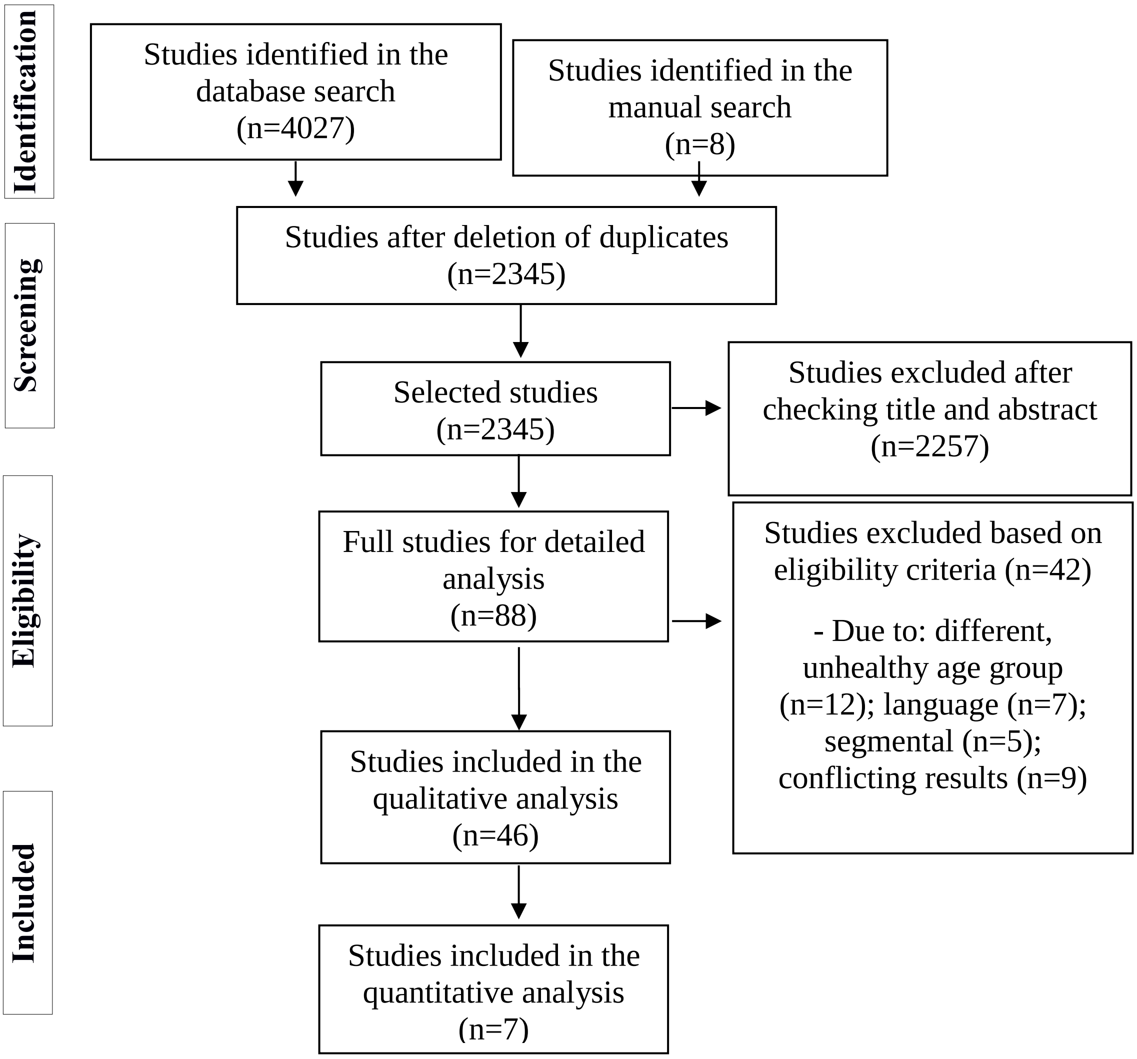


Figure 2. Flowchart of included studies according to PRISMA[18]
Source: The authors

Table 1. Characteristics of included studies

1st Author	Sample	Instrument	Assessed Aspect	Results
Measuring tape				
Bandy[19]	n=63	Sternum-bed distance	Intra-rater reproducibility of lumbar extension (experienced and unexperienced raters)	Experienced ICC: 0.90-0.91; unexperienced ICC: 0.82-0.86
Beattie[20]	n=100	Modified Schöber's test	Intra- and inter-rater reproducibility (n=11) of lumbar extension	Intra-rater ICC: 0.90; Inter-rater ICC: 0.94
Burdett[21]	n=23	Modified Schöber's test	Intra-rater reproducibility of lumbar flexion.	ICC:0.72
Dopf [22]	n=30	Modified Moll's and modified Schöber's tests	Intra- and inter-rater reproducibility of lumbar flexion and extension.	Intra-rater reproducibility: Flexion: *r*: 0.89; Extension: *r*: 0.66; Inter-rater reproducibility: Flexion: *r*: 0.76; Extension: *r*: 0.54.
Frost[23]	n=24	Finger-floor distance and C7-S2	Intra- and inter-rater repeatability and reproducibility of trunk flexion and extension	Repeatability: Flexion: *r*: 0.98; Extension: r: 0.96; Intra-rater reproducibility: Flexion: *r*: 0.98; Extension: r: 0.79; Inter-rater reproducibility: Flexion: *r*: 0.94; Extension: *r*: 0.78
Gill[24]	n=10	Modified Schöber's test and finger-floor distance	Repeatability of lumbar flexion and extension	Flexion: Modified Schöber: CV: 0.9-1.5; Finger-floor: CV: 14.1; Extension: Modified Schöber : CV: 2.8-2.9
Merritt[25]	n=50	Modified Schöber's, Moll's, Loebl's tests, and finger-floor distance	Intra- and inter-rater reproducibility of trunk flexion and extension	Intra-rater reproducibility: Flexion: Finger-floor distance: mean CV: 76.4; Schöber: mean CV: 6.6; Loebl: mean CV: 13.4; Extension: Moll: mean CV: 7.3; Loebl: mean CV: 50.7; Inter-rater reproducibility: Flexion: Finger-floor distance: mean CV: 83.0; Schöber: mean CV: 6.3; Loebl: mean CV: 9.6; Extension: Moll: mean CV: 9.5; Loebl: mean CV: 65.4
Ronchi[26]	n=23	Modified Schöber's test	Intra- and inter-rater reproducibility of lumbar flexion	Intra-rater reproducibility: ICC: 0.77; Inter-rater reproducibility: ICC: 0.74
Miller[27]	n=50	Modified Schöber's test	Inter-rater reproducibility of lumbar flexion	*r*:0.71

Table 1. Characteristics of included studies (continued...)

