Do physiological limits of human palpation prevent clinicians from appreciating differences in spinal stiffness?

Gregory Neil Kawchuk¹*, Stephen Miazga¹, Isabelle Pagé¹,², Michael Swain³,⁴, Diana De Carvalho⁵, Martha Funabashi¹, Alexander Breen⁶, Arnold YL Wong⁷

¹ Department of Physical Therapy, University of Alberta, Edmonton, AB, Canada
² Department of Anatomy, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada
³ Musculoskeletal Health Sydney, School of Public Health, Sydney Medical School, University of Sydney, Australia.
⁴ Department of Chiropractic, Faculty of Science and Engineering, Macquarie University, Australia
⁵ Faculty of Medicine, Memorial University of Newfoundland, St. John’s, NL, Canada
⁶ Institute for Musculoskeletal Research and Clinical Implementation, Anglo-European College of Chiropractic, Bournemouth, UK
⁷ Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong SAR, China

*Corresponding author

E-mail: greg.kawchuk@ualberta.ca (GNK)
Abstract

Objective: To determine if the physiological limits of human palpation influence traditional examination procedures.

Methods: In this study, the threshold at which a change in spinal stiffness was detected was quantified in 12 experienced clinicians by changing the differential stiffness in two inflatable targets until the clinician could no longer identify which was stiffer. In the second part of the study, clinicians were then asked to palpate pre-identified pairs of vertebrae in an asymptomatic volunteer and to identify the stiffer of the pair (T7&L3, T7&L4, L3&L4) while the biomechanical stiffness of each vertebral pair was quantified objectively by a validated instrument.

Results: The mean stiffness detection threshold for the clinicians was 8%. Objective measurement of the stiffness differential between vertebral pairs was 30% for T7* & L3 and 20% for T7* & L4 and 10% for L3* & L4 (*denotes the stiffer of the pair). Ten of 12 clinicians correctly identified T7 as more stiff when compared to L3 and T7 as more stiff than L4. Alternatively, when the differential vertebral pair stiffness was similar to the stiffness detection threshold (~8%), clinicians were less successful in identifying the stiffer vertebra of the pair; 4/12 clinicians correctly identified L3 as being more stiff compared to L4.

Conclusion: These results suggest that the physiological limits of human palpation may limit the ability of clinicians to identify small alterations in spine stiffness.
Introduction

Low back pain (LBP) is a very common condition with an estimated lifetime prevalence as high as 84% worldwide.\textsuperscript{1,2} Characterized by recurrent or intermittent episodes with a third of cases becoming chronic,\textsuperscript{3,4} LBP is the leading cause of global disability.\textsuperscript{5} Presently, the cause of LBP is notoriously difficult to identify in any given person, often resulting in inefficient approaches toward prevention, diagnosis and treatment. Consequently, LBP is responsible for substantial societal burdens in terms of morbidity, disability and cost.\textsuperscript{6–8}

Despite these challenges, a contemporary theoretical framework exists with which to contextualize back pain with the aim of reducing these burdens. This framework, the biopsychosocial model, comprises three domains (biological, psychological and social).\textsuperscript{9,10} While there are many well-developed tools that measure self-reported psychological and social domains of LBP, such as the Start Back Screening Tool,\textsuperscript{10} the Tampa Scale for Kinesiophobia\textsuperscript{11} and the Fear-Avoidance Beliefs Questionnaire\textsuperscript{12}, the same cannot be said for tools that measure physical aspects of the biological construct. In the same way that the absence of objective tests of mechanical heart function would negatively impact cardiac care, the absence of objective physical measures that quantify clinically meaningful mechanical function could limit, or possibly bias, our current understanding of LBP within the biopsychosocial model.

Toward this, our group has recently demonstrated that in people with LBP, those who report clinical improvement following conservative care, also experience immediate and significant changes in three objective biomechanical measures: spinal stiffness, dynamic paraspinal muscle thickness and disc perfusion.\textsuperscript{13} Those who report no improvement with the same
intervention do not experience these biomechanical changes. These findings suggest that there may be specific phenotypes of LBP.

