

RESEARCH

Open Access

The bench-top accuracy of the VerteTrack spinal stiffness assessment device



Anika Young^{1*} , Michael S. Swain¹, Gregory N. Kawchuk², Arnold Y. L. Wong³ and Aron S. Downie¹

Abstract

Background: The assessment of spinal stiffness by manual palpation in clinical settings has demonstrated both poor accuracy and reliability. More recently, mechanical methods for assessment of spinal stiffness have demonstrated superior accuracy and reliability. However, mechanical methods of spinal stiffness assessment can be expensive, time consuming and/or unsuited to clinical practice. While a new device has been designed to address these issues (VerteTrack), its benchtop performance remains unknown.

Aim: To measure the bench-top performance of VerteTrack.

Methods: A series of laboratory-based experiments were conducted in February 2018 to investigate the accuracy (precision and bias) of load and displacement measurements obtained by VerteTrack and then were compared against an appropriate reference standard. Measurements of both multiple-level continuous assessment (multiple spinal levels measured), and single-level assessment (single spinal level measured) were performed on a viscoelastic foam medium (AIREX® balance beam, Switzerland) and the resulting stiffness calculated.

Results: VerteTrack demonstrated high precision at all loads and displacements. There was minimal systematic measurement bias identified for applied versus reference load (mean bias = -0.123 N; 95%CI -0.182 to 0.428 N, $p < .001$), and no systematic measurement bias for measured versus reference displacement (mean difference = 0.02 mm; 95%CI -0.09 to 0.14 mm, $p < .001$). The magnitude of stiffness obtained during multiple-level continuous assessment was on average 0.25 N/mm (2.79%) less than that for single-level assessment (95%CI -0.67 to 0.17 N/mm, $p < .001$).

Conclusions: VerteTrack demonstrated high accuracy (high precision, low bias) under bench-top conditions. The difference in stiffness found between multiple versus single spinal levels should be considered in the research context, but is unlikely to be clinically relevant. The results of this study demonstrate that VerteTrack may be suitable for both single and multi-level spinal stiffness measurements in-vivo.

Keywords: Spinal stiffness, Spinal stiffness assessment, Mechanical spinal stiffness device, VerteTrack and instrumented spinal stiffness measurements

* Correspondence: annie.young@mq.edu.au

¹Department of Chiropractic, Faculty of Science and Engineering, Macquarie University, Sydney, Australia

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Manual therapists’ commonly use manual spinal stiffness assessment (MSSA) to guide diagnosis and treatment decisions for patients with non-specific spinal pain [1, 2]. In MSSA, therapists apply a force to the spine by hand in a posterior to anterior direction then subjectively interpret the resultant resistance to displacement [3]. Although MSSA is an accessible option for clinical practice, the reliability and criterion validity of this method is poor [3–10]. Mechanical devices were created as an objective alternative to MSSA [11, 12]. Fig. 1 illustrates the evolution of spinal stiffness assessment devices that typically assess spinal stiffness at one segmental level per indentation (single-level mechanical assessment) [12–21] which can be time-consuming and requires large amounts of data transcription.

VerteTrack is a new device developed to address the limitations of existing mechanical devices that quantifies bulk measurement of spinal stiffness [22]. A novel feature of the VerteTrack is a pair of rolling indentation wheels that enables stiffness assessment of an entire spinal region thereby minimising assessment time [22]. The VerteTrack has recently demonstrated excellent within-session test-retest reliability (intraclass correlation coefficient $ICC_{3, k}$ 0.95–100) and good to excellent between-day reliability ($ICC_{3, k}$ 0.82–0.93) in the clinical setting [23], however, its accuracy is unknown. Benchtop performance of a test instrument can be evaluated through measurement of precision (random error) and bias (systematic error) of the system under test (Fig. 2)

[24, 25]. Therefore, this study aimed to measure the bench-top performance of the VerteTrack under both single-level and multiple-level continuous test conditions.

Methods

Study design, setting and equipment

This was a laboratory-based accuracy study conducted in February 2018 [24]. Experiments were designed and conducted following the International Organization for Standardization (ISO 5725-1) for the accuracy of measurement methods and results [24].

VerteTrack overview

The VerteTrack frame (width 1080 mm × height 1090 mm × length 1, 510 mm) suspends an aluminium gantry that supports movement of a rolling indenter head (RIH) in three axes: X-axis (longitudinal, superior-inferior), Y-axis (transverse, left-right) and Z-axis (vertical, posterior-anterior) via stepper motors (resolution = 0.007 mm, www.stepperonline.com, China) (Fig. 3). A string potentiometer (resolution = 0.020 mm, accuracy ±0.010 mm, TE Connectivity, USA) is used to record Z-axis displacement. A vertically-oriented laser assists the operator to align the RIH upon pre-determined anatomical landmarks (GLX Laser Site, Barska). During spinal stiffness assessment, the VerteTrack applies discrete loads via addition of weighted plates (“plates”) with a nominal mass of 1 kg each (RIH + k plates; k = 0, 6). These loads were selected as they represent loads that have previously been used in VerteTrack

