Instrumented Measurement of Spinal Stiffness: A Systematic Literature Review of Reliability

Léonie Hofstetter, BSC, Melanie Häusler, BSC, Brigitte Wirth, PhD, and Jaap Swanenburg, PhD

ABSTRACT

Objective: The purpose of this review was to critically appraise the quality of studies evaluating the reliability of spinal stiffness assessment devices.

Methods: An electronic search of the MEDLINE, PubMed, CINAHL, PsycINFO, PEDro, and Embase databases up to September 2016 was performed. Information on participants, measurement protocols, reliability, and accuracy were extracted. Two reviewers independently applied the COnsensus-based Standards for the selection of health Measurement INstruments checklist to assess the methodological quality of the measurement properties reliability and measurement error, which were rated as excellent, good, fair, or poor. The overall score was determined using the worst score counts method.

Results: In total, 1,728 studies were identified and 9 studies were included in this review. All included studies showed high reliability, with intraclass correlation coefficient values ranging from 0.65 to 0.99. In the quality assessment, 2 studies were rated as fair and 7 studies as poor, mainly because of sample sizes.

Conclusion: The studies demonstrated favorable high-reliability values but low methodological quality. In the future, high-quality studies with larger sample sizes are needed. (J Manipulative Physiol Ther 2018;41:704-711)

Key Indexing Terms: *Spine; Reliability*

INTRODUCTION

The measurement of spinal stiffness in patients with back pain is part of the daily routine of a manual therapist.[1] Biomechanically, stiffness is resistance against an applied force to make a change in position.[2] In a clinical environment, stiffness is resistance to the elastic deformation of a tissue, joint, or limb.[3] In spinal stiffness, clinicians manually assess the posterior-to-anterior stiffness by pressing on the spinous process while the patient is lying in a prone position. The resulting movement of the vertebra is determined by the patient's pain level, presence of muscle spasm, and physical resistance (ie, stiffness).[4] Clinicians consider the relationship between the applied force and the resulting displacement of the spine[5] to determine whether a joint is hypo- or hypermobile. A study by Fritz et al showed that the judgment about hypo- or hypermobility provides useful information for clinical decision-making.[6] Flynn et al found that spinal stiffness is 1 of the 5 predictors of response to spinal manipulation in participants with low back pain (LBP).[7] Information about spinal stiffness can therefore help clinicians to identify the restricted spinal segment and to determine the diagnosis and the options for further treatment.[8]

Spinal stiffness is also influenced by other factors such as pain or intra-abdominal pressure. Several studies have demonstrated that spinal stiffness is increased in patients with LBP compared with people with little or no pain.[9,10] Hodges et al described augmentation of lumbar spinal stiffness owing to increased intra-abdominal pressure caused by the stimulation of the phrenic nerves, provoking a contraction of the diaphragm.[11] Shirley et al investigated the changes of stiffness during the respiratory cycle. They found higher values for measurements above and below the normal tidal volume compared with values obtained after a normal exhalation.[12]

Even though the manual assessment of stiffness often is done in clinical practice, many studies have found that it has poor intra- and interrater reliability.[13-15] Some reasons are examiner variability in the magnitude[16] or the direction[17] of the applied force and the variation in individual perceptions of stiffness.[18] Maher et al investigated a

Integrative Spinal Research, Department of Chiropractic Medicine, Balgrist University Hospital, Zürich, Switzerland.

Corresponding author: Jaap Swanenburg, PhD, Integrative Spinal Research, Department of Chiropractic Medicine, Balgrist University Hospital, Forchstrasse 340, CH-8008, Zürich, Switzerland. Tel.: +41 44 510 73 82. (e-mail: *jaap.swanenburg@balgrist.ch*).

Paper submitted October 10, 2017; in revised form March 30, 2018; accepted March 30, 2018.

