A simulated passive intervertebral motion task: observations of performance in a cross-sectional study

Charles R. Hazle Jr, Arthur J. Nitz

University of Kentucky, Hazard, KY, USA

Introduction: Passive motion palpation is an integral component in examination, diagnosis or classification, and treatment of persons with mechanical disorders of the cervical spine. If the magnitude of force application during passive movement assessment is associated with greater palpatory accuracy has not been established.

Methods: This investigation used a novel mechanical model as a basis for assessing the palpatory force of students and clinicians. The model included multiple palpable resistance and displacement levels similar to that observed in humans. The ability of the subjects to discriminate the various levels of resistance and displacement offered by the model was concurrently measured.

Results: Large variability occurred in the amount of force applied by the subjects in completing the palpatory examination. The data indicated no major differences in palpatory accuracy across the student and clinician groups with different training and experience levels beyond basic competency. Those subjects applying less force in the palpatory exam demonstrated greater accuracy of palpatory assessment with one measure.

Discussion: The data indicate training and experience had minimal relationship to palpatory interpretation precision beyond the basic level and individual natural discriminatory ability may be a factor in accuracy of palpatory skill. The results demonstrate remarkable inconsistency in palpatory force among examiners and suggest that palpatory accuracy may be related to less force application.

Keywords: Passive intervertebral motion, Palpatory examination, Cervical spine mechanical model, Joint play, Cervical spine examination

Introduction

Evidence has accumulated over the past decade to indicate benefit from manual therapy interventions in treating patients with mechanical neck pain disorders and associated clinical syndromes, particularly when combined with exercise.¹⁻⁵ Criteria have also been identified as predictors of favorable responses for the use of manual therapy.⁶⁻⁸ To target where to apply these techniques has not been clarified. Simple palpation for detection of tenderness in the cervical spine to localize the origin of pathology has yielded mixed results.^{9,10} Limited and conflicting evidence exists as to the validity of passive intervertebral motion (PIVM) in identifying symptomatic segments in the cervical spines of subjects with neck pain. Manual therapy practitioners have frequently cited the study by Jull *et al.*¹¹ as a validation of PIVM examination and its value in guiding the course of clinical care. In a more eloquently designed subsequent study that has

been vigorously defended,¹² King *et al.*¹³ challenged the presumed value of the cervical PIVM examination. Both studies relied solely on the skills and abilities of a single examiner.

The application of segmentally specific manual therapy has been suggested to be insignificant.¹⁴⁻¹⁶ The results of these studies suggest that specific manual palpation may not be required to identify mechanical lesions for focal delivery of effective mobilization or manipulation techniques. Motion irregularities are thought to be associated with mechanical pain syndromes on a consistent basis and identification of these lesions is considered an aim of examination.¹⁷⁻¹⁹ For clinicians to render optimal care, including the appropriate inclusion or exclusion of manual therapy techniques, precision as to where to deliver those techniques has historically been accepted clinical reasoning. Thus, despite several studies, little is understood with certainty about the performance of PIVM and the ability of clinicians to detect changes in tissue resistance, potentially associated with mechanical neck pain syndromes.

Correspondence to: Charles R. Hazle Jr, Center for Excellence in Rural Health, University of Kentucky, 750 Morton Boulevard, Hazard, KY 41701, USA. Email: crhazl00@uky.edu

A systematic review stated palpatory diagnostic tests are not reliable, simultaneously noting the faulty methodology in many of the studies.²⁰ Their collected findings suggest the reliability of examiners' judgments were only marginally above chance agreement. The authors considered all studies collectively, regardless of significant differences in techniques and study design.

A small number of studies have suggested higher levels of agreement. In a particularly noteworthy result, blinded examiners agreed on palpatory results of subjects with congenital block vertebrae in the cervical spine, analogous to detection of hypomobile segments, with kappa values of 0.46–0.76.²¹ Similarly, kappa values of 0.57, 0.60, and 0.84 were achieved among the three combinations of examiners in another carefully controlled study.²²

Addressing many of the methodological issues in these studies with PIVM in subjects being palpated has proven difficult because of a number of practical matters. As an example, the variance among persons with symptomatic and asymptomatic cervical spines has been an issue pertaining to validity and reliability in the passive motion palpation studies. Bias in analysis of reliability is readily introduced when subjects of PIVM palpation are known to be symptomatic or asymptomatic as occurred in multiple investigations.^{23–28} Knowledge of patient history, specifically pertaining to the cervical spine, has been found to affect the prevalence of positive findings.²⁹ The categorization of palpatory findings is also a consideration in these studies. A small number of studies used percent agreement among examiners, but failed to account for chance agreement.^{30–32}