1st Author	Sample	Instrument	Assessed Aspect	Results
Paternostro-Sluga[28]	n=16	Ott's, Schöber's, Modified Schöber's test, and finger-floor distance	Intra- and inter-rater reproducibility of trunk flexion	Good reproducibility (ICC values not specified)
Hyytiäinen[29]	n=30	Schöber's test	Intra- and inter-rater reproducibility of lumbar flexion	Intra-rater reproducibility: *r*: 0.88; inter-rater: *r*: 0.87
Van Den Dolder[30]	n=60	Author's own methodology	Intra- and inter-rater reproducibility of lumbar flexion	Intra-rater reproducibility: ICC: 0.95; Inter-rater reproducibility: ICC: 0.96
Inclinometer				
Bø[31]	n=16	Digital inclinometer	Intra- and inter-rater reproducibility of thoracolumbar flexion and extension	Flexion: intra-rater ICC: 0.84-0.92; Inter-rater ICC: 0.83-0.92. Extension: intra-rater ICC: 0.85-0.86; Inter-rater ICC: 0.68-0.88
Breum[32]	n=47	Modified inclinometer (BROM II)	Intra-, inter-rater reproducibility and validity of lumbar flexion and extension (dual inclinometer)	Intra-rater reproducibility: Flexion: ICC: 0.91; Extension: ICC: 0.63; Inter-rater reproducibility: Flexion: ICC: 0.77; Extension: ICC: 0.35; Validity: Flexion: ICC: 0.75; Extension: ICC: 0.63
Dopf[22]	n=30	Dual inclinometer	Intra- and inter-rater reproducibility of lumbar flexion and extension.	Intra-rater reproducibility: Flexion: *r*: 0.92; Extension: *r*: 0.93; Inter-rater reproducibility: Flexion: *r*: 0.71; Extension: *r*: 0.78
Gill[24]	n=10	Dual inclinometer	Repeatability of lumbar flexion and extension	Flexion: CV: 9.3-33.9; Extension: CV: 2.8-4.7
Kolber[33]	n=30	Inclinometer and mobile device (inclinometer - iPhone)	Intra-, inter-rater reproducibility and validity (inclinometer) of lumbar and trunk flexion and extension.	Intra-rater reproducibility: Flexion: iPhone: lumbar: ICC: 0.88; thoracolumbopelvic: ICC: 0.97; Inclinometer: lumbar: ICC: 0.83; thoracolumbopelvic: ICC: 0.96; Extension (thoracolumbopelvic only): iPhone: only: 0,80; Inclinometer: ICC: 0.88; Inter-rater reproducibility: Flexion: iPhone: lumbar: only: 0.88; thoracolumbopelvic: ICC: 0.98; Inclinometer: lumbar: ICC: 0,81; thoracolumbopelvic: ICC: 0.97; Extension (thoracolumbopelvic only): iPhone: ICC: 0,81; Inclinometer: ICC: 0.91; Validity: Flexion: lumbar: ICC: 0.86-0.87; thoracolumbopelvic: ICC: 0.97-0.98; Extension (thoracolumbopelvic only): ICC: 0.89-0.91

Table 1. Characteristics of included studies (continued...)

1st Author	Sample	Instrument	Assessed Aspect	Results
Mayer[34]	n=18	Inclinometer and electroinclinometer	Intra- and inter-rater repeatability and reproducibility of lumbar flexion	All instruments presented repeatability (*r* 0.89) and intra- (F=1,39, df=13.319) and inter-rater (F=1.62, df=1.319) reproducibility
Mellin[35]	n=27	Inclinometer	Intra-rater reproducibility of thoracolumbar flexion and extension.	Flexion: *r*: 0.91-0.95; Extension: *r*: 0.72-0.92
Ng[36]	n=12	Modified inclinometer	Intra-rater reproducibility of lumbar flexion and extension	Flexion: ICC: 0.87; Extension: ICC: 0.92
Ronchi[26]	n=23	Dual inclinometer	Intra and inter-rater reproducibility of lumbar flexion and extension	Intra-rater reproducibility: Flexion: ICC: 0.95; Extension: ICC: 0.94; Inter-rater reproducibility: Flexion: ICC: 0.89; Extension: ICC: 0.91
Chiarello[37]	n=12	Electroinclinometer	Inter-rater reproducibility of lumbar flexion and extension	Inter-rater reproducibility: Flexion: ICC: 0.74; Extension: ICC: 0.65-0.85
Rondinelli[38]	n=8	An inclinometer, dual inclinometer and electroinclinometer (Back ROM)	Intra- and inter-rater reproducibility of lumbar flexion	Intra-rater reproducibility: An inclinometer: ICC: 0.85-0.86; Dual inclinometer: ICC: 0.70-0.81; Back ROM: ICC I: 0.81-0.90; Inter-rater reproducibility: An inclinometer: ICC: 0.76; Dual inclinometer: ICC: 0.69; Back ROM: ICC: 0.77
Boocock[39]	n=12	Inclinometer	Intra-rater reproducibility of lumbar ROM	Intra-rater reproducibility: *r*: 0.96
Goniometer				
Bedekar[40]	n=30	iPod Mobile Device (goniometer)	Intra- and inter-rater reproducibility of lumbar flexion, concurrent validity (dual inclinometer)	Intra-rater: ICC: 0.92; Inter-rater: ICC: 0.81; Validity: *r*: 0.95
Chiarello[37]	n=12	Two goniometers	Inter-rater reproducibility of lumbar flexion and extension	Inter-rater reproducibility: Flexion: ICC: 0.57; Extension: CCI: 0.59-0.67
Burdett[21]	n=27	Modified Gravity Goniometers and Parallelogram	Intra-rater reproducibility of lumbar flexion and extension. Validity of lumbar flexion and extension	Intra-rater reproducibility: Flexion: Gravity Goniometer: ICC: 0.91; Parallelogram Goniometer: ICC: 0.92; Extension: Gravity Goniometer: ICC: 0.71; Parallelogram Goniometer: ICC: 0.60; Validity: Flexion: Gravity Goniometer: ICC: -0.11; Parallelogram Goniometer: ICC: 0.19; Extension: Gravity Goniometer: ICC: -0.73; Parallelogram Goniometer: ICC: -0.71

Table 1. Characteristics of included studies (continued...)

1st Author	Sample	Instrument	Assessed Aspect	Results
Salisbury[41]	n=17	Goniometer	Intra-rater reproducibility of lumbar flexion and extension	Flexion: MAD: 3.80±2.95; Extension: MAD: 3.10±1.98
Salisbury[41]	n=17	Goniometer	Intra-rater reproducibility of lumbar flexion and extension	Flexion: MAD: 3.80±2.95; Extension: MAD: 3.10±1.98
Boocock[39]	n=12	Electrogoniometer	Intra-rater reproducibility of lumbar ROM	Intra-rater reproducibility: *r*: 0.78.
Paquet[42]	n=10	Electrogoniometer	Repeatability and validity (two inclinometers) of trunk flexion	Validity: *r*: 0.97; Repeatability: ICC: 0.98
Tojima[43]	n=7	Electrogoniometer	Intra-rater reproducibility of lumbar flexion and extension	Flexion: ICC: 0.80; extension: ICC: 0.63
Motion Analysis System				
Gill[44]	n=15	Video Motion Analysis System	Intra- and inter-rater reproducibility (10 individuals) of trunk flexion and extension.	Intra-rater reproducibility: Flexion: *r*: 0.87; Extension: *r*: 0.85; Inter-rater reproducibility: Flexion: *r*: 0.93; Extension: *r*: 0.96
Mannion[45]	n=11	3D motion analysis system OSI CA-6000 and Space Fastrak	Repeatability of lumbar flexion and extension	Repeatability: *r*: 0.82-0.99, with high ICC (values not specified)
Petersen[46]	n=21	3D motion analysis system (OSI CA-6000)	Intra and inter-rater reproducibility (raters with and without experience) of thoracolumbar flexion and extension	Intra-rater reproducibility: Flexion: ICC: 0.90-0.96; Extension: ICC: 0.96; Inter-rater reproducibility: Flexion: ICC: 0.93; Extension: ICC: 0.95
Pearcy[47]	n=10	3D motion analysis system (3 SPACE Isotrak)	Repeatability of lumbar flexion and extension	RMS error: 0.079
Dopf [22]	n=30	3D motion analysis system (OSI CA-6000)	Intra- and inter-rater reproducibility of lumbar flexion and extension.	Intra-rater reproducibility: Flexion: r*:* 0.94; Extension: *r*: 0.94; Inter-rater reproducibility: Flexion: *r*: 0.76; Extension: *r:* 0.84.
Tojima[43]	n=7	3D motion analysis system (VICON)	Intra-rater reproducibility of lumbar flexion and extension	Flexion: ICC: 0.77; extension: ICC: 0.80.
Troke[48]	n=22	3D motion analysis system (OSI CA-6000)	Intra- and inter-rater reproducibility of lumbar flexion and extension	Intra-rater reproducibility: ICC: 0.81-0.94; Inter-rater reproducibility: ICC: 0.73-0.82

Table 1. Characteristics of included studies (continued...)