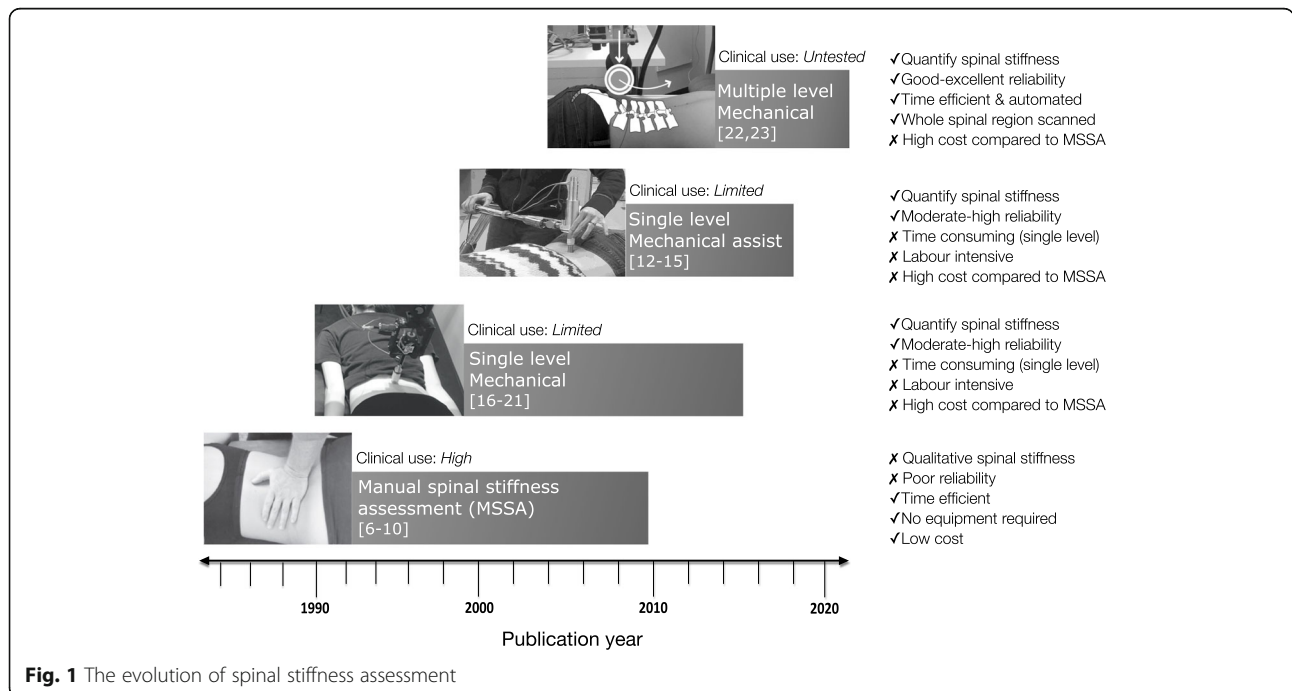
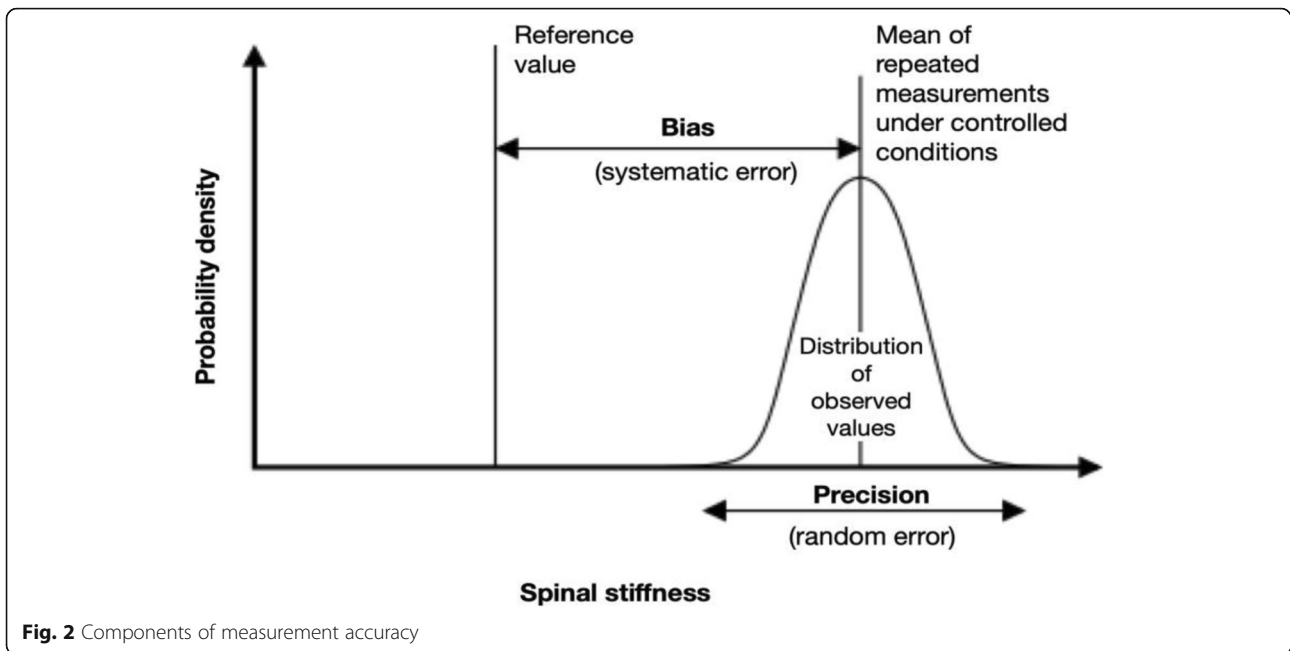