0161-4754

https://doi.org/10.1016/j.jmpt.2018.03.002

reference-based protocol, where examiners received a reference stiffness stimulus to rate the stiffness more objectively. The reliability of this revised protocol was acceptable, with an intraclass correlation coefficient (ICC) of 0.77, but the validity was not. The result suggested that the examiners and the mechanical device measured different things.[19]

Thus, for a more objective quantification of spinal stiffness, researchers developed instrumented measurements. The first mechanical device, called a spinal physiotherapy simulator, was constructed by Lee and Svensson in 1990 and was found to be highly reliable.[20] In 1996, Latimer et al designed a lighter, more easily transportable device to measure spinal stiffness in patients with LBP. This was the first study that demonstrated highly reliable measurement results in participants with LBP.[21] However, these mechanical devices primarily were used for research applications because their size and complexity made it difficult to use them in clinical practice. Thus, there was a need for smaller devices, and handheld mechanically assisted devices were developed.[22,23] With these, the examiner manually applies force for the tissue indentation and receives feedback from a computerized force and displacement monitoring system. Even though the reliability was slightly less than that reported for the mechanical devices, the ICC values of 0.79 and 0.93 showed clearly better reliability than the manual posterior-to-anterior pressure test.[22,23]

Although the early mechanical devices, like the spinal physiotherapy simulator[20] or the portable stiffness device,[21] applied the concept of indented force and resulting displacement of the skin, these handheld mechanically assisted devices use different methods to assess spinal stiffness. The device used by Evans et al measures the compliance of the vertebra based on the theory of impulse response.[24] PulStar, another mechanically assisted device, compares consecutive spinal segments and identifies variance in resistance. The examiner produces a preload force of 3 to 4 lbs before the handheld impulse head generates an impulse. With use of a piezoelectric sensor, the stiffness resulting from this impulse can be monitored.[25] Another newer mechanical indentation device uses a motorized indentation probe and a compressive-tension load cell transducer to measure the resulting displacement.[26]

As highly reliable methods are needed for precision in diagnosis and treatment evaluation, mechanical or mechanically assisted devices will become increasingly important. Nevertheless, to our knowledge, there are no systematic overviews of the reliability of instrumented measurements of spinal stiffness. Wong et al included a narrative review about the reliability and accuracy of indentation devices in their study,[26] but more devices have been developed since then.[27] Therefore, the aim of this review was to systematically review the literature on the reliability of instrumented measurements of spinal stiffness in healthy participants and in participants with neck pain and LBP.

Methods

Search Strategy

An electronic search strategy was developed and performed by a librarian at Zurich University. The databases MEDLINE, PubMed, CINAHL, PsycINFO, PEDro, and Embase were included in the search. The search was restricted to English- and German-language literature up to September 2016. The search strategy used for these databases is shown in more detail in the supplementary files. Moreover, the reference lists of the identified studies were scanned for potentially useful supplementary articles.

Study Selection

After duplicates were removed, 1 reviewer (J.S.) screened the search results for eligibility. First, the titles and abstracts were screened, and irrelevant studies were excluded. The full text of potentially eligible articles then was screened to assess eligibility. The remaining studies were checked for eligibility by 2 other authors (B.W. and L. H.), and any disagreements were resolved through discussion. Articles were included if they met the following criteria: (1) mechanical spinal stiffness assessment was the main topic, (2) reliability was reported, and (3) they were in English or German. Studies were excluded for the following reasons: (1) manual spinal stiffness assessment was the focus, (2) spinal stiffness assessment was done during spine fusion surgery/anesthesia, (3) stiffness assessment was of cadaveric or nonhuman material, or (4) full text was not available.

Data Extraction

One author (L.H.) extracted the following data from each study: author, year of publication, name of the device assessed, number of participants, and whether the participants were symptomatic or asymptomatic. Information about the level of assessment, patient position, and maximal applied load also was collected. Furthermore, details pertaining to the reliability of the device, such as reliability type, time interval, averaged measurements, and ICC, were extracted. Furthermore, data about the accuracy of the device were collected. If uncertainty arose, it was discussed with a second author (J.S.).

Study Quality Assessment

The methodological quality of all studies was assessed using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. This checklist is a standardized tool for evaluating the methodological quality of studies on measurement properties. It consists of 12 boxes, 9 of which contain standards for different measurement properties.[28,29]

Because this review focuses on the reliability of mechanical devices, the COSMIN box B (reliability) was used. Because it is also important to know the influence of the measurement error on the accuracy of the measuring devices, box C (measurement error) also was used. Two reviewers (B.W. and J.S.) independently rated each of the items of these 2 boxes using a 4-point rating scale containing the response options poor, fair, good, and excellent. The methodological quality score for these boxes was determined by taking the lowest rating of any item (worst score counts).[29] Any disagreements were resolved through discussion.