1st Author	Sample	Instrument	Assessed Aspect	Results
Schuit[49]	n=10	3D Motion Analysis System (OSI CA-6000) and X-Ray	Inter-rater reproducibility and validity of trunk flexion and extension	Inter-rater reproducibility: Flexion: X-ray: ICC: 0.93; OSI: CCI: 0.99; Extension: X-ray: ICC: 0.85; OSI: ICC: 0.98; Validity: Flexion: *r*: 0.100; Extension: *r*: 0.394
Schuit[49]	n=10	3D Motion Analysis System (OSI CA-6000) and X-Ray	Inter-rater reproducibility and validity of trunk flexion and extension	Inter-rater reproducibility: Flexion: X-ray: ICC: 0.93; OSI: ICC: 0.99; Extension: X-ray: ICC: 0.85; OSI: ICC: 0.98; Validity: Flexion: *r*: 0.100; Extension: r: 0.394
Flexicurve				
Tillotson[50]	n=20	Flexicurve	Intra-rater reproducibility and concurrent validity (X-ray) of lumbar flexion and extension.	Intra-rater reproducibility: Flexion: *r*: 0.95-0.97; Extension: *r*: 0.96-0.97; Validity: ROM: *r*: 0.98.
Burton[51]	n=15	Flexicurve	Intra- and inter-rater reproducibility of lumbar flexion and extension, validity (X-rays, n=1)	Intra-rater reproducibility: *r*: 0.95-0.97; Inter-rater reproducibility: *r*: 0.82-0.99; validity: the flexicurve presented superior angulation (greater by 1º) to X-ray
Burton[52]	Not presented	Flexicurve	Intra and inter-rater reproducibility of lumbar flexion and extension	Intra- (9% error) and inter-rater (7-15%) reproducibility
Youdas[53]	n=10	Flexicurve	Intra- and inter-rater reproducibility of lumbar flexion and extension	Intra-rater reproducibility: Flexion: ICC: 0.90-0.95; Extension: ICC: 0.96-0.98; Inter-rater reproducibility: Flexion: ICC: 0.84-0.91; Extension: ICC: 0.97-0.98
Boocock[39]	n=12	Flexicurve	Intra-rater reproducibility of lumbar ROM	Intra-rater reproducibility: *r*: 0.86
Accelerometers				
Alqhtani[54]	n=18	Triaxial accelerometer	Reproducibility of thoracolumbar flexion and extension.	Thoracic: flexion (ICC: 0.97-0.99) and extension (ICC: 0.92-0.96); Lumbar: flexion (ICC: 0.95-0.98) and extension (ICC: 0.96-0.97)
Consmuller[55]	n=30	Accelerometer	Intra-rater reproducibility of thoracolumbar flexion and extension.	Intra-rater reproducibility: Flexion: ICC: 0.86; Extension: ICC: 0,84
Ronchi[26]	n=23	Accelerometer	Intra and inter-rater reproducibility of lumbar flexion and extension	Intra-rater reproducibility: Flexion: ICC: 0.99; Extension: ICC: 0.98; Inter-rater reproducibility: Flexion: ICC: 0.95; Extension: ICC: 0.95

Table 1. Characteristics of included studies (continued...)

1st Author	Sample	Instrument	Assessed Aspect	Results
Photogrammetry				
Tederko[56]	n=12	Photometry	Repeatability and reproducibility of thoracic ROM	ICCs between 0.951 and 0.958 (no expressed isolated values per movement and assessed aspect)
Gill[24]	n=10	Photogrammetry	Repeatability of lumbar flexion and extension	Flexion: CV: 6.0-22.3; Extension: CV; 11.3-12.4
Edmondston[57]	n=14	Photogrammetry	Validity of thoracic extension ROM (X-ray).	*r*:0.69
Inertial System				
Ha[58]	n=26	Inertial System (Xsems MTx)	Validity (with Fastrak) of lumbar flexion and extension	Flexion: *r*: 0.88; Extension: *r*: 0.66
Yun[59]	n=19	Inertial system	Intra-rater reproducibility of lumbar flexion and extension	Intra-rater reproducibility: ICC: 0.90-0.98
Other Instruments				
Roussel[60]	n=61	Isokinetic dynamometer	Inter-rater reproducibility of lumbar flexion and extension	Inter-rater reproducibility: Flexion: ICC: 0.77; Extension: ICC: 0.93-0.94;
Williams[61]	n=13	Fiber Optic System	Repeatability and validity (3D motion analysis system) of lumbar flexion	Repeatability: *r*: 0.94-0.97; Validity: *r*: 0.86-0.95
Lee[62]	n=19	3D Gyroscope	Repeatability of lumbar flexion and extension	Multiple correlation coefficient: 0.97-0.99
Salisbury[41]	n=17	<u>Kyphometer</u>, Goniometer and Flexicurve, measuring tape and ultrasound	Intra-rater reproducibility of lumbar flexion and extension	Flexion: kyphometer MAD: 2.95 ± 2.96; MAD Goniometer: 3.80 ± 2.95 and Flexicurve MAD: 3.15±2.0. Extension: kyphometer MAD: 3.16±2.24; Goniometer MAD: 3.10±1.98 and Flexicurve MAD: 4.18±3.58
Cohn[63]	n=19	Electromagnetic Sensors	Intra and inter-rater reproducibility of lumbar flexion and extension	Intra- and inter-rater reproducibility with ICC> 0.9.
Fölsch[64]	n=28	Ultrasonic analysis system	Intra-rater reproducibility of thoracic flexion and extension.	Flexion: ICC: 0.71; Extension: ICC: 0.34

Source: The authors

In the evaluation of methodological quality and risk of bias only 18 studies were considered of high quality (score ≥ 60%). The mean of the studies' methodological quality was 53.11% and can be seen in Table 2.

Table 2. Methodological quality evaluation

1st Author	1	2	3	4	5	6	7	8	9	10	11	12	13	Quality	Quality %
Alqhtani[54]	y	n	n/a	n	n	y	n/a	y	n/a	y	n/a	n	y	5	55.56
Bandy[19]	y	y	n/a	y	s	y	n/a	y	n/a	y	n/a	n	y	8	88.9
Beattie[20]	y	n	n/a	n	n	y	n/a	y	n/a	y	n/a	n	y	5	55.6
Bedekar[40]	y	y	n/a	y	s	y	n/a	y	n/a	y	n/a	n	y	8	88.9
Bø[31]	y	y	n/a	y	s	y	n/a	y	n/a	y	n/a	n	y	8	88.9
Boocock[39]	y	n	y	n	n	y	y	y	y	y	y	n	y	9	69.2
Breum[32]	y	n	y	n	n	y	y	y	n	y	y	n	y	8	61.5
Burdett[21]	y	y	y	n	n	n	n	n	n	y	y	n	y	6	46.2
Burton[51]	n	n	n	n	n	n	y	y	y	n	n	n	y	4	44.4
Burton[52]	n	n	n	n	n	n	n	n	n	n	n	n	n	0	0.0
Chiarello[37]	y	n	n/a	n	n	y	n/a	n	n/a	y	n/a	n	y	4	44.4
Cohn[63]	y	y	n/a	n	n	y	n/a	y	n/a	y	n/a	n	y	6	66.7
Consmuller[55]	y	n	n/a	n/a	n	n	n/a	y	n/a	y	n/a	n	y	4	50.0
Dopf[22]	y	n	y	n	n	y	y	y	n	y	y	n	y	8	61.5
Edmondston[57]	y	y	y	n	n	n	n	n	y	y	y	n	n	6	46.15
Fölsch[64]	y	n	n/a	n	n	n	n/a	n	n/a	y	n/a	y	y	4	44.4
Frost[23]	y	n	n/a	n	n	y	n/a	y	n/a	y	n/a	y	y	6	66.7
Gill[24]	y	n	n/a	n	n	n	n/a	y	n/a	n	n/a	y	y	4	44.4
Gill[44]	y	n	n/a	n	n	y	n/a	y	n/a	y	n/a	n	y	5	55.6
Ha[58]	y	n	y	n/a	n/a	n	n	n	y	y	y	n	y	6	66.7
Hyytiainen[29]	y	y	n/a	n	n	y	n/a	y	n/a	y	n/a	n	n	5	55.6
Kolber[33]	y	n	y	s	s	n	y	y	y	y	y	n	y	10	76.9
Lee[62]	y	n	n/a	n	n	n	n/a	n	n/a	y	n/a	n	n	2	22.2
Mannion[46]	y	n	n/a	n	n	y	n/a	n	n/a	y	n/a	n	y	4	30.8
Mayer[34]	y	y	n/a	n	n	y	n/a	n	n/a	y	n/a	n	n	4	44.4
Mellin[25]	y	n	n/a	n	n	n	n/a	n	n/a	y	n/a	y	n	4	44.4
Merritt[36]	y	y	n/a	n	n	n	n/a	y	n/a	y	n/a	n	n	4	44.4
Miller[27]	y	y	n/a	y	n/a	y	n/a	n	n/a	y	n/a	n	n	5	62.5
Ng[36]	y	y	n/a	n/a	n	n	n/a	n	n/a	y	n/a	n	y	4	50.0
Paquet[42]	y	n	y	n	n	n	n	y	n	y	y	n	y	6	46.2
Paternostro-Sluga[28]	y	n	n/a	y	s	y	n/a	y	n/a	y	n/a	n	n	6	66.7
Pearcy[47]	y	n	n	n	n	n	n	y	n	n	n	n	n	2	15.4
Petersen[46]	y	y	n/a	n	n	n	n/a	y	n/a	y	n/a	n	y	5	38.8
Ronchi[26]	y	y	n/a	n	n	n	n/a	y	n/a	y	n/a	n	y	5	38.8
Rondinelli[40]	y	y	n/a	n	n/a	y	n/a	y	n/a	y	n/a	n	y	6	46.2
Roussel[60]	y	n	n/a	n	n	n	n/a	n	n/a	y	n/a	n	y	3	33.3
Salisbury[41]	y	n	n/a	n	n	y	n/a	n	n/a	y	n/a	n	n	3	33.3
Schuit[49]	y	n	y	n	n	y	y	n	y	y	y	n	y	8	61.5
Tederko[56]	y	n	n/a	n	n	n	n/a	n	n/a	y	n/a	n	y	3	33.3
Tillotson[50]	y	y	y	n	n	n	n	y	y	y	y	n	y	8	61.5
Tojima[43]	y	n	y	n	n	y	y	y	y	y	y	n	y	9	69.2
Troke[48]	y	y	n/a	n	n	y	n/a	y	n/a	y	n/a	n	y	6	66.7
Van DenDolder[30]	y	y	n/a	y	s	y	n/a	n	n/a	y	n/a	n	y	7	77.8
Williams[61]	y	y	y	n/a	n/a	n/a	n	n/a	y	y	y	n	y	7	77.8
Youdas[53]	y	y	n/a	n	n	n	n/a	n	n/a	y	n/a	n	y	4	44.4
Yun[59]	y	y	n/a	n	n	n	n/a	y	n/a	y	n/a	n	y	5	55.6