Fig. 1 The evolution of spinal stiffness assessment

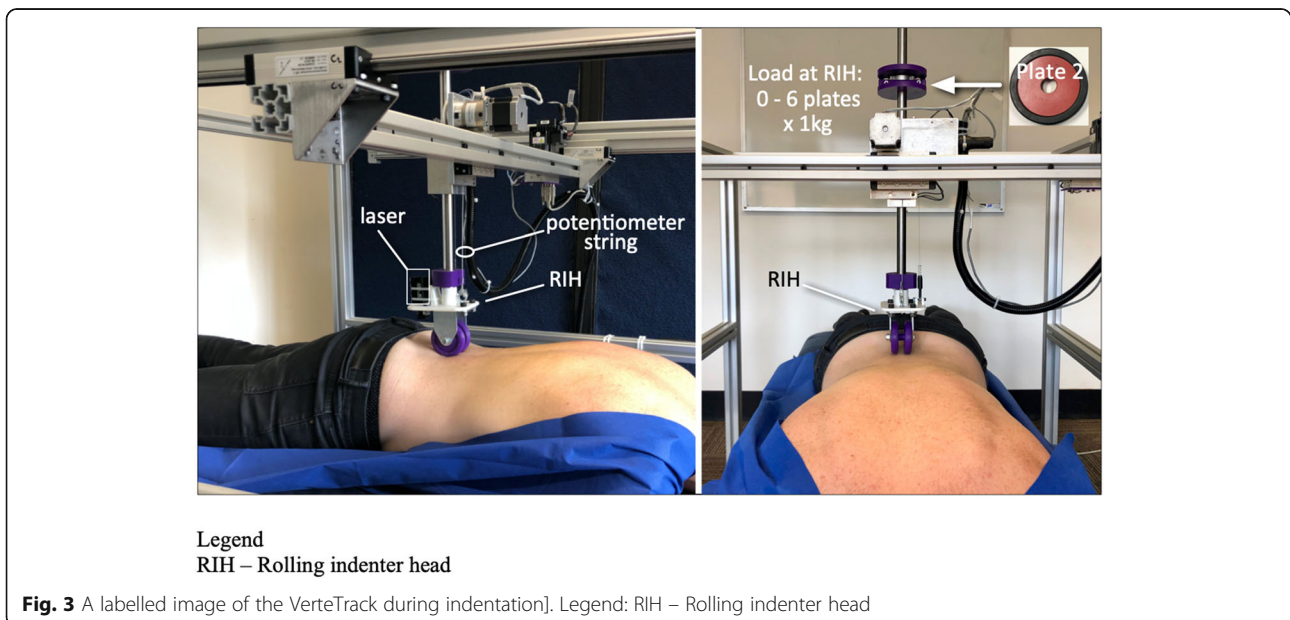


studies [22, 23, 26] and are comparable to loads applied in other mechanical indentation studies [14, 20]. Plates were numbered and always added in the same order for each indentation cycle. For more detail about the VerteTrack see Brown et al. 2017 [22].

Methods of indentation

The VerteTrack can perform two modes of indentation testing: single-level and multiple-level continuous indentation. Single-level indentation assesses a single spinal level and requires the operator to position the RIH

directly above the target tissue. Loads are then applied incrementally to the spine in a posterior to anterior direction with the resulting deformation of the spinal tissues recorded (Z-axis displacement). Multiple-level continuous indentation requires the operator to first identify the spinal trajectory that the RIH will travel within the horizontal (X-Y) plane. This is achieved by manually aligning each spinous process (determined by palpation or ultrasonography) with the RIH using the embedded laser pointer. The laser points are memorised by the device and then replayed to move the RIH



continuously along the same pre-defined trajectory for each successive load. The resolution of the RIH is identical to the resolution of stepper motors (0.007 mm).

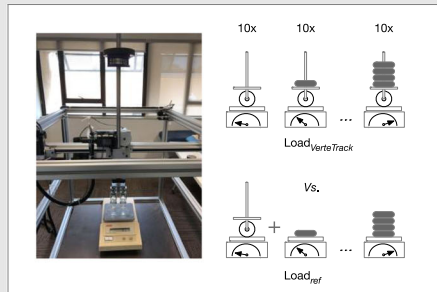
Load and displacement precision

Load precision (random error) of the VerteTrack was estimated by the coefficient of variation ($CV = SD / \text{load mean}$) over 10 repetitions for each load. The RIH was measured using recently calibrated digital scales (OHAUS, model TS4KD: Resolution 0.1 g, accuracy ± 0.07 g) (Fig. 4, panel a). Each plate was added to the RIH, then repeated up to a total of 5 plates. Loads were converted to Newtons (N) using mass (kg) x gravity (9.81 m/s^2). Displacement precision (z-axis, depth) of the

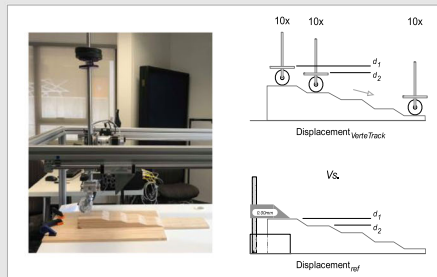
VerteTrack was also estimated using coefficient of variation over 10 repetitions at each of 6 discrete levels of the RIH on a custom-engineered wooden wedge to simulate tracking of a spinal sagittal curve (Fig. 4, panel b).