Results

Search Outcome

The electronic search strategy yielded a total of 1728 studies. As shown in the flowchart of our study selection process (Fig 1), 9 studies were included in this review.

Study Characteristics

Descriptions of the 9 included studies are given in Table 1.[20-27,30] The number of participants ranged between 3 and 52, with only 1 study including more than 50 participants.[22] Three studies measured the spinal stiffness of symptomatic participants,[21,22,26] whereas the others assessed healthy participants. With the exception of 2 studies that assessed spinal stiffness while sitting[24] and standing,[27] the studies assessed spinal stiffness in the prone position.[24,27] Half of the studies measured stiffness of the lumbar spine.[20,23,26] The other studies investigated stiffness of the cervical spine,[30] the cervical and thoracic spine,[24] the thoracic spine,[27] or the whole spine.[25] The maximal load varied from 25 to 115 N. On average, the studies applied a maximal load of around 80 to 100 N. Only 3 studies evaluated the accuracy of the devices reporting the standard error of measurement.[22,26,27]

Reliability

Data on the reliability of the instruments are shown in Table 1. All studies measured the intrarater reliability,[20-23,26,27,30] and 2 of them additionally tested interrater reliability.[24,25] Four studies determined within-day reliability,[21,23,24] 2 studies between-day reliability,[20,27] and 3 studies a combination of these.[22,26,30] One study did not specify the time interval between the 2 measurements.[25] To quantify reliability, 7 studies used ICCs,[20-26] 1 study used the dependability coefficient Φ,[27] and 1 study used the coefficient of multiple determination.[30] The reliability ranged between ICC values of 0.65 to 0.99, with 4 studies attaining a value above 0.90.[21,23,24,26]

Methodological Quality

The resulting quality scores for the COSMIN boxes B (reliability; relative measures) and C (measurement error; absolute measures) are shown in Table 1. Items 1 to 10 of box C are the same as those in box B. Therefore, we show only item 11 from box C (11C). In both boxes, 2 studies were rated as fair[22,27] and 7 studies as poor.[20,21,23-26,30] None of the studies were classified as good or excellent.

The COSMIN items of box B and item 11 of box C are shown in Figure 2. The COSMIN mean percentage score per item is shown in Figure 3. In all of the studies, item 4 (2 measurements available) and 10 (important flaws) were rated as excellent. Item 9 (similar test conditions for both measurements) also showed high scores; 67% were evaluated as good[20-22,24,27,30] and the rest as excellent.[23,25,26] For item 7 (stability of patients), 89% of the studies received a good rating.[20-24,26,27,30] Seven studies had a poor rating for item 3 (adequate sample size).[17,21,23-26,30] Items 1 and 6 (statement of missing items and time interval, respectively) offered only the following 2 response options: (1) excellent (11%)[26] and good (82%)[20-25,27,30] and (2) excellent (89%)[20-24,26,27,30] and fair (11%),[25] respectively. Six studies calculated the ICC and described a model or formula for the ICC, which resulted in an excellent rating for item 11 of box B (calculation of ICC).[20-23,26,27] Two studies were estimated as good[24,25] and 1 study as fair.[30] For item 11 of box C (calculation of standard error of measurement, smallest detectable change, or limits of agreement), 45% of the studies achieved an excellent rating,[22,26,27] with the remaining receiving a poor rating.[20,21,23-25,30]

Rater Agreement

The percentage of agreement between the 2 raters was 92% (Cohen's κ = 0.94). The quality assessment differed twice for item 5 (independent administration) and item 8 (appropriate time interval)[25,27] and once for item 6 (statement of time interval) and item 9 (similar conditions for both measurements).[26] These disagreements were resolved through discussion.

Discussion

This systematic review evaluated 9 studies assessing the reliability of instrumented measurements of spinal stiffness. The included studies showed high variability in reliability estimates, patient populations, methodological quality, and study design.