Note: 1- Sample adequacy; 2- rater description adequacy; 3- explanation of reference standard; 4- Inter-rater blinding; 5- Intra-rater blinding; 6- Evaluation order variation; 7- Period of time between evaluated test and reference standard; 8- Period between repeated measures; 9- Independence of reference standard from evaluated test; 10- Adequacy of the evaluated test procedure's description; 11- Adequacy of the description of the reference standard's procedure; 12- Sampling loss Explanation; 13- Appropriate statistical methods.
Source: The authors

The main areas of methodological weakness found were: explanation about sampling loss, justified for being cross-sectional studies; intra- and inter-rater blinding; period of time between evaluated test and reference standard; independence of reference standard from evaluated test; explanation and adequacy of the description of the reference standard's procedure, and rater description adequacy.

With regard to quality of evidence, taking into account the heterogeneity of studies, especially concerning the methodological rigor, it is possible that other researches have an important impact and probably change the results presented by the present systematic review, which gives the present review low strength of evidence based on the main criteria established by GRADE[17].

Regarding quantitative analysis results, only seven studies were included in the meta-analysis, supporting that there is scientific evidence on the inter-rater reproducibility of the measuring tape instrument in the modified Schöber's test for lumbar flexion movement, and the intra-rater reproducibility of the Flexicurve and video analysis system instruments for lumbar extension and flexion movements (APPENDIX).

Discussion

The studies presented in Table 1 show the use of numerous instruments to assess spinal flexibility, of which the most commonly employed are: measuring tape, inclinometers, goniometers/electrogoniometers, 3D motion analysis systems, Flexicurves and accelerometers. In addition, some instruments were mentioned in a few studies, such as: photogrammetry, ultrasound, inertial system, optical fiber system, kyphometer, electromagnetic sensors, 3D gyroscope, and isokinetic dynamometer. Besides the variety of instruments, the protocols used are numerous for each one of them, making it even more difficult to compare the studies.

Measuring tape is an instrument that has been frequently described in studies for assessment of flexion and extension ROM of the thoracic and lumbar spine, with several measurement protocols, such as the modified Schöber's test[20-22,24-29], finger-floor distance[24,25,28], modified Moll's test[22,25], among others. It should be noted that low cost, easy handling and the fact of providing quantitative results, presenting values in centimeters (cm), are factors that can facilitate the widespread use of this instrument. Furthermore, measurement protocols, in general, have adequate intra- and inter-rater repeatability and reproducibility (Table 1), which makes it possible to use them to follow up spinal training and treatments, since measurements can be reliably performed at different times, as well as by different raters. In addition, it is possible to affirm, through meta-analysis, that the inter-rater reproducibility of the modified Schöber's test for assessment of lumbar flexion (APPENDIX) is already elucidated and confirmed, that is, it is very likely that the results from the test are similar, though carried out by different raters. However, it should be pointed out that, when it comes to questions related to statistical analysis applied in the studies[22,23,27,29], there is discrete misconception when using only tests that verify correlation (values correlate, that is, behave in a direct way – one increases, the other increases, one decreases, the other decreases – or inversely – one increases, the other decreases, and vice versa, in the same proportion, but they are not necessarily similar or close; in this case, the relevant statistical tests are Pearson's and Spearman's) and not agreement (when the difference between one value and another is null or very close to that, the values are identical or nearly identical; in this case, the relevant statistical test is the Intraclass Correlation Coefficient).

However, there were no studies that assessed the concurrent validity of the protocols (internal comparisons between different measurement methodologies, taking into account the agreement between them) and used measuring tape; thus, the fact of not knowing the real variable analyzed in these protocols, that is, whether it is really about assessment of spinal flexibility or whether other factors may be influencing the results obtained, is understood as an important limitation. Another limitation that can be emphasized is the fact that assessment is usually done based on only two reference points, which does not allow representing the curvature of the spine.

Another widely used instrument is the inclinometer (Table 1), which consists of a gravity-driven 360° protractor. It has variations; for instance, the dual inclinometer, the modified inclinometer (BROM II) and the electroinclinometer (Back ROM). Among the included studies, only two assessed the validity of inclinometers. One of them compared a new inclinometer model, called BROM II, with a dual inclinometer and found adequate concurrent validity, with correlation varying from moderate to excellent[42]. The second study verified the concurrent validity of a mobile device inclinometer (iPhone) compared to a traditional inclinometer and found excellent correlation results between both instruments. Regarding reproducibility, in general, all studies showed moderate to high correlation for the inclinometer, and most of them performed adequate statistical analysis to measure the psychometric index[26,31-33,36-38]. However, when we take into account the meta-analysis results (APPENDIX), a high methodological heterogeneity between studies is evidenced, which prevents assertions and extrapolations about metric measurements of intra-rater reproducibility, requiring new studies, in addition to the fact that analysis was only possible with the use of Pearson's *r*-related information. A good reason to use dual inclinometers is that they are recommended by the American Medical Association (AMA), in the Guides to the Evaluation of Permanent Impaiment[40,65]. However, inclinometers presented concurrent validity tested only with the same instrument, which limits knowledge about the true value obtained, besides being relatively expensive, difficult to handle and may contain marking and assembling errors; therefore, it is necessary to master this technique in order to obtain precise measures[65].

The goniometer is an instrument that measures joint positions and ROM of almost all joints. Similar to protractors, they are transparent plastic tools used to measure or construct angles. They differ from inclinometers for not depending on the action of gravity. Variations found for the goniometer are the electrogoniometer[39,42,43], the mobile device goniometer (iPod)[40] and traditional goniometers[21,37,41].