Load and displacement bias

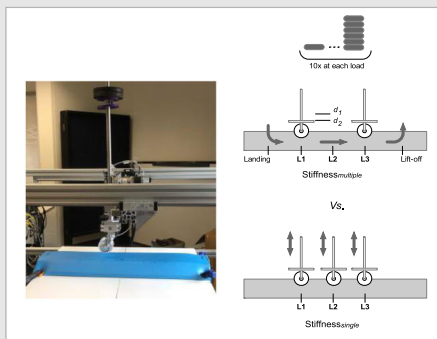
Load bias (systematic error) was estimated by comparing each load delivered through the VerteTrack against the same load externally. Mean load bias was estimated by calculating the differences between reference loads and loads measured by the VerteTrack, and the 95% confidence interval of the difference [25]. Reference loads were calculated by the addition of successive plates



Panel (a) Experiment one methodology: load applied by the VerteTrack ($Load_{VerteTrack}$) vs. reference load ($Load_{ref}$)



Panel (b) Experiment two methodology: displacement measured by the VerteTrack ($Displacement_{VerteTrack}$) vs. reference displacement ($Displacement_{ref}$)



Panel (c) Experiment three methodology: stiffness measurements using dynamic indentation ($Stiffness_{multiple}$) vs. static indentation ($Stiffness_{single}$)

Legend
 RIH – Rolling indenter head, d_0 – displacement 0, d_1 – displacement 1, d_2 – displacement 2, d_3 – displacement 3, d_4 – displacement 4

Fig. 4 a Experiment one methodology: load applied by the VerteTrack ($Load_{VerteTrack}$) vs. reference load ($Load_{ref}$). **b** Experiment two methodology: displacement measured by the VerteTrack ($Displacement_{VerteTrack}$) vs. reference displacement ($Displacement_{ref}$). **c** Experiment three methodology: stiffness measurements using multiple-level continuous indentation ($Stiffness_{multiple}$) vs. single-level indentation ($Stiffness_{single}$).
 Legend: RIH – Rolling indenter head, d_0 – displacement 0, d_1 – displacement 1, d_2 – displacement 2, d_3 – displacement 3, d_4 – displacement 4

placed directly upon the digital scale (i.e. not through the VerteTrack RIH) plus the load measured through the RIH alone. Each reference load (k plates; $k = 1, 5$) was measured ten times. Displacement bias was also estimated using the same method employed to determine load bias. Mean displacement bias was determined over 10 repetitions at each of 6 discrete levels as reported by the VerteTrack, compared to an external digital calliper (Wixey, WR200: Resolution = 0.05 mm, accuracy ± 0.025 mm) (Fig. 4, panel b).

Comparison of single-level and multiple-level continuous operation

A method-comparison experiment was conducted to evaluate the performance of VerteTrack for measurement of stiffness during multiple-level continuous and single-level (reference) modes of operation. Terminal stiffness values (i.e. the ratio of the maximum load to the maximum displacement) [26] were used in our analysis. The stiffness of a deformable foam test medium (AIREX® balance beam, Switzerland) was measured during both single-level and multiple-level continuous modes of operation to simulate measurement at a single vertebral level and across multiple vertebral levels respectively. The test medium was chosen to emulate the physiological stiffness encountered for the in vivo adult lumbar spine (range: 2–10 N/mm) [12, 20, 26]. Five equidistant locations (5 cm apart) were marked on the foam medium along a straight line (RIH landing, L1, L2, L3 and RIH lift-off) for stiffness assessment (Fig. 4, panel c).

Precision during single-level and multiple-level continuous indentation

Precision of the VerteTrack during measurement of stiffness on the test medium was estimated by the coefficient of variation ($CV = SD / \text{stiffness mean}$) over 300 trials for both single-level and multiple-level continuous indentation. Stiffness was measured during multiple-level continuous indentation ($\text{Stiffness}_{\text{multiple}}$) and single-level indentation ($\text{Stiffness}_{\text{single}}$) at three discrete locations (L1, L2, L3) on the medium. Incremental loads (plates) were added to the RIH in a predefined sequence (RIH + k ; $k = 1, 6$). Between each trial, 90 s elapsed to allow for any residual deformation to resolve. Between each cycle (six trials of increasing load), an additional 5 min elapsed to allow any residual deformation to resolve after the maximum load was applied to the medium. A total of ten cycles were performed.

Single-level versus multiple-level continuous indentation

Each trial for $\text{Stiffness}_{\text{multiple}}$ was compared to $\text{Stiffness}_{\text{single}}$ to quantify bias between indentation methodologies we calculated the stiffness differences and 95% confidence intervals of the difference. Bias calculation,

and a plot of raw stiffness data were conducted to assist interpretation. In addition, Lin's Concordance Correlation Coefficient (LinCCC, R_c) was reported for load and displacement. LinCCC tests both agreement and linearity [27]. The strength of agreement was graded as "almost perfect" ($R_c > 0.99$), "substantial" ($R_c > 0.95-0.99$), "moderate" ($R_c > 0.90-0.95$), or "poor" ($R_c < 0.90$) [28]. Alpha was set at 0.05 for all statistical significance tests of agreement.