The majority of the studies used a simple Cohen's kappa statistic in assessing inter-examiner reliability, but often did so in a manner to dichotomize the examiners assessment of the joint mobility into two rather stark categories of either pathomechanically limited or not limited.^{17,18,24,26–28,33,34} This type of designation essentially negates the ranges of mobility which naturally occur in the joints. Only two studies used a weighted kappa statistic,^{23,25} perhaps more consistent with gradations of mobility as might actually occur in the population.

Another methodological issue to be noted is that all palpatory PIVM techniques were grouped together in the aforementioned systematic review. For example, side-glide palpatory technique data were grouped together with simple posterior-to-anterior (PA) pressure technique data, disregarding the potential for one method yielding better outcomes than another.

Interpreting normal as opposed to pathological movement in the palpatory exam is problematic as substantial normative data do not exist. Objective measurement methods for the palpatory process are not available; thus, the practice of performing the

exam and subsequent instruction to others is necessarily subjective. Similarly, the act of an examiner completing PIVMs on an individual's cervical spine may change the mechanical properties of the tissues being examined. In effect, mobilization of suspect segments has been completed by the first PIVM examiner, providing the subsequent examiners with a different experience. As a result, direct sequential comparisons between examiners may be affected.

While the forces used in PIVM have been studied by several investigators, an association of those levels of force with palpatory accuracy has not been evaluated.^{35–40} The current study sought to explore the performance reliability and interpretive accuracy of examiners with multiple levels of education and experience in a frequently used examination skill, including use of force measurements. A mechanical model was constructed to allow comparison between examiners without substantive alteration in the interpretive experience as a result of the examination process. The effect of possible change in the subjects being palpated due to the examination process was eliminated, addressing one issue in many prior studies. Construction of the model included properties of the human cervical spine not previously reported in the literature in other simulation devices. To date, simulation devices have used elastic resistance, which does not replicate the viscoelastic properties of human tissue. The model for this study incorporated an element of viscoelastic resistance, more closely simulating human tissue resistance. This model was then used to investigate the accuracy of examiners across four cohorts along with analyzing their performance characteristics. Experience and training levels were assessed for their relative accuracies, addressing another convention of palpatory skill. Subsequently, the results were used to consider the relationship of the examiners' interpretation accuracy and their performance attributes.

Methods

To minimize many of the methodological issues in prior passive motion studies, this investigation used a mechanical model of the cervical spine. Three key biomimetic properties were incorporated into the model. Viscoelastic resistance was accomplished by using a polymer gel (Perma-GelTM, Permal-Gel Inc, OR, Albany, USA) with a modulus of elasticity consistent with human tissue.^{41,42} Three-point bending was included to allow greater motion in the middle of the unit than at either end, simulating greater mobility potential in the middle of the cervical spine than at the transitional segments.^{43,44} Perhaps most importantly, the force–displacement curves observed by Tuttle *et al.*⁴⁵ served as the basis for the range of curves available within the model (Fig.1).

Figure 1 Range of force-displacement curves measured with mechanical model and curve measured by Tuttle *et al.*⁴⁵

For the purposes of testing, the design of the model also allowed for interchangeable modules permitting the capability to vary the resistance and allowable motion at each simulated vertebra over the course of multiple trials during testing with subjects. The design of the model is demonstrated in Fig. 2.

A total of 66 subjects participated in this study, consisting of physical therapy students and clinicians in four cohorts: novice students, experienced students, novice clinicians, and experienced clinicians.

The novice students group (10 females and 6 males) was comprised of physical therapy students in their first semester of study of a single, entry-level clinical doctorate program. Their course content, at the time of data collection, had included no instruction in manual examination and treatment.

The experienced students group (13 females and 2 males) consisted of physical therapy students nearing the completion of their ninth academic term of a single, entry-level clinical doctorate program. All students had, at the time of data collection, completed four courses (12 semester hours) of academic work featuring manual examination and treatment skills in classroom and laboratory settings. Additionally, all students had 10 weeks of clinical education experience, albeit not necessarily in clinical settings allowing use of cervical spine palpation examination skills.