Reliability

Overall, all mechanical devices except 1[24] showed acceptable to excellent ability to reliably measure spinal stiffness. These studies used different measurement protocols with various time intervals. Thus, these factors do not

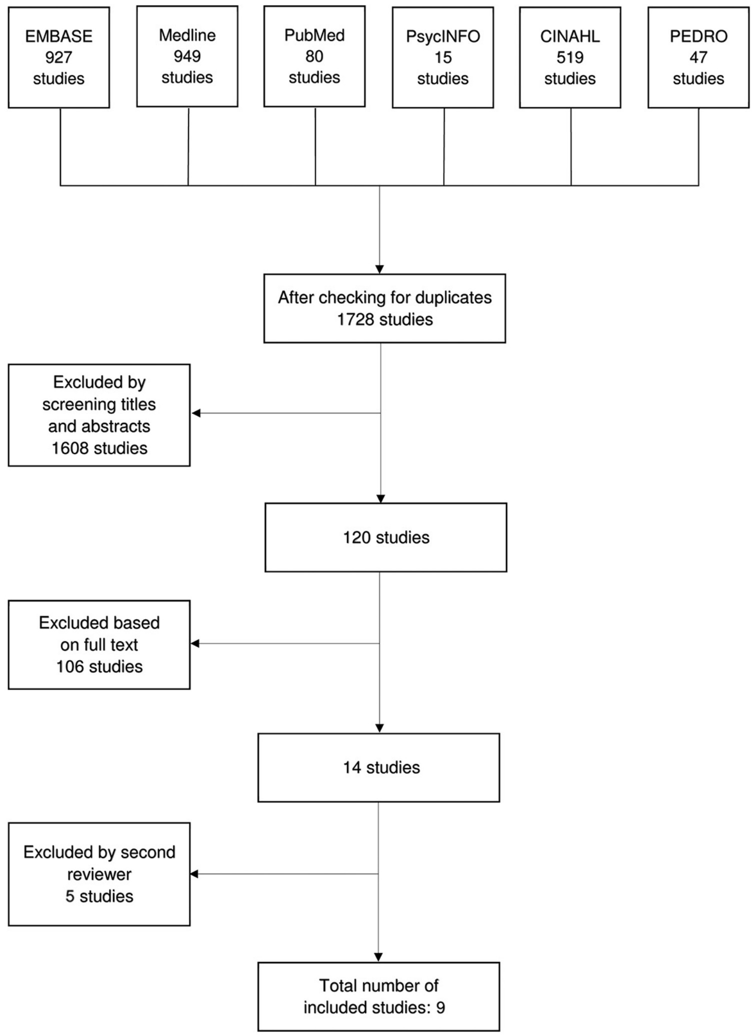

Fig 1. *Search flowchart of study selection process.*

seem to influence the reliability of the devices. All studies except 2 used the ICC to calculate reliability.[27,30] The only study that measured spinal stiffness in a free-sitting position had the lowest interrater ICC value.[24] This position provides less stabilization of the spine during measurement than the prone position, which might explain this finding. Although most studies measured asymptomatic participants, 3 studies evaluated the reliability of spinal stiffness in symptomatic participants.[21,22,26] However, there was no difference in reliability from those studies testing patients.

Methodological Quality

The low methodological quality of the studies limits the strength of conclusions. Seven studies[20,21,23-26,30] were rated as poor and 2 studies[22,27] as fair for both COSMIN box B (reliability; relative measures) and box C (reliability; absolute measures). The main reason for the low quality assessment in most studies was the sample size, with 7 of the included studies using a sample size <30[20,21,23-25,27,30] and only 1 study including a sample size of >50.[22] The stability of the participants during and between the measurements is crucial for repeated measurements. In most studies, it was presumable that the participants were stable, which led to a good score.[20-24,26,27,30] Leach et al received a poor rating because several participants experienced temporary pain because of an accident or trauma.[25] The time interval between 2 measurements was stated in most of the studies.[20-24,26,27,30] However, there was no standard or uniform measurement protocol, and so the interval varied widely between seconds and weeks. Another reason for the low quality rating was item 11 of box C (reliability; absolute measures). Only 3 studies calculated the measurement error in the form of a standard error of measurement[22,27] or a minimum