Results

Load and displacement precision

Six measures relating to RIH loading ($\text{Load}_{\text{VerteTrack}}$) were obtained beginning with no load 16.557 N (95%CI: 16.470 to 16.645, RIH only) then increasing by the addition of single masses. The coefficient of variation (CV) ranged from 0.03 to 0.27% depending upon the applied load (Table 1).

Six discrete RIH displacements were then measured ($\text{Displacement}_{\text{VerteTrack}}$) beginning at a baseline value of 60.03 mm (95%CI: 60.01 to 60.05 mm, highest level) then increasing to 12.08 mm (95%CI: 12.00 to 12.16 mm, lowest level). The CV ranged from 0 to 0.32% depending upon the level of the wedge (Table 2).

Load and displacement bias

The calculated reference loads (Load_{ref}) ranged from 27.757 N (95%CI: 27.701 to 27.814 N, RIH + 1 plate) to 71.461 N (95%CI: 71.420 to 71.502 N, RIH + 5 plates). There was a statistically significant ($p < .001$) systematic mean bias for the VerteTrack load ($\text{Load}_{\text{VerteTrack}}$), compared to the reference load (Load_{ref}) of -0.123 N (95%CI: -0.182 to 0.428 N, $p < .001$) (supplementary figure 1, panel a). Lin's Concordance Correlation Coefficient showed almost perfect agreement ($R_c = 1.0$, 95% CI: 1.0 to 1.0) between $\text{Load}_{\text{VerteTrack}}$ and Load_{ref} (supplementary figure 2, panel a).

The reference displacement ($\text{Displacement}_{\text{ref}}$) as measured by the digital calliper ranged from 12.03 mm (95%CI: 11.98 to 12.08 mm) to 60.08 mm (95%CI: 60.02 to 60.13 mm). There was no statistically significant ($p = .001$) systematic bias for the VerteTrack displacement ($\text{Displacement}_{\text{VerteTrack}}$) as compared to the reference displacement ($\text{Displacement}_{\text{ref}}$) (mean difference = 0.02 mm, 95%CI: -0.09 to 0.14 mm, $p < .001$) (supplementary figure 1, panel b). Lin's Concordance Correlation Coefficient demonstrated an almost perfect agreement ($R_c = 1.0$, 95% CI: 1.0 to 1.0) between $\text{Displacement}_{\text{VerteTrack}}$ and $\text{Displacement}_{\text{ref}}$ (supplementary figure 2, panel b).

Stiffness measurements

Performance at different loads For the purposes of this study, we classified low load conditions as 27.757 N and 38.662 N (1 and 2 plates), and moderate to high loads as

Table 1 Precision of VerteTrack applied load

Indenter head loading	Load _{VerteTrack} ^a (N)	95%CI (N)	SD	CV
RIH only	16.557	16.470 to 16.645	0.045	0.27%
RIH + 1 plate	27.757	27.701 to 27.814	0.029	0.10%
RIH + 2 plates	38.662	38.589 to 38.735	0.037	0.10%
RIH + 3 plates	49.583	49.479 to 49.688	0.053	0.11%
RIH + 4 plates	60.687	60.592 to 60.783	0.049	0.08%
RIH + 5 plates	71.461	71.420 to 71.502	0.021	0.03%

CV Coefficient of variation, RIH Rolling indenter head, SD standard deviation

^aAverage of 10 measurements at each load. All loads measured with digital scale (OHAUS, model TS4KD: Resolution = 0.1 g, accuracy ± 0.07 g. Equivalent to resolution = 0.001 N, accuracy ± 0.0007 N)

49.583 N, 60.687 N and 71.461 N (3–5 plates). The terminal stiffness values for low load conditions ranged from 6.09 to 8.81 N/mm, and moderate-high load conditions yielded stiffness values ranging from 5.70 to 6.38 N/mm. Under low load conditions, the terminal stiffness grand mean was 7.43 N/mm. As for moderate to high load conditions, the terminal stiffness grand mean was 6.03 N/mm (Fig. 5). Figure 5 shows a graphical representation of the effects of low versus moderate-high loading on stiffness values.

Performance of single-level versus multiple-level continuous indentation modes To determine the precision of stiffness measured by the VerteTrack (Stiffness_{multiple} and Stiffness_{single}) at three locations on the AIREX balance beam, CV was calculated for each load (Table 3). The CV at each load for Stiffness_{single} ranged from 2.0 to 2.3% and Stiffness_{multiple} ranged from 1.4 to 3.2%.

Multiple-level continuous indentation (Stiffness_{multiple}) and single-level indentation (Stiffness_{single}) were compared at three discrete locations (L1, L2 and L3) on the AIREX balance beam. There was a negative systematic bias for Stiffness_{multiple}, compared to Stiffness_{single} of -0.25 N/mm (95%CI -0.67 to 0.17 , $p < 0.001$) (supplementary figure 1, panel c).

Discussion

This is the first mechanical spinal stiffness device to be evaluated for the bench-top performance (accuracy), which is essential to establish internal validity of the VerteTrack. Both single-level and multiple-level continuous indentation modes demonstrated high levels of precision and agreement, despite a small negative systematic bias for multiple-level continuous compared to single-level indentation (-0.25 N/mm, equivalent to 4% lower stiffness). It is unclear if this difference is clinically relevant as there is currently no published data to support a minimal clinically important difference (MCID) for the assessment of spinal stiffness, nor standards for different indentation modes [3, 11]. More broadly, mechanical devices must first be used to collect baseline spinal stiffness data in a human population in order to determine a MCID, while on the other hand, an MCID cannot be calculated without understanding the performance of a measurement as is described here. Such baseline data will allow for more robust conclusions regarding differences between single-level and multiple-level continuous indentation.