The novice clinician group (12 females and 4 males) consisted of newly licensed physical therapists, having recently graduated from an entry-level master's degree program. All clinicians at the time of data collection were within the first 5 months of practice as licensed physical therapists in a variety of practice settings. All were geographically located in Kentucky, except for one in West Virginia.

The experienced clinician group (12 males and 7 females) was comprised of licensed physical therapists with clinical experience ranging from 5 to 32 years with a mean of 16.5 years. All had an outpatient

Figure 2 Mechanical model diagram. 1: plywood base; 2: 2" x 4" wood base; 3: 2" x 4" wood upright; 4: turnbuckle; 5: extension spring; 6: pulley; 7: pulley; 8: cable; 9: chrome ring; 10: 'S' hook; 11: gutter bracket; 12: cable; 13: extension spring; 14: turnbuckle; 15: screw eye.

orthopedic setting as the primary locus of clinical practice. The experienced clinicians graduated from multiple institutions in obtaining their degrees ranging from bachelor's degrees to clinical doctorate degrees. All had completed extensive continuing education, predominantly not of an academic setting, for manual examination and treatment skills with a variety of philosophical approaches represented. All experienced clinicians for this study were located in Kentucky.

Student subjects were recruited from announcements posted in the area of their classrooms along with mass electronic mail announcements. Specific provisions were made to avoid any academic coercion as part of the recruitment activity. Novice and experienced clinicians were recruited by the principal investigator by electronic mail solicitation and by direct contact. The potential list of experienced clinicians was derived from prior professional association and the principal investigator's knowledge of general practice patterns of these clinicians being inclusive of manual examination and treatment of patients with spinal disorders. The electronic mail solicitation included a description of the study and the participation requirements in regards to time commitment. Clinicians were also informed that the data collection equipment and process would be brought to their clinical site at a time of their selection in order to facilitate their participation.

The previously mentioned mechanical model was used as a simulation instrument in data collection for assessing the accuracy of subjects in palpating and judging various levels of PIVM. Additionally, data collection included measurement of their mechanical

Figure 3 Diagram of testing apparatus and configuration.

performance characteristics during the activity while blinded to all conditions and not receiving performance feedback of any kind during the activity. The apparatus for data collection featured the mechanical spine model placed on a treatment table. A force platform (AccuSway Plus, Advanced Mechanical Technology, Inc., Watertown, MA, USA) was positioned adjacent to the treatment table with the midpoint of the force plate aligned with the middle module on the model. The force plate was connected to a laptop computer operating with Balance Clinic software (version 2.02.01). A testing protocol over 25 seconds duration was established at a default sampling rate of 50 Hz. The apparatus is represented in Fig. 3.

The subjects were asked to replicate performing PA passive movements on the mechanical model in the same manner, as was their practice to yield an assessment during the clinical examination of a patient's cervical spine. They were not directed to perform the technique in a specific way, but the subjects were aware of the need to detect potential differences in the model as a result of their examination. The forces applied to the model were measured using an indirect method. By virtue of the subjects standing on the force-plate, the forces placed downwardly onto the model were subtracted from the force plate and measured as negative values from the starting points of the subjects' body weights. For the purposes of this study, only those forces applied vertically (Z-plane) were considered as this vector is consistent with the orientation of PA pressures.

Subjects were asked to rate the resistance and mobility to their applied PA pressures to five simulated vertebrae of the model using a scale of one to five. The midpoint of the scale at a rating of three most closely replicated the force-displacement curves as measured by Tuttle *et al.*⁴⁵ Before data collection, each subject was allowed to complete three

practice trials on each module of the model to allow accommodation to the model and the testing procedure (45 oscillations total). Each subject moved left to right across the five simulated vertebrae, applying three oscillations of PA pressures. At the conclusion of each trial during testing, the subject was asked to rate each simulated vertebra. Each subject completed a total of 15 trials with five ratings in each trial, thus a total of 75 specific assessments. All subjects completed comparable trials of the model arrangement during data collection.