Table 1. *Study Characteristics*

Study	Device	No.	Pain	Level	Position	Max. load (*N*)	Reliability Type	Time Interval	AM	Reliability (ICC)	Accuracy (SEM)	Quality: Box B	Quality: Box C
Evans, 1998[24]	Compliance detection instrument with function recording and analysis system	3 20	-	C0-T3	Sitting	?	Inter-rater Intrarater	 Within-day: seconds	2 2	0.65 0.90, 0.93	-	Poor	Poor
Larivière, et al, 2015[27]	Trunk perturbation apparatus	30	No	T8	Standing	?	Intrarater	Between-day: 8 wk	?	Φ = 0.41-0.85	1.1-11.9	Fair	Fair
Latimer, et al, 1996[21]	Portable stiffness device	22	Yes	L2-5[a]	Prone	105	Intrarater	Within-day: 5 min	4	0.96	-	Poor	Poor
Leach, et al, 2003[25]	PulStar	18	No	C0-L5	Prone	?	Inter-rater Intrarater	 Within-day:?	2 2	0.87 0.78, 0.89	-	Poor	Poor
Lee and Svensson, 1990[20]	Spinal physiotherapy simulator	11	No	L3	Prone	115	Intrarater	Between-day: days	5	0.88	-	Poor	Poor
Owens, et al, 2007[22]	Posterior-to-anterior spinal stiffness device	52	Yes	L1-L5	Prone	80	Intrarater	Combined within-/ between-day: 10-wk period	≤4	0.79	1.62	Fair	Fair
Stanton, et al, 2009[23]	Assisted indentation device	23	No	L4	Prone	100	Intrarater	Within-day: 2 min	9	0.93	-	Poor	Poor
Tuttle, et al, 2008[30]	Passive movement assessment device	10	No	C2-C7	Prone	25	Intrarater	Within-day: minutes Between-day: 1 day	3 3	CMD = 0.99 CMD = 0.96	-	Poor	Poor
Wong, et al, 2013[26]	Mechanical indentation device	12 14	Yes No	L3	Prone	60	Intrarater	Within-day: 2 min Between-day: 1-4 d	3 3	0.99 0.98	MDC95 = 3.51%-5.36% MDC95 = 8.33%-8.93%	Poor	Poor

AM, averaged measurement; *CMD*, coefficient of multiple determination; *ICC*, intraclass correlation coefficient; *MDC95*, minimal detectable change based on 95% confidence interval; *SEM*, standard error of the mean; *Φ*, dependability coefficient.

[a] The most painful level.

1.	Was the percentage of missing items given?
2.	Was there a description of how missing items were handled?
3.	Was the sample size included in the analysis adequate?
4.	Were at least 2 measurements available?
5.	Were the administrations independent?
6.	Was the time interval stated?
7.	Were patients stable in the interim period on the construct to be measured?
8.	Was the time interval appropriate?
9.	Were the test conditions similar for both measurements (eg, type of administration, environment, instructions)?
10.	Were there any important flaws in the design or methods of the study?
11B.	Was an intraclass correlation coefficient (ICC) calculated?
11C.	Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated?

Fig 2. *COnsensus-based Standards for the Selection of Health Measurement INstruments (COSMIN) item list of Box B and item 11 of Box C.*