The plot of raw single-level and multiple-level continuous stiffness values demonstrated that higher stiffness values were obtained under low loads compared with moderate-high loads where stiffness remained at approximately 6 N/mm (Fig. 5). This is likely attributed to properties of the

Table 2 Precision of the VerteTrack RIH displacement

Wedge level	RIH displacement relative to table-top ^a (mm)	95%CI (mm)	SD	CV
d ₀ (landing point)	60.03	60.01 to 60.05	0.04	0.01%
d ₁ (highest level)	60.02	60.02 to 60.02	0.01	0.00%
d ₂	48.30	48.29 to 48.31	0.01	0.00%
d ₃	36.13	36.12 to 36.13	0.01	0.01%
d ₄	23.82	23.82 to 23.83	0.01	0.01%
d ₅ (lowest level)	12.08	12.00 to 12.16	0.14	0.32%

CV Coefficient of variation, RIH Rolling indenter head, SD standard deviation

^aAverage of 10 measurements at each displacement. All displacements were measured by the string potentiometer relative to the table-top (TE Connectivity, USA, Resolution = 0.020 mm, accuracy ± 0.010 mm)

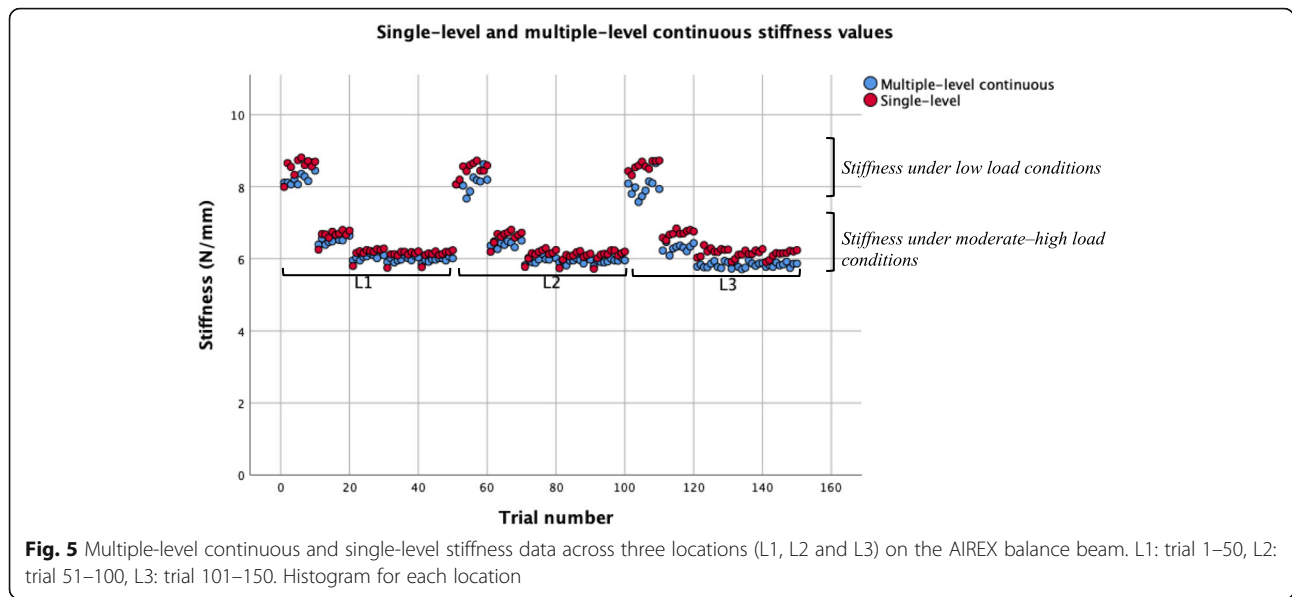


Fig. 5 Multiple-level continuous and single-level stiffness data across three locations (L1, L2 and L3) on the AIREX balance beam. L1: trial 1–50, L2: trial 51–100, L3: trial 101–150. Histogram for each location

viscoelastic foam medium. In vivo testing observed the inverse relationship between load and stiffness, that is a positive relationship between load and stiffness [23]. It would be advantageous to identify on a human population a specific load that yields the most useful spinal stiffness information. As reducing the number of loads would further reduce assessment time. In addition, results from in vivo testing suggest that the device provided reliable stiffness values, irrespective of load [23].

Limitations

This study was performed on a viscoelastic foam medium, without the presence of physiological properties known to influence spinal stiffness (such as breathing, spinal extensor muscle contraction and abdominal muscle contraction) [3, 11]. Also, it is unclear to what extent the observed phenomena can be attributed to the medium and whether a human population would emulate similar findings. To quantify bias, the level of agreement between multiple-level continuous stiffness measurements was compared to a reference standard. Single-level indentation was used as a proxy reference standard, given that it is the more established method of indentation reported in the literature. Unfortunately, there is no ‘gold standard’ to ascertain spinal stiffness in human participants.