The goniometer instrument was analyzed in several studies; however, because of this wide variety of types for this instrument, analyzing them together has its limits. In general, the goniometer has been described for lumbar mobility assessment. Only in the study by Paquet *et al.*[42], this instrument was used for trunk assessment. The electrogoniometer had its intra-rater reproducibility tested in two studies, which showed excellent results[39,43], and excellent concurrent validity when compared to the inclinometer[39,42]. Another type of goniometer that presented excellent results was the mobile device goniometer (iPod), with correlations above 0.8 for both intra- and inter-rater reproducibility and concurrent validity[40], supporting the agreement between the measures taken by different raters and at different times. Gravity goniometers and the parallelogram also presented excellent intra-rater reproducibility[21]. However, when two goniometers were used to assess lumbar flexion and extension, ICC results were lower, with moderate correlations[37].

The goniometer is considered a low-cost instrument, easy to use and carry; however, it is emphasized that goniometers require technical knowledge from raters, since their difficulty of alignment with body regions, especially in flexion and extension, may interfere with the precision of results[65]. Paquet *et al.*[42] pointed out some important limitations to the use of electrogoniometers, such as assessment only in the sagittal plane and the need for system calibration for each individual. In addition, both instruments do not allow representing the curvature of the assessed spine.

3D motion analysis systems allow determining the position and orientation of body segments, seeking to measure parameters of linear or angular displacements, speed and acceleration in these segments[8]. Among studies that assessed ROM in flexion and extension, eight were conducted with 3D motion analysis system. Of these, six verified reproducibility, with results varying from moderate to high[22,41,44,46,48-49], and it is worth highlighting that there is evidence confirmed by meta-analysis to support intra-rater reproducibility for lumbar flexion and extension movements (the statistical matter of exclusive use of tests for verification of correlation of measurements is reiterated, without information on agreement). However, inter-rater reproducibility still needs to be investigated with greater methodological rigor in order to fill in the gap caused by the heterogeneity of results between studies (APPENDIX); two studies verified the reproducibility of lumbar flexion and extension, obtaining high ICC values[41] and low RMS error between measurements[42], and only one study verified the validity of the 3D video system with X-ray examination[49], with this methodology being considered of very low validity for flexion, and low validity for trunk extension.

With the advent of technology, 3D analysis methods have expanded rapidly, mainly because they provide many possibilities of assessed parameters and present adequate precision in the results provided[66]. Nevertheless, these instruments need proper environment for assessments, experienced raters and high cost, being unfeasible for use in clinical practice.

The Flexicurve instrument is a flexible lead ruler, 30 to 80 cm long, easy to use, low-cost, and serves as a diagnostic means and evolutionary treatment indicator for field studies in large populations[67]. The concurrent validity of Flexicurve in assessing flexibility with X-rays was tested in two studies[50,51]. However, Burton's study[51] assessed only one individual, and its results only showed superior angulation (greater by one degree) when compared to X-rays. Tillotson and Burton[50], in their turn, assessing the validity in lumbar flexion and extension of Flexicurve, obtained excellent results for both movements.

The other studies[29,50-53] presented results on the reproducibility of Flexicurve, showing correlations ranging from high to very high. However, to date, there is only evidence, based on meta-analysis, to affirm the intra-rater reproducibility of lumbar spine flexion and extension movements (APPENDIX), supported also by the high agreement between measurements by the same rater in the study by Youdas *et al*[53]. Flexicurve has been described as an easy-to-assess instrument and has the advantage of providing a graphical representation of assessed curvatures. However, despite adequate intra-rater and inter-rater reproducibility and validity results, this instrument is only described for lumbar region assessment, restricting its possibility of use, since it has not been tested in thoracic spine assessment.

Accelerometers are devices that measure acceleration and are generally used in positioning systems, inclination sensors, and vibration sensors. Studies such as those by Alqhtani *et al*.[54], Consmuller *et al*.[55] and Ronchi *et al*.[26] used accelerometers to assess spinal ROM. All of them assessed intra and inter-rater reproducibility, obtaining very high results of agreement between measurements.

Other instruments such as photogrammetry[57], optical fiber system[61] and electromagnetic device (3 Space Isotrack System)[47] have been described in few studies for spinal ROM assessment, and still lack further information on their validity aspects.

The above clearly show the wide variety of instruments available to assess spinal ROM, but, mostly, the instruments present very well-defined results only for the reproducibility of the systems, as in the case of measuring tape, inclinometers, goniometers and accelerometers. Regarding the concurrent validity of the instruments, the studies that tested it presented limitations; for instance, the concurrent validity of some inclinometers and goniometers were compared to inclinometers, which are not considered the gold standard for movement assessment[68]. In addition, the Flexicurve instrument, which presented reproducibility and concurrent validity results, is limited to lumbar region assessment. In this sense, it is possible to affirm that the literature lacks validated instruments for spinal ROM assessment in both the thoracic and lumbar regions.

Conclusions

There are 14 instruments available for assessment of joint ROM in the thoracic and lumbar spine tested as to their repeatability and/or reproducibility, and only six instruments that were assessed for concurrent validity. However, there is scientific evidence only to support the inter-rater reproducibility of the measuring tape instrument only in the modified Schöber's test for lumbar flexion movement, and the intra-rater reproducibility of the Flexicurve and video analysis system instruments for lumbar extension and flexion movements. Nevertheless, adequacy limitations in the statistical analyses in the included studies are highlighted.

Based on GRADE criteria, the results presented in this systematic review indicate little scientific evidence on the validity, repeatability and reproducibility of the instruments and methods indicated for assessment of joint ROM in the thoracic and lumbar spine.

References

1. Behnke R. Anatomia do movimento. Porto Alegre: Artmed; 2014.
2. Correia VG, Foganholi G, Macedo CSG. Lumbar flexion and functional disability: a comparative study between asymptomatic subjects and patients with low back pain. J Health Sci 2015;17(3)194-197.
3. Alter, MJ. Ciência da flexibilidade. Porto Alegre: Artmed; 2010.
4. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380(9859):2163–2196. DOI: 10.1016/S0140-6736(12)61729-2
5. Panjabi MM. The stabilizing system of the spine. Part I. Function, dysfunction, adaptation and enhancement. J Spinal Disord 1992a;5(4):383-9. DOI: 10.1097/00002517-199212000-00001
6. Kisner C, Colby LA. Exercícios terapêuticos: fundamentos e técnicas. São Paulo: Manole; 2005.
7. Zatsiorsky VM. Kinematics of human motion. Champaign: Human Kinetics; 1998.
8. Winter D. Biomechanics and motor control of human movement. New Jersey: Johnwiley e songs; 2009.
9. Vaughan C, Davis B, Jeremy C. Dynamics of human gait. Cape Town: Kiboho Publishers; 1992.
10. Leardini A, BiagiF,Merlo A, Belvedere C, Benedetti MG. Multi-segment trunk kinematics during locomotion and elementary exercises. Clin Biomechanics 2011;26:562–571. DOI:10.1016/j.clinbiomech.2011.01.015
11. International vocabulary of metrology – Basic and general concepts and associated terms. 3rd. ed. Joint Committee for Guides in Metrology; 2012.
12. Associação Brasileira de NormasTécnicas – ABNT. Norma Brasileira Registrada – NBR 10536 – Statistics: Vocabulary and Symbols; 1988.
13. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons, Ltd; 2011.
14. Van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain: a systematic review of randomized controlled trials of the most common interventions. Spine 1997;22(18):2128-56.
15. Brink Y, Louw Q. Clinical instruments: reliability and validity critical appraisal. J Eval Clin Pract 2012;18(6):1126-32. DOI:10.1111/j.1365-2753.2011.01707.x
16. Barret E, McCreesh K, Lewis J. Reliability and validity of non-radiographic methods of thoracic kyphosis measurement: a systematic review. Man Ther 2014;19:10-17. DOI: 10.1016/j.math.2013.09.003
17. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. Grade Guidelines: 1. Introduction Grade evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383–94. DOI: 10.1016/j.jclinepi.2010.04.026

18. Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):01-06. DOI: 10.1016/j.jclinepi.2010.04.026
19. Bandy WD, Reese NB. Strapped versus unstrapped technique of the prone press-up for measurement of lumbar extension using a tape measure: differences in magnitude and reliability of measurements. Arch Phys Med Rehabil 2004;1(85):99-103. DOI:10.1016/S0003-9993(03)00430-1
20. Beattie P, Rothstein JM, Lamb R. Reliability of the attraction method for measuring lumbar spine backward bending. Phys Ther 1987;67(3):364-369.
21. Burdett RG, Kathryn PF, Brown PF, Michael PF. Reliability and validity of four instruments for measuring lumbar spine and pelvic positions. Phys Ther 1986;66(5):677-684.
22. Dopf CA, Mandel SS, Geiger D, Mayer P. Analysis of spine motion variability using a computerized goniometer compared to physical examination. A prospective clinical study. Spine 1994;19(5):586-595.
23. Frost M, Stuckey S, Smalley LA, Dorman G. Reliability of measuring trunk motions in centimeters. Phys Ther 1982; 62:1431-1437.
24. Gill K, Krag HM, Johnson GB, Haugh LD, Popo MH. Repeatability of four clinical methods for assessment of lumbar spinal motion. Spine 1988;13(1):50-53. DOI: 10.1097/00007632-198801000-00012
25. Merritt JL, Mclean TJ, Erickson RP, Offord, K. Measurement of trunk flexibility in normal subjects: Reproducibility of three clinical methods. Mayo Clinic Proc 1986;61(3):192-197. DOI: 10.590/1809-2950/13088921042014
26. Ronchi AJ, Lech M, Taylor NF, Cosic I. A reliability study of the new Back Strain Monitor based on clinical trials. 30th Annual Internacional IEEE Conference. Vancouver; 2008, p.693-696. DOI: 10.1109/IEMBS.2008.4649247
27. Miller SA, Mayer T,Cox R, Gatchel, R.Reliability problems associated with the modified Schober technique for true lumbar flexion measurement. Spine 1992;17(3):345-348. DOI: 10.1097/00007632-199203000-00017
28. Paternostro-Sluga T, Preisinger E, Resh KL, Ernst E. How reproducible is the functional assessment of the spine? Eur J Phys Rehabil Med 1995;5(4):122-125.
29. Hyytiainen K, Salminen JJ, Suvitie T, Wickström G, Pentti J. Reproducibility of nine tests to measure spinal mobility and trunk muscle strength. Scand J Rehabil Med 1991;23(1):3-10.
30. Van Den Dolder PA, Ferreira PH, Refshauge K. Intra and inter-rater reliability of a modified measure of hand behind back range of motion. Man Ther 2014;19(1):72-76. DOI: 10.1016/j.math.2013.08.002
31. Bo K, Storheim HK. Intra- and interobserver reproducibility of Cybex ED1 320 measuring spinal mobility. Scand J Med Sci Sports 1997;7:140-143. DOI: 10.1111/j.1600-0838.1997.tb00130.x
32. Breum J, Wiberg J, Bolton JE. Reliability and concurrent validity of the BROM II for measuring lumbar mobility. J Manipulative Physiol Ther1995;18(8):497-502.
33. Kolber MJ, Pizzini M, Robinson A, Yanez D, Hanney WJ. The reliability and concurrent validity of measurements used to quantify lumbar spine mobility: an analysis of an iphone (R) application and gravity based inclinometry. Int J Sports Phys Ther 2013;8(2):129-137.
34. Mayer RS, Chen I, Lavender SA, Trafimow JH, Andersson GBJ. Variance in the meansurement of sagital lumbar spine range od motion among examiners, subjects, and instruments. Spine 1995;20(13):1489-1493. DOI: 10.1097/00007632-199507000-00008
35. Mellin G, Kiiski R, Weckström A. Effects of subject position on measurements of flexion, extension, and lateral flexion of the spine. Spine 1991;16(9):1108-1110.
36. Ng JKF, Kippers V, Richardson C, Parnianpour M. Range of motion and lordosis of the lumbar spine: Reliability of measurement and normative values. Spine 2001;26(1):53-60.
37. Chiarello CM, Savidge R. Inter-rater reliability of the Cybex EDI-320 and fluid goniometer in normals and patients with low back pain. Arch Phys Med Rehabil 1993;74:32-37.
38. Rondinelli R, Murphy J, Esler A, Marciano T, Cholmakjian C. Estimation of normal lumbar flexion with surface inclinometry. A comparison of three methods. Am J Phys Med Rehabil 1992;71(4):219-224.
39. Boocock MG, Jackson JA, Burton AK, Tillotson KM. Continuous measurement of lumbar posture using flexible electrogoniometers. Ergonomics 1994;37(1):175-185. DOI: 10.1080/00140139408963636
40. Bedekar N, Suryawanshi M, Rairikar S, Sancheti P, Shyam A. Inter and intra-rater reliability of mobile device goniometer in measuring lumbar flexion range of motion. J Back Musculoskelet Rehabil 2014;27(2):161-166. DOI: 10.3233/BMR-130431
41. Salisbury PJ, Porter R. Measurement of lumbar sagittal mobility: A comparison of methods. Spine 1987;2(2):190-193.
42. Paquet N, Malouin F, Richards C, Dionne JP, Comeau F. Validity and reliability of a new electrogoniometer for the measurement of sagittal dorsolumbar movements. Spine 1991;16(5):516-519.