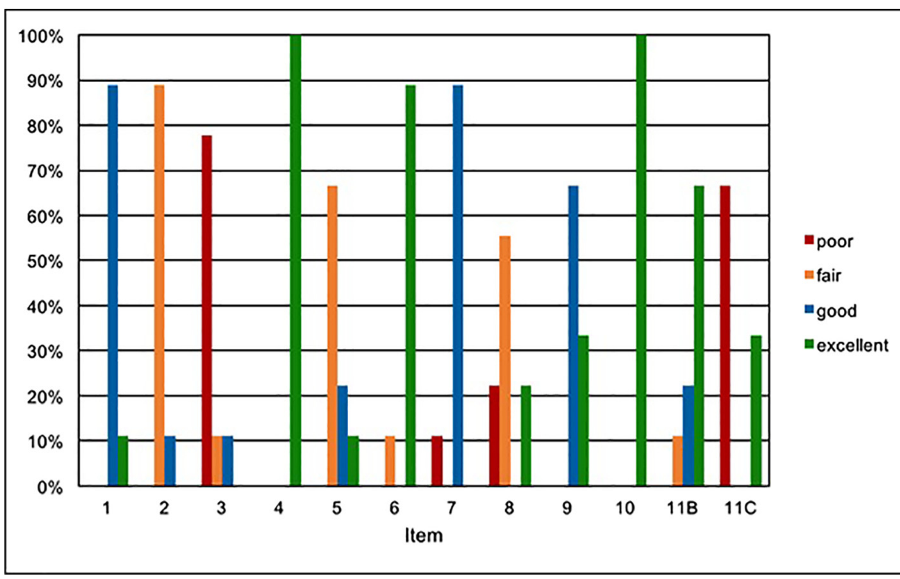

Fig 3. *Mean percentage in quality distribution of each COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) item.*

detectable change.[26] The rest of the studies did not give any details about the accuracy of their device, which would be important for achieving reliable measurements.

Design

This review demonstrates that the reliability of the instrumented measurements is excellent for both asymptomatic and symptomatic participants. It also shows that spinal stiffness generally was tested in a prone position. However, usually back pain is worst when sitting or standing. Both Edmondston et al (lumbar spine)[31] and Snodgrass et al (the neck)[32] showed that body position influences spinal stiffness. In the study by Evans et al, participants were sitting unrestricted on a chair or bench during measurement, which reduced reliability.[24] Therefore, when measuring spinal stiffness in positions other than lying prone, a reliable definition of the body position is required.

Compared with the manual assessment of spinal stiffness, the instrumented measurement showed considerably higher reliability. For the manual posterior-to-anterior pressure test, Binkley et al and Maher et al found poor reliability, with ICC values of 0.03 to 0.37.[14,15] Nevertheless, it is doubtful whether mechanical devices are used in a clinical setting because of their size, their cumbersome nature, and necessary examiner training.[23] However, if highly reliable values are needed, as for example in research, instrumented measurements are useful tools for the objective quantification of spinal stiffness.

For an acceptable confidence interval around the estimated reliability value, an adequate sample size is important,[33] and a sample size of at least 50 participants is required for reliability studies.[33,34] However, this review found that most of the studies used a study population of <30 participants. Thus, future studies with larger sample sizes will be necessary to verify the high reliability values of the mechanical devices measuring spinal stiffness.

Limitations

Some reliability studies on instrumented stiffness assessment had to be excluded during the study selection process because they used cadaveric or nonhuman materials. For example, Vaillant et al used springs of varying stiffness to assess the reliability of the device.[35] In the study by Snodgrass et al, the aim was to measure the stiffness of the cervical spine. However, for the reliability testing of the mechanical device, they measured foams of different densities.[36] Similarly, the mechanical-assisted device of Wang et al was tested using different kinds of foam.[37] Nevertheless, all these studies showed very high ICC values of 0.98 to 1.00. This confirms the highly reliable results of the mechanical devices that were investigated in this review, although stiffness of the human spine is difficult to measure because it is influenced by several factors.[11,12,38] Moreover, Kawchuk and Fauvel presented several sources of variation that may reduce the reliability of devices, such as indentation site relocation, participant movement, and intra-abdominal pressure or muscular response.[39] Another limitation of this review was that the literature search was restricted to the English and German languages. We therefore may have missed some eligible studies in other languages. Moreover, unpublished data were not included, which could have caused a certain publication bias because unpublished studies more likely found negative results.

Conclusion

Only a few studies were found that assessed the reliability of devices mechanically measuring stiffness of the spine. Both mechanical and mechanically assisted devices demonstrated favorable high ICC values. The intrarater reliability ranged between 0.78 and 0.99, and the interrater reliability ranged between 0.65 and 0.87. The results also showed that spinal stiffness of both asymptomatic and symptomatic participants can be measured reliably with these devices. However, the findings were limited by the low quality of the included studies, which was mainly due to the small sample sizes. Furthermore, most studies tested spinal stiffness in a prone position, but the stiffness of the human spine is influenced by many factors, including body position.[9-12,32,33,39] Thus, further reliability studies that measure spinal stiffness in other positions, such as sitting or standing, are needed. Overall, there is a lack of high-quality studies including large sample sizes. Furthermore, standardized measurement protocols are needed to acquire reliable spinal stiffness measurements.