Clinical utility of mechanical measurement of spinal stiffness

It is unclear whether the use of mechanical spinal stiffness measurement devices in a clinical setting would aid in the patient’s diagnosis, prognosis, treatment, or clinical outcomes. There are inconsistencies in the literature regarding the relationship between pain, disability and spinal stiffness, however emerging research into the subgrouping of patients into responders and non-responders to spinal manipulative therapy have yielded promising results [16]. Mechanical spinal stiffness devices currently have limited utility in clinical practice until further research can identify specific populations that may benefit from the assessment procedure.

Conclusion

The Vertetrack demonstrated good bench-top performance through high precision, linearity, and low systematic bias compared to reference values. When combined with recently published clinical reliability data, the VerteTrack demonstrates high levels of internal validity. The novel multiple-level continuous indentation mode offers potential for increased time efficiency in future clinical trials without compromise in stiffness measurement or patient comfort.

Table 3 Coefficient of variation for $Stiffness_{multiple}$ and $Stiffness_{single}$

VerteTrack mode	Coefficient of variation (CV) at each load					Mean
	RIH + 1	RIH + 2	RIH + 3	RIH + 4	RIH + 5	
Single-level	2.3%	2.2%	2.1%	2.0%	2.1%	2.2%
Multiple-level continuous	3.2%	2.2%	2.2%	1.7%	1.4%	2.2%

RIH Rolling indenter head

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12998-020-00331-8>.

Additional file 1: Figure S1. Panel a) The Bland-Altman plot demonstrates a statistically significant bias ($p < .001$) for loads delivered by the VerteTrack compared to the calibration sample (-0.123 N; 95%CI -0.182 to 0.428 N, $p < .001$). Open circles (50 data points) represent the magnitude of bias (N) = $\text{Load}_{\text{ref}} - \text{Load}_{\text{VerteTrack}}$. **Panel b)** The Bland-Altman plot demonstrates no statistically significant bias ($p = .001$) for displacement as measured by the VerteTrack compared to a digital calliper ($+0.02$ mm, 95% CI -0.09 to 0.14 mm, $p < .001$). Open circles (60 data points) represent the magnitude of bias (mm) = $\text{Displacement}_{\text{ref}} - \text{Displacement}_{\text{VerteTrack}}$. **Panel c)** The Bland-Altman plot demonstrates a statistically significant ($p < .001$) negative bias for multiple-level continuous vs. single-level stiffness, of 0.25 N/mm (95%CI -0.67 to 0.17 N/mm, $p < 0.001$). Open circles (150 data points) represent the magnitude of bias (N) = $\text{Stiffness}_{\text{multiple}} - \text{Stiffness}_{\text{single}}$. Legend: RIH – Rolling indenter head, d_0 – displacement 0, d_1 – displacement 1, d_2 – displacement 2, d_3 – displacement 3, d_4 – displacement 4. **Figure S2. Panel a)** Lin's Concordance Correlation Coefficient for VerteTrack load vs. the reference sample to demonstrate almost perfect agreement ($R_c = 1.0$, 95% CI 1.0 to 1.0). Open circles (50 data points) represent co-ordinates (Load_{ref} , $\text{Load}_{\text{VerteTrack}}$) at loads (RIH + k plates; $k = 1, 5$). **Panel b)** Lin's Concordance Correlation Coefficient for VerteTrack displacement vs. the digital calliper demonstrated an almost perfect agreement ($R_c = 1.0$, 95% CI 1.0 to 1.0). Open circles (60 data points) represent co-ordinates ($\text{Displacement}_{\text{ref}}$, $\text{Displacement}_{\text{VerteTrack}}$) for each wedge level (d_0 – d_5). Legend: RIH – Rolling indenter head, d_0 – displacement 0, d_1 – displacement 1, d_2 – displacement 2, d_3 – displacement 3, d_4 – displacement 4.

Abbreviations

MSSA: Manual spinal stiffness assessment; RIH: Rolling indenter head; ICC: Intraclass correlation coefficient; MCID: Minimally clinically important difference; CV: Coefficient of variation; CI: Confidence interval; N: Newtons; L1: Location 1; L2: Location 2; L3: Location 3; d_0 : Displacement 0; d_1 : Displacement 1; d_2 : Displacement 2; d_3 : Displacement 3; d_4 : Displacement 4

Acknowledgements

We would also like to acknowledge the Department of Chiropractic, Macquarie University, Sydney for supplying materials required for this project through a Masters of Research fund and for granting access to the VerteTrack.

Authors' contributions

ASD, MSS and GNK formulated the study design and methodology. AY and GNK collected the data. AY conducted the statistical analysis with the guidance of ASD, MSS and GNK. AY and ASD wrote and formatted the manuscript. ASD, MSS, GNK and AYLW were involved in reviewing and editing the manuscript. ASD constructed the evolution of spinal stiffness assessment (Fig. 1) and the graphic describing accuracy, precision and bias (Fig. 2). The author(s) read and approved the final manuscript

Funding

We would like to acknowledge the generous scholarship support (Masters of Research scholarship) received from the Australian Chiropractors' Association (ACA).

Availability of data and materials

The full data set used for the analysis is available on request from the corresponding author.

Ethics approval and consent to participate

Not applicable.

Consent for publication

The model in Fig. 3 consented for the photo to be used for the publication.

Competing interests

GK developed the VerteTrack device, however did not conduct the analysis of the data presented in this study. The authors state that there is no conflict of interest.

Author details

¹Department of Chiropractic, Faculty of Science and Engineering, Macquarie University, Sydney, Australia. ²Department of Physical Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, Canada. ³Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Hong Kong.