43. Tojima M, Ogata N, Yozu A, Sumitani M, Haga N. Novel 3-dimensional motion analysis method for measuring the lumbar spine range of motion repeatability and reliability compared with an electrogoniometer. Spine 2013;38(21):1327-1333. DOI: 10.1097/BRS.0b013e3182a0dbc5
44. Gill KP, Callaghan MJ. Intratester and intertester reproducibility of the lumbar motion monitor as a measure of range, velocity and acceleration of the thoracolumbar spine. Clin Biomechanics 1996;11(7):418-421. DOI:10.1016/0268-0033(96)00031-9
45. Mannion A, Troke M. A comparison of two motion analysis devices used in the measurement of lumbar spinal mobility. Clin Biomechanics 1999;14(9):612-619. DOI: 10.1016/S0268-0033(99)00017-0
46. Petersen CM, Johnson RD, Schuit D, Hayes K. Intraobserver and interobserver reliability of asymptomatic subjects thoracolumbar range of motion using the OSI CA 6000 spine motion analyzer. J Orthop Sports Phys Ther 1994;20(4):207-212. DOI: 10.2519/jospt.1994.20.4.207
47. Pearcy MJ, Hindle RJ. New method for the non-invasive three-dimensional measurement of human back movement. Clin Biomechanics 1989;4:73-79. DOI:10.1016/0268-0033(89)90042-9
48. Troke M, Schuit D, Petersen CM. Reliability of lumbar spinal palpation, range of motion, and determination of position. BMC MusculoskeletDisord2007;8(103):1-6. DOI:10.1186/1471-2474-8-103
49. Schuit D, Petersen C, Johnson R, Levine P, Knecht H, Goldberg D. Validity and reliability of measures obtained from the OSI CA-6000 Spine Motion Analyzer for lumbar spinal motion. Man Ther 1997;2(4):206-215. DOI: 10.1054/math.1997.0301
50. Tillotson KM, Burton AK. Noninvasive measurement of lumbar sagittal mobility an assessment of the flexicurve technique. Spine 1991;16(1):29-33.
51. Burton AK. Regional lumbar sagittal mobility: Measurement by flexicurves. Clin Biomechanics 1986;1:20-26. DOI: 10.1016/0268-0033(86)90032-X
52. Burton AK. Measurement of regional lumbar sagittal mobility. J Orthop Sports Phys Ther1987;166-169.
53. Youdas JW, Suman VJ, Garrett TR. Reliability of measurements of lumbar spine sagittal mobility obtained with the flexible curve. J Orthop Sports Phys Ther 1995;21(1):13-20. DOI:10.2519/jospt.1995.21.1.13
54. Alqhtani RS, Jones MD, Theobald P, Williams J. Reliability of an accelerometer- based system for quantifying multiregional spinal range of motion. J Manipulative PhysiolTher 2015;38(4):275-81. DOI: 10.1016/j.jmpt.2014.12.007
55. Consmuller T, Rohlmann A, Weinland D, Druschel C, Duda G, Taylor W. Comparative evaluation of a novel measurement tool to assess lumbar spine posture and range of motion. Eur Spine J 2012;21(11):2170-2180. DOI: 10.1007/s00586-012-2312-1
56. Tederko P, Krasuski M, Maciejasz P. Restrainer of pelvis and lower limbs in thoracic and lumbar range of motion measurement – preliminary report. OrtopTraumatolRehabil 2007;2(6):156-167.
57. Edmondston SJ, Christensen M, Keller S, MClinPhysio PT, Steigen L, Barclay L. Functional radiographic analysis of thoracic spine extension motion in asymptomatic men. J Manipulative PhysiolTher 2012;35(3):203-208. DOI: 10.1016/j.jmpt.2012.01.008
58. Ha T, Sheikh K, Moore AP, Jones MP. Measurement of lumbar spine range of movement and coupled motion using inertial sensors - A protocol validity study. Man Ther 2013;18:87-91. DOI: 10.1016/j.math.2012.04.003
59. Yun w, Kim H, Ahn JH, Park Y, Park Y. Individual characteristics of reliable lumbar coupling motions. Eur Spine J 2015;24:1917–1925. DOI: 10.1007/s00586-015-4081-0
60. Roussel N, Nijs J, Truijen S, Breugelmans S, Claes I, Stassijns G. Reliability of the Assessment of Lumbar Range of Motion and Maximal Isometric Strength. Arch Phys Med Rehabil 2006;87(4)576-582. DOI: 10.1016/j.apmr.2006.01.007
61. Williams, JM, Haq, Lee RY. Dynamic measurement of lumbar curvature using fibre-optic sensors. Med Eng Phys 2010;32(9):1043-1049. DOI: 10.1016/j.medengphy.2010.07.005
62. Lee RYW, Laprade J, Fung EHK. A real-time gyroscopic system for three-dimensional measurement of lumbar spine motion. Med Eng Phys 2003;25:817–824. DOI: 10.1016/S1350-4533(03)00115-2
63. Cohn ML, Machado AF, Cohn SJ. Low-Frequency magnetic field technology: Quantifying spinal range of motion. Arch Phys Med Rehabil 1999;70(6):455-457. DOI: 10.1016/0003-9993(89)90006-3
64. Fölsch C, Schlögel S, Lakemeier S, Wolf U, Timmesfeld N, Skwara A. Test-Retest Reliability of 3D Ultrasound Measurements of the Thoracic Spine. AAPMR 2012;4(5):335-341. DOI: 10.1016/j.pmrj.2012.01.009
65. Norkin CC, White DJ. Medida do movimento articular: manual do goniômetro. 2 ed. Porto Alegre: Artes médicas;1997.
66. Hamill J, Knutzen K. Bases biomecânicas do movimento humano. 3. ed. Barueri: Manole; 2012.
67. Oliveira TS, Candotti CT, La Torre M, Pelinson PT, Furlanetto TS, Kutchak FM, et al. 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. Rehabil Res Pract 2012;12:01-09. DOI: 10.1155/2012/186156
68. Wu G. Letter to the editor. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion— part i: ankle, hip, and spine. J Biomech 2002;35:543-548.

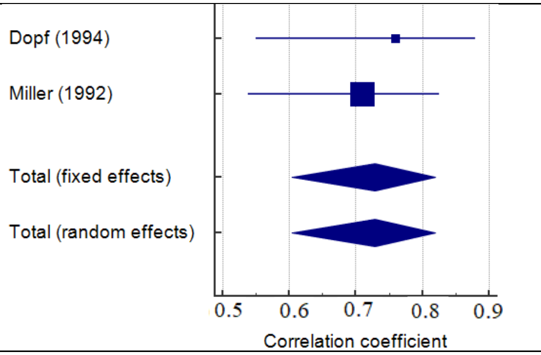
Received on Mar, 16, 2017.
Reviewed on Jun, 02, 2017.
Accepted on Oct, 30, 2017.

Author address: Marja Bochehin do Valle - Avenida Mariland 156/603, Porto Alegre - RS – E-mail: marjabv@hotmail.com

APPENDIX

Table 3. Meta-analysis result for inter-rater reproducibility of the measuring tape instrument in the modified Schöber's test for lumbar flexion movement

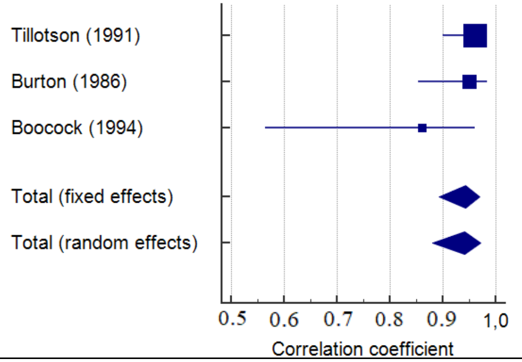
Study	Sample Size	Correlation Coefficient	95% IC	z	p	Weight (%)	
						Fixed	Random
Dopf[22]	30	0.76	0.55 – 0.88			36.49	36.49
Miller[27]	50	0.71	0.54 – 0.83			63.51	63.51
Total (fixed effects)	80	0.73	0.60 – 0.82	7.97	<0.001	100.00	100.00
Total (random effects)	80	0.73	0.60 – 0.82	7.97	<0.001	100.00	100.00



Note: Heterogeneity test: Q=0.20; DF=1; p=0.652; I^2=0.00%; 95% IC for I^2=0.00 – 0.00.

Table 4. Meta-analysis for intra-rater reproducibility of the Flexicurve instrument for lumbar extension movement

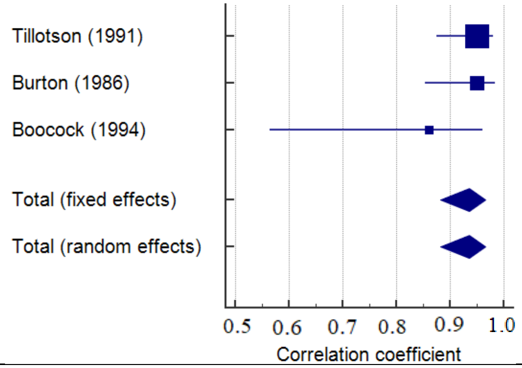
Study	Sample Size	Correlation Coefficient (r)	95% IC	z	p	Weight (%)	
						Fixed	Random
Tillotson[50]	20	0.96	0.90 – 0.98			44.74	41.86
Burton[51]	15	0.95	0.86 – 0.98			31.58	32.38
Boocock[39]	12	0.86	0.57 – 0.96			23.68	25.76
Total (fixed effects)	47	0.94	0.89 – 0.97	10.82	<0.001	100.00	100.00
Total (random effects)	47	0.94	0.88 – 0.97	9.3	<0.001	100.00	100.00



Note: Heterogeneity test: Q=2.61; DF=2; p=0.271; I^2=23.33%; 95% IC for I^2=0.00 – 97.43.

Table 5. Meta-analysis result for intra-rater reproducibility of the Flexicurve instrument for lumbar flexion movement

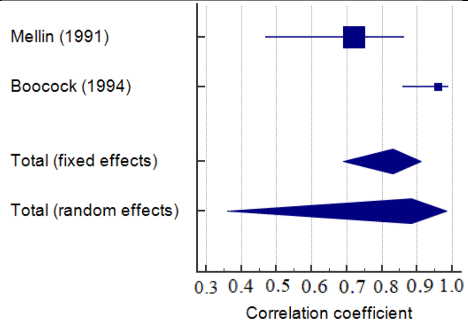
Study	Sample Size	Correlation Coefficient (r)	95% IC	z	p	Weight (%)	
						Fixed	Random
Tillotson[50]	20	0.95	0.88 – 0.98			44.74	44.74
Burton[51]	15	0.95	0.85 – 0.98			31.58	31.58
Boocock[39]	12	0.86	0.57 – 0.96			23.68	23.68
Total (fixed effects)	47	0.94	0.88 – 0.97	10.51	<0.001	100.00	100.00
Total (random effects)	47	0.94	0.88 – 0.97	10.51	<0.001	100.00	100.00



Note: Heterogeneity test: Q=1.99; DF=2; p=0.369; I²=0.00%; 95% IC for I²=0.00 – 96.63.