Funding Sources and Conflicts of Interest

No funding sources or conflicts of interest were reported for this study.

Contributorship Information

Concept development (provided idea for the research): L.H., J.S.
Design (planned the methods to generate the results): L.H., B.W., J.S.
Supervision (provided oversight, responsible for organization and implementation, writing of the manuscript): L.H., J.S.
Data collection/processing (responsible for experiments, patient management, organization, or reporting data): L.H., M.H., B.W., J.S.
Analysis/interpretation (responsible for statistical analysis, evaluation, and presentation of the results): L.H., M.H., B.W., J.S.
Literature search (performed the literature search): L.H., M.H., J.S.
Writing (responsible for writing a substantive part of the manuscript) L.H.
Critical review (revised manuscript for intellectual content, this does not relate to spelling and grammar checking): L.H., M.H., B.W., J.S.

Practical Applications

- Most studies testing mechanical devices showed acceptable to excellent ability to reliably measure spinal stiffness.
- Most studies testing mechanical devices showed low methodological quality, mainly because of small sample size.
- Most studies tested mechanical devices in a prone position.

References

1. Henderson CNR. The basis for spinal manipulation: chiropractic perspective of indications and theory. *J Electromyogr Kinesiol.* 2012;22(5):632-642.
2. Oatis C. The use of a mechanical model to describe the stiffness and damping characteristics of the knee joint in healthy adults. *Phys Ther.* 1993;73(11):740-749.