The collected force plate data, default sampled at 50 Hz, was processed through MatLab (version 7.7.0-R2008b) with a fourth-order low-pass Butterworth filter to minimize unwanted signal. As part of this process, residual analysis was undertaken, which indicated filtering at 4 Hz to be optimal for signal processing. The goal of the filtering was to allow the 15 force peaks representing the oscillations to be more readily identifiable by eliminating unwanted, higher frequency signal. At 4 Hz, however, some peaks consistent with the oscillatory force applications were eliminated or made more difficult to identify when compared to the raw data. Subsequent trials of filtering allowed determination that a frequency of 10 Hz offered the best clarity and sensitivity. From this processed data, the 15 identifiable peaks, usually occurring in five clusters of three repetitions were identified and their peak force values designated for analysis.

The groups were subsequently compared for accuracy of palpatory assessment, magnitude of force application, and performance characteristics. Specific analyses were conducted to determine intergroup differences. Additionally, analyses were completed to determine if any particular performance attributes were associated with greater accuracy.

At the conclusion of data collection, repeat testing of the model demonstrated a subtle shift to the right of the force-displacement curve of the level 3 modules. In effect, the level 3 modules became more like level 4 and discrimination between the two levels became more difficult. Thus, to compensate for this change, the comparisons of levels 3 and 4 were eliminated and the calculations adjusted commensurately.

The performance characteristics of all subjects were analyzed for their relationship to the accuracy of the palpatory interpretations concurrently completed during the data collection from the subjects. Mean peak force, frequency of the oscillations, and reliability of force delivery were analyzed for any effect on the subjects' abilities to be precise in rating the perceived level of force and displacement. This analysis was undertaken to determine if particular performance characteristics tend to facilitate more accurate assessment of PIVM.

Table 1 Absolute accuracy scores of the four groups

Group	Mean estimate	Standard error
First-year students	42.51	2.38
Third-year students	51.98	2.44
Novice clinicians	51.39	2.10
Experienced clinicians	53.46	1.93

The absolute accuracy (count of matched subject ratings to model settings) and relative error (based on magnitude of error on each non-identical match) scores of all subjects were compared to the performance characteristics of all subjects using a general linear model. This statistical model was chosen because of a lack of basis for assumption of a normal distribution for the response variables in question. Additionally, analysis of the outcomes of the absolute accuracy and relative error scores each reveals a resemblance to a negative binomial distribution.

In comparing the settings of the model to the ratings provided by the examiner, the number of identically matching decisions to settings were scored by a cumulative count. A perfect score would yield a value of 75, while the lack of a match on each possible decision would result in a score of zero. Additionally, the relative accuracy of the subject rating of the model setting was considered. The intent for this was to allow greater credit for being only one interval away from the model setting as opposed to being a larger number of intervals away from the model setting. Scores for both parameters were calculated for each subject and then calculated within each of the four cohorts. A mixed model analysis of variance, which controlled for any bias because of the time of data collection over the 3-month period, was used to compare the scores across the groups for both the number of identically matching scores and the relative error pertaining to accuracy across the numerical scale. Statistical analysis was completed using multiple tools including data management through Excel 2007® (Microsoft Corporation, Redmond, WA, USA) and processing with SAS® (SAS Institute, Inc., Cary, NC, USA).

Table 2 Mean relative error scores among groups

Group	Mean estimate	Standard error
First-year students	44.24	3.79
Third-year students	29.73	3.89
Novice clinicians	30.79	3.05
Experienced clinicians	32.49	3.32

Table 3 Description of forces applied to model within and across groups

Group	Subject mean ranges within groups (N)	Group mean (N)	Standard deviation
First-year students	5.71–48.86	20.38	13.66
Third-year students	4.61–63.82	21.20	14.86
Novice clinicians	3.71–38.99	17.54	11.20
Experienced clinicians	4.98–124.52	29.46	30.22

Despite this study containing a relatively large number of subjects in comparison to past passive movement studies, a greater number of subjects would have provided more impressive results. The achieved effect size for the measurements of absolute accuracy was 0.47 at alpha-level 0.05. *Post hoc* power analysis was calculated to be 0.89. The relative error measurements yielded an effect size of 0.40 at alpha-level 0.05. Analysis of power performed retrospectively results in a value of 0.77. To achieve power at 0.80 as proposed by Cohen,⁴⁶ a total sample size of 72 would be required at the same effect size.

Results

With comparisons across the four groups, only the difference of the first-year students to the other groups revealed a significant mean difference for the count of matching subject ratings to model settings ($P < 0.01$). Among the experienced student, novice clinician, and experienced clinician groups, no differences were detected in the means with P values ranging from 0.47 to 0.86. Two first-year students (12.5%), for whom this included novel tasks and judgments, scored above the mean of the experienced clinician group of this measure. In Table 1, the least squares means of the four groups are listed in reference to the number of identically matching subject ratings and model settings.