This observation raises the question of whether clinicians would have the ability to detect stiffness changes of this magnitude using traditional examination procedures such as manual palpation. Although palpation remains a widely taught skill to evaluate spinal stiffness and guide conservative treatment plans\textsuperscript{14,15} previous research suggests that palpation has poor intra- and inter-rater reliability.\textsuperscript{16–18} In addition, palpation is a sense that has not yet benefited from technology in the same way that the stethoscope has improved auscultation or optics have improved vision. Therefore, there is a possibility that technological advances that improve the human sense of palpation may outperform traditional techniques of manual palpation.

Given this small, but significant magnitude of change in spinal stiffness quantified in our previous research,\textsuperscript{13} we asked the following question: “Do physiologic limitations in the ability of human touch prevent clinicians from detecting clinically important changes in tissue stiffness?” The objectives of the current study were to quantify the threshold at which clinicians can detect a difference in spinal stiffness via palpation, and then determine if this detection threshold would impact a clinician’s ability to identify changes in spinal stiffness as measured by an objective instrument. As such, our hypothesis was that a clinician’s ability to identify the stiffer of two vertebrae through manual palpation would decrease as the relative stiffness between the vertebrae approached the clinician’s physiological limit of stiffness detection.

**Materials and methods**

**Participants**
Clinicians from the greater Edmonton region (Alberta, Canada) were recruited for this controlled laboratory study. Specifically, physical therapists from the Edmonton region were recruited directly through email using a list provided by Physiotherapy Alberta College + Association, and chiropractors were recruited by email through the Chiropractic Society. No eligible clinicians were refused from participating. This study received ethics approval from the University of Alberta Human Research Ethics Review Board (Protocol number Pro00041451) and all participants signed an informed consent prior to testing.

Detection of stiffness threshold

To determine the detection threshold of palpated stiffness, two inflatable palpation targets were used. Stiffness of these targets could be controlled remotely by two independent hand pumps with pressure relief valves (Fig 1). The same investigator operated this equipment throughout the trial to achieve the desired pressure in each target by monitoring digital pressure readouts. At all times, the investigator's actions and the digital readouts were blocked from the participants view (Fig 1). The differential stiffness of these targets was altered using the staircase methodology. A detailed description of the staircase method can be found elsewhere.19 Briefly, this method is an adaptive psychometric technique that can be used to determine sensation thresholds perceived by participants by presenting a series of stimuli in a pattern of varying intensities ranging from large (obvious) to small.19 Digital pressure gauges were used to record the pressure in each bladder before the clinician began testing any specific combination of pressures. The pressure differential was then reduced in a step-like manner until it was too small for the participant to judge without making a mistake. At this point, the pattern was reversed and intensities increased again to the last identifiable level and then decreased again from that point. This pattern was repeated until the minimally detectable threshold was identified. All clinicians were given standardized instruction at the start of the study. Specifically,
they were instructed to apply posteroanterior forces vertically to the inflatable targets and the spine during stiffness assessments. The clinicians wore opaque goggles to block their vision during all palpation tasks and they were instructed to evaluate the spinal stiffness only (Fig 1). Clinicians were instructed to palpate the targets using the pisiform area of their palm placed overtop of a rigid metal disc placed in the centre of the inflatable target. The disc was of the same diameter of the instrument used to measure spinal stiffness (below).