3. Edmondston SJ, Allison GT, Althorpe BM, McConnell DR, Samuel KK. Comparison of ribcage and posteroanterior thoracic spine stiffness: an investigation of the normal response. *Man Ther.* 1999;4(3):157-162.
4. Maitland GD. *Vertebral Manipulation*. 4th ed. Butterworths; 1977.
5. Snodgrass S, Rivett D, Robertson V. Manual forces applied during posterior-to-anterior spinal mobilization: a review of the evidence. *J Manipulative Physiol Ther.* 2006;29(4):316-329.
6. Fritz JM, Whitman JM, Childs JD. Lumbar spine segmental mobility assessment: an examination of validity for determining intervention strategies in patients with low back pain. *Arch Phys Med Rehabil.* 2005;86(9):1745-1752.
7. Flynn T, Fritz J, Withman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine.* 2002;27(24):2835-2843.
8. Abbott H, Flynn T, Fritz J, Hing W, Reid D, Whithman J. Manual physical assessment of spinal segmental motion: intent and validity. *Man Ther.* 2009;14(1):36-44.
9. Latimer J, Lee M, Adams R, Moran C. An investigation of the relationship between low back pain and lumbar posteroanterior stiffness. *J Manipulative Physiol Ther.* 1996;19(9):587-591.
10. Brodeur R, DelRe L. Stiffness of the thoraco-lumbar spine for subjects with and without low back pain. *J Neuromusculoskelet Syst.* 1999;7(4):127-133.
11. Hodges PW, Eriksson AE, Shirley D, Gandevia SC. Intra-abdominal pressure increases stiffness of the lumbar spine. *J Biomech.* 2005;38(9):1873-1880.
12. Shirley D, Hodges PW, Eriksson AEM, Gandevia SC. Spinal stiffness changes throughout the respiratory cycle. *J Appl Physiol.* 2003;95(4):1467-1475.
13. Mootz R, Keating J, Kontz H, Milus T, Jacobs G. Intra- and inter observer reliability of passive motion palpation of the lumbar spine. *J Manipulative Physiol Ther.* 1989;12(6):440-445.
14. Maher C, Adams R. Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. *Phys Ther.* 1994;74(9):801-809.
15. Binkley J, Stratford P, Gill C. Interrater reliability of lumbar accessory motion mobility testing. *Phys Ther.* 1995;75(9):786-792.
16. Harms M, Bader D. Variability of forces applied by experienced therapists during spinal mobilization. *Clin Biomech (Bristol, Avon).* 1997;12(6):393-399.
17. Caling B, Lee M. Effect of direction of applied mobilization force on the posteroanterior response in the lumbar spine. *J Manipulative Physiol Ther.* 2001;24(2):71-78.
18. Nicholson L, Adams R, Maher C. Manual discrimination capability when only viscosity is varied in viscoelastic stiffness stimuli. *J Manipulative Physiol Ther.* 2003;26(6):365-373.
19. Maher C, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgements made using a reference-based protocol. *Phys Ther.* 1998;78(8):829-837.
20. Lee M, Svensson N. Measurement of stiffness during simulated spinal physiotherapy. *Clin Phys Physiol Meas.* 1990;11(3):201-207.
21. Latimer J, Goodsell M, Lee M, Maher C, Wilkinson B, Moran C. Evaluation of a new device for measuring responses to posteroanterior forces in a patient population, part 1: reliability testing. *Phys Ther.* 1996;76(2):158-165.
22. Owens E, DeVocht J, Wilder D, Gudavalli M, Meeker W. The reliability of a posterior-to-anterior spinal stiffness measuring system in a population of patients with low back pain. *J Manipulative Physiol Ther.* 2007;30(2):116-123.
23. Stanton T, Kawchuk G. Reliability of assisted indentation in measuring lumbar spinal stiffness. *Man Ther.* 2009;14(2):197-205.
24. Evans JM. Differential compliance measured by the function recording and analysis system in the assessment of vertebral subluxation. *J Vertebral Subluxation Res.* 1998;2(1):15-21.
25. Leach RA, Parker PL, Veal PS. PulStar differential compliance spinal instrument: a randomized interexaminer and intraexaminer reliability study. *J Manipulative Physiol Ther.* 2003;26(8):493-501.
26. Wong AYL, Kawchuk G, Parent E, Prasad N. Within- and between-day reliability of spinal stiffness measurements obtained using a computer controlled mechanical indenter in individuals with and without low back pain. *Man Ther.* 2013;18(5):395-402.
27. Larivière C, Ludvig D, Kearney R, Mecheri H, Caron J-M, Preuss R. Identification of intrinsic and reflexive contributions to low-back stiffness: medium-term reliability and construct validity. *J Biomech.* 2015;48(2):254-261.
28. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* 2010;19(4):539-549.
29. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21(4):651-657.
30. Tuttle N, Barrett R, Laakso L. Postero-anterior movements of the cervical spine: repeatability of force displacement curves. *Man Ther.* 2008;13(4):341-348.
31. Edmondston SJ, Allison GT, Gregg CD, Purden SM, Svansson GR, Watson AE. Effect of position on the posteroanterior stiffness of the lumbar spine. *Man Ther.* 1998;3(1):21-26.
32. Snodgrass SJ, Rhodes HR. Cervical spine posteroanterior stiffness differs with neck position. *J Electromyogr Kinesiol.* 2012;22(6):829-834.
33. De Vet HCW, Terwee CB, Mokkink LB, Knol DL, eds. *Measurement in Medicine: Practical Guides to Biostatistics and Epidemiology*. New York, NY: Cambridge University Press; 2011.
34. Terwee CB, Bot SDM, De Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
35. Vaillant M, Pickar JG, Kawchuk GN. Performance and reliability of a variable rate, force/displacement application system. *J Manipulative Physiol Ther.* 2010;33(8):585-593.
36. Snodgrass SJ, Rivett DA, Robertson VJ. Measuring the posteroanterior stiffness of the cervical spine. *Man Ther.* 2008;13(6):520-528.
37. Wang T, Huang J, Wang Y, Huang X, Zhang J. Development of the measurement system of the postero-to-anterior movements of the cervical spine. International conference on modelling, identification and control; Wuhan, China; 2012:1055-1060.
38. Nicholson L, Maher C, Adams R, Phan-Thien N. Stiffness properties of the human lumbar spine: a lumped parameter model. *Clin Biomech (Bristol, Avon).* 2001;16(4):285-292.
39. Kawchuk GN, Fauvel OR. Sources of variation in spinal indentation testing: indentation site relocation, intraabdominal pressure, subject movement, muscular response, and stiffness estimation. *J Manipulative Physiol Ther.* 2001;24(2):84-91.