In similar fashion, the relative error scores were also compared between groups. The magnitude of variance of the subject's rating compared to the actual model setting was the basis for this score. The largest possible error score on each decision is four and the lowest (matching) is zero; this was totaled for each subject for each decision across all trials. A worst possible score was 300 and a perfect score of matching on all 75 opportunities would yield a value of zero. These are shown in Table 2.

The relative error scores were similar to the absolute accuracy measures. The first-year student group mean reflects less accuracy compared to all other groups with P values ranging between < 0.01 and 0.03. Among the other three groups, no between group differences existed (P values 0.60–0.84).

The ranges of forces imparted into the model to simulate PIVM were highly variable across all subjects and within the groups (Table 3). These values were calculated based upon the 15 identifiable peaks of force in each of the 15 trials for all 66 subjects.

The results reveal the experienced clinicians, as a group, applied the most force and the novice clinicians the least force. The first and third-year students were approximately the same and were between the two groups of clinicians. Thus, all four groups contained widely ranging individual mean force applications in completing the simulated PIVMs. This is most evidenced by the group standard deviations. Most notably, the experienced clinician group standard deviation is larger than the magnitude of the mean, reflecting a remarkable distribution across the range of forces. Significant differences in groups as to peak mean forces of the PIVM oscillations were seen between all pairs of groups (all $P < 0.01$), except for the first and third year students ($P = 0.10$). Again, a notable lack of consistency in force application is demonstrated.

The analysis of performance attributes and accuracy was completed with applying a negative binomial distribution and the general linear model. The absolute accuracy of all subjects with the coefficient for mean peak force was estimated to be < -0.01 with the concurrent probability of 0.04. Thus, a relationship of mean peak force and accuracy as measured by count of precise matches of examiner rating and the model setting is suggested. The application of greater force (measured negatively from the force platform) was associated with less absolute accuracy. A similar analysis was completed for the relative error scores with the force values. The coefficient for mean peak force was estimated to be < -0.01 with the concurrent probability of 0.15, not reaching a level of significance with the application of greater force and higher relative error scores.

Among the accuracy scores and force application measurements are individual values worthy of note. The most accurate individual, a clinician of 26-years experience, identically nominated the exact rating to the model setting on 72 of 75 decisions (96%). His mean peak force was measured at 6.37 N (1.43 pounds). Also within the experienced clinician group, an individual with over 30 years of practice, correctly matched the rating to the model setting only 33 of the 75 opportunities (44%). Her mean peak force was 124.51 N (27.99 pounds). Thus, a wide range of ability was demonstrated among experienced clinicians who routinely perform palpation examination procedures as part of daily clinical care. The second most accurate subject was a novice clinician correctly identifying 65 of 75 ratings (87%) and whose practice experience of 3 months at the time of data collection focused on in-patient acute rehabilitation. This subject's actual experience in completing cervical PIVM was predominantly in instructional laboratories, while in the physical therapy curriculum with only minimal experience completing cervical passive

motion in the clinical setting. The ability to perform the requested task at this level of precision would apparently be consistent with her individual abilities rather than the result of lengthy devoted practice and accumulated clinical experience. Her mean peak force was 21.06 N (4.73 pounds).

Discussion

The primary purpose of this study was to assess if particular performance attributes, such as magnitude of force application, were associated with assessment accuracy of a simulated palpation exam. Prior studies of PIVM have been unable to approach this aspect because of having no standard for comparison across multiple subjects. To reiterate, the act of assessing PIVM may easily change the status of the segments being examined in live subjects such that the reliability of the experience of palpating the examinee's spine cannot be established. With the reliability offered by a mechanical model across several examiner subjects, accuracy would theoretically be studied in greater expanse. Prior investigations using mechanical devices in the study of PIVM have essentially examined minimal detectable differences of elastic resistance. In this study, a mechanical model of the cervical spine was developed to approximate the force-displacement curve available in the human cervical spine when a PA PIVM was applied. The model also incorporated viscoelastic resistance and three-point bending with the application of PA PIVM to more closely simulate the experience of assessment of the human cervical spine for the subjects.