Fig 1: Participant conducting the staircase protocol with an examiner recording the pressure from two inflatable bags connected to digital pressure meters

**Vertebral pair testing**

Clinicians were then asked to palpate pre-identified pairs of vertebrae in prone asymptomatic volunteers and identify the stiffer of the pair (T7&L3, T7&L4, L3&L4). These locations were marked in advance by an expert clinician using ultrasonography (Fig 2) and did not change over the course of the study. Briefly, this assessment involves the application of a manual posteroanterior force to a specific body landmark (e.g. spinous process) or to a spinal region in general (e.g. lumbar) by a clinician and then the clinician perceives the corresponding spinal movement or stiffness. As with testing on the inflatable targets, palpation consisted of the clinician using the pisiform area of their palm to apply direct pressure to the spinous process of pre-identified vertebra within a time limit of five minutes or less. This was done while the human volunteer was asked to fully exhale without exertion. While there are many spinal motion palpation techniques, spinal stiffness assessment performed in this way is a widely used skill among manual therapy professionals. Although we did not time the clinician’s palpation of the subject, each took approximately less than 10 seconds in making their assessment.
Fig 2: Skin markings to identify vertebral pair targets at the T7, L3 and L4 levels. Marks are to identify the vertebral segment without identifying the spinous process itself.

Following palpation by the clinician, the biomechanical stiffness of each vertebral pair in the asymptomatic volunteer was quantified with a validated instrument (Fig 3). This device has demonstrated excellent within-day reliability and has been used in multiple studies to quantify spinal stiffness in people with and without LBP.\textsuperscript{13,23–25} In brief, the device consists of a computer-controlled indentation probe that can be moved vertically at a predetermined speed of 2 mm/s. The probe is attached in-series with a compression-tension load cell transducer (Entran, Fairfield, NJ, USA) and a rotary encoder (Dual Motion Motor, Haydon Kerk Motion Solutions Inc., Waterbury, CT, USA) to measure indentation forces and the corresponding displacements, respectively. These measures are obtained from placing the probe on the skin surface above the spinous process of the lumbar vertebrae of interest while the subject is asked to exhale completely as in the manual palpation procedure. A customized programming language (Labview, National Instruments Inc., Austin, TX, USA) was used to control data collection from the above sensors at 200 Hz. The indentation process could be stopped at any time via hardware or software emergency switches.

The determination of a correct palpation (yes/no) was made when the clinician’s estimate of which vertebra is stiffer agreed (or disagreed) with the results obtained from the indentation instrument.

Fig 3: Indentation equipment used to obtain spinal stiffness (photograph taken without skin markings).
Data processing and analysis

The stiffness detection threshold was calculated as the average minimal difference between stiffness values obtained by palpating the inflatable objects that were correctly detected by the participants and then expressed as a percentage of the maximal detectable difference (i.e. Weber fraction). For the vertebral pair testing, stiffness (N/mm) was calculated as the maximally applied force (60 N) divided by the resulting displacement of the blunt indentation probe (mm). The percentage difference was calculated by considering the highest stiffness of the pair as 100%.

Results

Although 15 clinicians were recruited, three did not attend the testing session. Therefore, data from 12 clinicians were included in the analysis.

Table 1 presents data from the staircase stiffness threshold detection as well as from the vertebral pair testing. The staircase test revealed that the mean stiffness detection threshold for all participants was 8% (± 3%). Objective measurement of the stiffness differential between vertebral pairs was 30% for the T7* - L3 pair, 20% for the T7* - L4 pair and 10% for the L3* - L4 pair (*demotes the stiffer of the pair). While T7 was the stiffer vertebra when compared to L3 and L4, L3 was stiffer compared to L4 vertebra.
Table 1: Threshold detection results from manual palpation of vertebral pairings. The determination of correct (yes/no) is made when the clinician’s estimate of which vertebra is stiffer agrees (or disagrees) with the results obtained from the indentation instrument.