Table 6. Meta-analysis result for intra-rater reproducibility of the inclinometer instrument for lumber extension movement

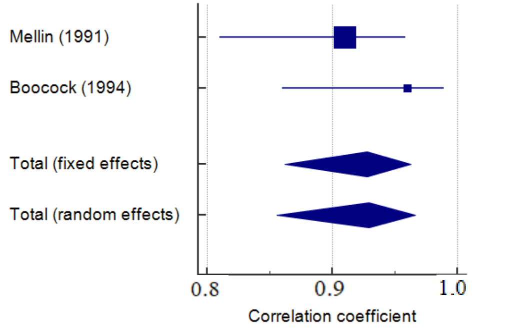
Study	Sample Size	Correlation Coefficient (r)	95% IC	z	p	Weight (%)	
						Fixed	Random
Mellin[35]	27	0.72	0.47 – 0.86			72.73	53.22
Boocock[39]	12	0.96	0.86 – 0.99			27.27	46.78
Total (fixed effects)	39	0.83	0.69 – 0.91	6.84	<0.001	100.00	100.00
Total (random effects)	39	0.88	0.36 – 0.98	2.69	0.007	100.00	100.00



Note: Heterogeneity test: Q=7.06; DF=1; p=0.079; I²=85.83%; 95% IC for I²=43.14 – 96.47.

Table 7. Meta-analysis result for intra-rater reproducibility of the inclinometer instrument for lumbar flexion movement

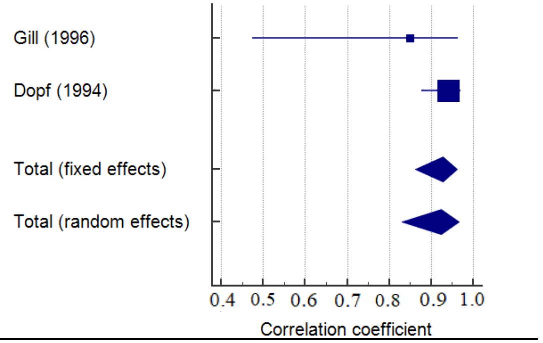
Study	Sample Size	Correlation Coefficient (r)	95% IC	z	p	Weight (%) Fixed	Weight (%) Random
Mellin[35]	27	0.91	0.81 – 0.96			72.73	69.84
Boocock[39]	12	0.96	0.86 – 0.99			27.27	30.16
Total (fixed effects)	39	0.93	0.86 – 0.96	9.43	<0.001	100.00	100.00
Total (random effects)	39	0.93	0.86 – 0.96	8.61	<0.001	100.00	100.00



Note: Heterogeneity test: Q=7.06; DF=1; p<0.079; I^2=85.83%; 95% IC for I^2=43.14 – 96.47.

Table 8. Meta-analysis result for intra-rater reproducibility of the video analysis system instrument for lumber extension movement

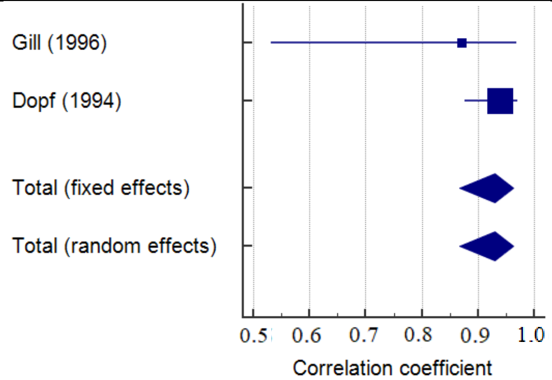
Study	Sample Size	Correlation Coefficient (r)	95% IC	z	p	Weight (%) Fixed	Weight (%) Random
Gill[44]	10	0.85	0.47 – 0.96			20.59	27.22
Dopf[22]	30	0.94	0.88 – 0.97			79.41	72.78
Total (fixed effects)	40	0.93	0.86 – 0.96	9.56	<0.001	100.00	100.00
Total (random effects)	40	0.92	0.83 – 0.97	7.49	<0.001	100.00	100.00



Note: Heterogeneity test: Q=1.29; DF=1; p=0.256; I^2=22.53%; 95% IC for I^2=0.00 – 100.00.

Table 9. Meta-analysis result for intra-rater reproducibility of the video analysis system instrument for lumbar flexion instrument

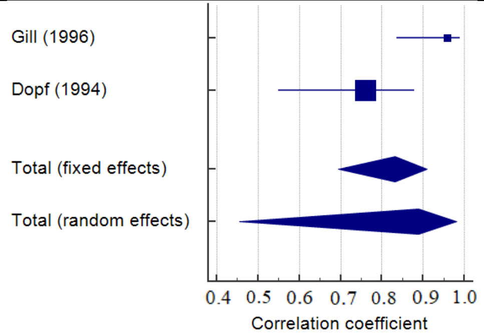
Study	Size Sample	Correlation Coefficient (r)	95% IC	z	p	Weight (%) Fixed	Weight (%) Random
Gill[44]	10	0.87	0.53 – 0.97			20.59	20.59
Dopf[22]	30	0.94	0.88 – 0.97			79.41	79.41
Total (fixed effects)	40	0.93	0.87 – 0.96	9.65	<0.001	100.00	100.00
Total (random effects)	40	0.93	0.87 – 0.96	9.65	<0.001	100.00	100.00



Note: Heterogeneity test: Q=0.92; DF=1; p=0.340; I²=0.00%; 95% IC for I²=0.00 – 0.00.

Table 10. Meta-analysis result for inter-rater reproducibility of the video analysis system instrument for lumbar extension movement

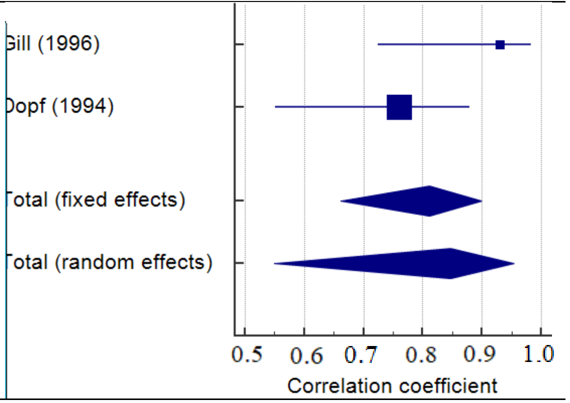
Study	Sample Size	Correlation Coefficient (r)	95% IC	z	p	Weight (%) Fixed	Weight (%) Random
Gill[44]	10	0.96	0.83 – 0.99			20.59	44.13
Dopf[22]	30	0.76	0.55 – 0.88			79.41	55.87
Total (fixed effects)	40	0.83	0.69 – 0.91	6.95	<0.001	100.00	100.00
Total (random effects)	40	0.89	0.46 – 0.98	3.00	0.003	100.00	100.00



Note: Heterogeneity test: Q=5.01; DF=1; p=0.025; I²=80.05%; 95% IC for I²=14.25 – 95.36.

Table 11. Meta-analysis result for inter-rater reproducibility of the video analysis system instrument for lumbar flexion movement

Study	Sample Size	Correlation Coefficient (r)	95% IC	z	p	Weight (%)	
						Fixed	Random
Gill[44]	10	0.93	0.72 – 0.98			20.59	37.93
Dopf[22]	30	0.76	0.55 – 0.88			79.41	62.07
Total (fixed effects)	40	0.81	0.66 – 0.90	6.60	<0.001	100.00	100.00
Total (random effects)	40	0.85	0.55 – 0.95	3.88	<0.001	100.00	100.00



Note: Homogeneity test: Q=2.44; DF=1; p=0.119; I^2=58.97%; 95% IC for I^2=0.00 – 90.33.