Received: 19 December 2019 Accepted: 18 June 2020

Published online: 18 August 2020

References

- Seffinger MA, Najm WI, Mishra SI, Adams A, Dickerson VM, Murphy LS, et al. Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. *Spine*. 2004;29(19):413–25.
- Haneline MT, Young M. A review of intraexaminer and interexaminer reliability of static spinal palpation: a literature synthesis. *J Manip Physiol Ther*. 2009;32(5):379–86.
- Snodgrass SJ, Haskins R, Rivett DA. A structured review of spinal stiffness as a kinesiological outcome of manipulation: its measurement and utility in diagnosis, prognosis and treatment decision-making. *J Electromyogr Kinesiol*. 2012;22(5):708–23.
- Jonsson A, Rasmussen-Barr E. Intra- and inter-rater reliability of movement and palpation tests in patients with neck pain: a systematic review. *Physiother Theory Pract*. 2018;34(3):165–80.
- Kawchuk GN, Miazga S, Pagé I, Swain M, De Carvalho D, Funabashi M, et al. Clinicians' ability to detect a palpable difference in spinal stiffness compared with a mechanical device. *J Manip Physiol Ther*. 2019;42(2):89–95.
- Jull G, Bullock M. A motion profile of the lumbar spine in an ageing population assessed by manual examination. *Physiother Pract*. 1987;3(2):70–81.
- Binkley J, Stratford PW, Gill C. Interrater reliability of lumbar accessory motion mobility testing. *Phys Ther*. 1995;75(9):786–92.
- Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. *Phys Ther*. 1998;78(8):829–37.
- Pool JJ, Hoving JL, de Vet HC, van Mameren H, Bouter LM. The interexaminer reproducibility of physical examination of the cervical spine. *J Manip Physiol Ther*. 2004;27(2):84–90.
- Piva SR, Erhard RE, Childs JD, Browder DA. Inter-tester reliability of passive intervertebral and active movements of the cervical spine. *Man Ther*. 2006; 11(4):321–30.
- Wong AYL, Kawchuk GN. The clinical value of assessing lumbar Posteroanterior segmental stiffness: a narrative review of manual and instrumented methods. *PM & R*. 2017;9(8):816–30.
- Stanton TR, Kawchuk GN. Reliability of assisted indentation in measuring lumbar spinal stiffness. *Man Ther*. 2009;14(2):197–205.
- Keller TS, Colloca CJ, Fuhr AW. 'In vivo' transient vibration assessment of the normal human thoracolumbar spine. *J Manip Physiol Ther*. 2000;23(8):521–30.
- Owens EF Jr, DeVocht JW, Wilder DG, Gudavalli MR, Meeker WC. The reliability of a posterior-to-anterior spinal stiffness measuring system in a population of patients with low back pain. *J Manip Physiol Ther*. 2007;30(2): 116–23.
- Björnsdóttir SV, Guðmundsson G, Auðunsson GA, Matthíasson J, Ragnarsdóttir M. Posterior-anterior (PA) pressure puffin for measuring and treating spinal stiffness: mechanism and repeatability. *Man Ther*. 2016;22:72–9.
- Wong AYL, Parent CE, Dhillon SS, Prasad NN, Kawchuk GN. Do participants with low back pain who respond to spinal manipulative therapy differ biomechanically from nonresponders, untreated controls or asymptomatic controls? *Spine*. 2015;40(17):1329–37.
- Lee M, Svensson NL. Measurement of stiffness during simulated spinal physiotherapy. *Clin Phys Physiol Meas*. 1990;11(3):201.
- Latimer J, Goodsell MM, Lee M, Maher CG, Wilkinson BN, Moran CC. Evaluation of a new device for measuring responses to posteroanterior forces in a patient population, part 1: reliability testing. *Phys Ther*. 1996; 76(2):158–65.
- Kawchuk GN, Fauvel OR. Sources of variation in spinal indentation testing: indentation site relocation, intraabdominal pressure, subject movement,

- muscular response, and stiffness estimation. *J Manip Physiol Ther.* 2001; 24(2):84–91.
20. Kumar S, Stoll S. Device, protocol and measurement of regional spinal stiffness. *J Electromyogr Kinesiol.* 2011;21(3):458–65.
 21. 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):1745–52.
 22. Brown BT, Blacke A, Carroll V, Graham PL, Kawchuk G, Downie A, et al. The comfort and safety of a novel rolling mechanical indentation device for the measurement of lumbar trunk stiffness in young adults. *Chiropr Manual Ther.* 2017;25:1.
 23. Hadizadeh M, Kawchuk GN, Parent E. Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants. *BMC Musculoskelet Disord.* 2019;20:1.
 24. ISO. 5725–1. Accuracy (trueness and precision) of measurement methods and results-part 1: general principles and definitions. 1st ed. Switzerland: International Organization for Standardization Geneva; 1998.
 25. Menditto A, Patriarca M, Magnusson B. Understanding the meaning of accuracy, trueness and precision. *Accred Qual Assur.* 2007;12(1):45–7.
 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. St-Pierre N. Validating mathematical models of biological systems: application of the concordance correlation coefficient. Sensitivity analysis of model output Los Alamos National Laboratory; 2005. p. 341–50.
 28. McBride G. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. NIWA client report: HAM2005–062; 2005.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