<table>
<thead>
<tr>
<th>Staircase Results</th>
<th>Vertebral Stiffness Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>Threshold (%)</td>
</tr>
<tr>
<td>1</td>
<td>5.2</td>
</tr>
<tr>
<td>2</td>
<td>7.5</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
</tr>
<tr>
<td>4</td>
<td>5.5</td>
</tr>
<tr>
<td>5</td>
<td>6.1</td>
</tr>
<tr>
<td>6</td>
<td>6.8</td>
</tr>
<tr>
<td>7</td>
<td>7.6</td>
</tr>
<tr>
<td>8</td>
<td>8.0</td>
</tr>
<tr>
<td>9</td>
<td>9.1</td>
</tr>
<tr>
<td>10</td>
<td>9.5</td>
</tr>
<tr>
<td>11</td>
<td>11.2</td>
</tr>
<tr>
<td>12</td>
<td>13.1</td>
</tr>
</tbody>
</table>
For the vertebral pair testing, when the stiffness difference within each vertebral pair exceeded the detection threshold of the clinician (~8%), clinicians had greater success in identifying the stiffer vertebra of the pair. Specifically, 10/12 clinicians correctly identified T7 as more stiff when compared to L3. The same number of clinicians (10/12) also correctly identified T7 as more stiff than L4. Alternatively, when the differential vertebral pair stiffness was similar to the stiffness detection threshold, clinicians had less success in identifying the stiffer vertebra of the pair. In this circumstance, only 4/12 clinicians correctly identified L3 as being more stiff compared to L4. Of note, the stiffness difference between L3 and L4 (~10%) observed in the current study was similar to the value observed in prior studies whose subjects reported improvement following spinal manipulation (~10%).

**Discussion**

This study found that clinicians are able to accurately detect large stiffness differences by using manual palpation, but were less able to accurately detect small differences in stiffness. Although previous studies have investigated both palpation and spinal stiffness separately, this is the first study to combine both in terms of a known threshold of detectable stiffness change.

Interestingly, when clinicians palpated the vertebral pair with the smallest stiffness differential as measured by instrumentation (L3 and L4), their responses were not distributed randomly.
between the two (i.e. 50% selecting L3 and 50% selecting L4). Further, because stiffness detection thresholds were determined individually, there were clinicians whose threshold were above and below the average value of 8%. These data suggest that when clinicians reach the limits of their ability to provide a confident answer regarding differential stiffness, other factors not measured here may come into play (e.g. guessing, ineffective incorporation of other sense on other senses).

Related studies
Multiple studies and reviews have accentuated poor intra- and inter-rater reliability of spinal motion palpation when assessing spinal stiffness manually.16,27–30 This low reliability may be attributed to the fact that many factors can affect spinal stiffness perceived by clinicians21 including vision31, technique31, intent32 and training.33 Although manual PA spinal stiffness assessment has face validity in assessing spinal biomechanical changes following manual therapy34, the sensitivity of clinicians in detecting differences/changes in spinal stiffness through manual palpation techniques has been unclear. While prior estimates suggest that clinicians can discriminate differences in stiffness as low as 11%,31,35 this threshold does nothing to suggest the magnitude of stiffness changes in pathological tissues which in theory, may occur below or above the detection threshold of clinicians.

Previously, Koppenhaver et al.36 demonstrated that estimates of spinal stiffness obtained by manual palpation were not correlated to measures obtained from an instrument. Specifically, their study compared manual stiffness assessments from the spinous processes of L1-L5 to a single measurement of stiffness at L3. One strength of the current study is that manual and instrumented assessments of stiffness were always taken from the same locations. In addition, the current study created a range of stiffness gradients which demonstrated that clinicians can
indeed discriminate changes in stiffness, but this discrimination is modulated by the magnitude of the change.

**Clinical advances as a result of assistive technologies**

Significant advances in our understanding of health conditions and their treatment occur when new technology augments a basic sense. From a rolled up tube of paper creating the first stethoscope, to modern day microphones, improving a clinician’s sense of hearing has revealed new pathologies, improved diagnosis and created more timely treatment. Similar advances have also occurred by augmenting vision with loupes and microscopes.

In contrast to hearing and sight, the sense of touch remains relatively unassisted by technology. This is likely the result of being unable to increase the sense of touch directly as would be the case for amplifying auditory signals in the ear or magnifying images provided to the eye. Therefore, sensory augmentation must be achieved in a different way. Technologies to evaluate aspects of touch that cannot be enhanced directly within the human experience do exist. An example of this would be thermometers which work not by improving the ability of humans to perceive temperature, but by providing a surrogate technology with the capability of visually displaying changes in temperature that are not perceptible with human touch.

In addition to temperature, humans can also feel changes in texture such as a nano-size wrinkles on a smooth surface. This is an extension of detecting changes in an object's stiffness, a regular human experience. Clinically, changes in stiffness can be linked to pathological alterations in tissue as in the case in glaucoma. However, where once clinicians evaluated the eye by gently pressing their fingers into the closed eyelid of their patients, the advent of tonometers, sophisticated technology capable of measuring ocular pressure, has
again led to important advances in understanding eye disease and its treatment.\textsuperscript{41,42} Similarly, technologies such as elastography are helping clinicians visualize changes in prostate stiffness related to cancer.\textsuperscript{43,44}

While these technologies have created important advancements in understanding diseases of relatively homogeneous tissues (eye, prostate), heterogeneous tissues pose a significant challenge in determining stiffness related to pathology, disease or injury. The spine is such an example where changes in its stiffness may reflect specific conditions.\textsuperscript{45,46} Unfortunately, there are few, if any, technologies that can quantify these changes to better aid clinicians. In addition, we do not know if the threshold of human stiffness detection is sufficient for detection of clinically important changes in spinal stiffness.

**Relevance of the findings to clinical practice**

This study revealed that while a majority of the clinicians could identify large differences in vertebral pair stiffness, the study sample had difficulty in successfully detecting differences in vertebral pairs when the stiffness gradient approximated the stiffness detected threshold. Since this threshold falls in the range of stiffness that has been observed to occur following clinical interventions such as spinal manipulation,\textsuperscript{13} it appears that practitioners may have difficulty in detecting these post-intervention changes in stiffness of magnitudes similar to the detection threshold.

If there is an inability to detect a small magnitude change in stiffness that is clinically significant, it is possible to imagine a scenario where a practitioner may inappropriately change, or maintain a course of therapy based on faulty sensory information. In this case, the development of technologies to make this information available in the decision-making processes, may improve outcomes for musculoskeletal conditions.
Limitations

As can be the case in the Weber faction, we report our results as percentages so that data can be compared between circumstances such as differences in the surfaces from which stiffness data was collected (air bladder on a rigid surface, human subject lying prone on a padded plinth) and the surface area of palpation (the area of the clinician/subject interface and the area of the indenter). As a result, comparison of absolute stiffness values is not possible. Although the staircase method is very efficient in identifying sensory thresholds, this method relies on the reporting of the participant.19 The possibility that a participant has manipulated the results purposefully or through unintentionally guessing cannot be ruled out. Although our sample was small and the generalizability of these results is limited, physiological thresholds of human senses tends to be similar.47 In addition, we designed the study to be representative of what clinicians do in practice which meant giving them a sufficient amount of time to make their palpation determination rather than giving them multiple attempts. Therefore, we cannot know from this work if multiple palpation attempts would have improved clinician performance but speculate it would not as the physiological limitations of the clinician would not have changed with additional attempts. Last, although we did not measure spinal stiffness in the subject following palpation, prior reliability testing of the device shows that in symptomatic subjects, this measure is stable within the testing period.24

Conclusions

While clinicians can detect large stiffness changes between spinal regions, they are less able to detect smaller changes in stiffness that are similar in magnitude to previously published change changes in stiffness following successful interventions. These results allow us to speculate that the physiological limits of human touch may act to restrict the ability of clinicians to appreciate
changes in spinal stiffness as they approach the physiological limits of palpation which may influence clinical decisions although we do not yet know the clinical meaning of these small changes or if they are common. Given that instrumentation is used commonly to assist clinical hearing (stethoscope) and sight (loupes, ultraviolet light), assistive technology to aid a clinician’s sense of touch (stiffness testing) may be helpful in better understanding how biological/biomechanical changes may be related to back pain within the biopsychosocial model.

Acknowledgements

The investigators would like to thank the physical therapists and chiropractors from who volunteered their time for this study.

References