Discrimination of the differing levels of resistance offered by the modules across the five positions on the model was a task requiring considerable sensorimotor skill. In examining the scores across the four groups, only the first-year students for whom this represented a new activity, demonstrated a significant difference from the other groups. Among third-year students and clinicians, regardless of experience level, no differences were demonstrated in their abilities to interpret force-displacement quantities design to approximate that observed in the human cervical spine. The accepted convention is expertise with manual skills is the direct result of continuing education and clinical experience, yet relatively few data support this notion, including the outcome of this investigation. The task required in this investigation was not an encompassing index of physical therapy manual skills, but rather primarily the subject's capacity to discriminate and rate relatively small differences allowable by palpation skill. The greater expanse of psychomotor skills of a more complex nature used in the clinical setting may be better represented in another task or a series of tasks.

Additionally, the ability of skilled clinicians to identify other palpable subtleties, such as changes of tissue texture or temperature, may account for the levels of expertise often credited to experienced clinicians. Other dimensions of palpatory skill were not explored in this study.

Perhaps most noteworthy in these results are the ranges of scores across all subjects and within each group. Large arrays of values were represented even within somewhat homogenous groups of having similar education, training, and experience levels. Among the possibilities for this outcome is the potential for individuals to develop diverse internal concepts of normal and abnormal passive joint motion, particularly given the absence of any normative standard against which to assess their own perception of passively delivered movement. Among these cohorts, the third-year student group was the most homogenous at the time of data collection because of their relatively similar backgrounds of education and experience. Their familiarity level with manual skills at the time of data collection consisted primarily of performing manual examination and treatment skills on one another. As a group, they would presumably have less context of understanding abnormal motion states compared to practicing clinicians. The least homogenous group is the experienced clinician cohort, which was comprised of individuals with the greatest variety of education and training backgrounds. The members of this group would also have the most opportunity to develop the greatest diversity of conceptual frameworks of manual examination and treatment. Similarly, these individuals or others of similar backgrounds would have served as clinical instructors for the novice clinicians and possibly influenced the members of that group in their developing abilities.

A limitation of the prior studies examining the abilities of examiners to complete and interpret PIVMs has been the small number of subjects. The largest number of subjects in prior studies was 12.⁴⁷ The present study included 66 subjects among four cohorts at presumed levels of clinical aptitude. This greater number of subjects was made possible by the model allowing for greater standardization of the task across so many participants. Those prior studies, many with basic design weaknesses, suggested the clinician performance and assessment of PIVM is highly variable. Consistent with those prior investigations, this study further detailed the widely disparate performances among all levels of experience and knowledge.

The greatest variance in force application was among the experienced clinician group. Perhaps the variance in the task completed by examiners of PIVM is a contributor to the poor inter-rater reliability

evidence in the PIVM studies. This finding would logically kindle the question as to whether greater standardization of the PIVM procedure would improve the results. The absolute accuracy measure suggested the application of lighter forces was associated with greater precision in interpretation of the mobility and resistance ratings. The second accuracy measure, intended to account for a scale of error, did not reveal a similar result. If a relationship exists with lighter force application and greater accuracy, this would be consistent with the findings of Tuttle *et al.*,⁹ whose work indicated the end point of the force-displacement curve during PIVM is not as critical diagnostically as the midrange. The concept of using increased palpatory resistance to PA pressure to represent pathology is supported by the work of multiple investigators,^{48,49} which demonstrates increased resistance being associated with symptomatic levels. These investigators showed that changes in the early to middle portions of the force-displacement curves were more revealing than the end points in identifying symptomatic levels.⁴⁹ This would appear to challenge long held beliefs about the end point of passively delivered movement in the spine being of greatest importance in describing mechanical lesions. The benefit of force feedback in learning manual skills has already been demonstrated with improved reliability.⁴⁷ If further research confirms this relationship with enhancing diagnostic accuracy, expanded use of objective feedback in learning manual skills may be indicated in educational models.

As noted previously, the most accurate clinicians were both an expert of 26 years' experience (96% correct) and a novice clinician (87% correct). A wide range of accuracies were demonstrated among experienced clinicians. These results suggest that the ability to perform and interpret palpation tasks may be more related to individual abilities than to years of training or experience. This possibility may be worthy of further study.

Limitations

Limitations are present in this study and should be considered when evaluating the results. A major premise of this study is that a mechanical model adequately allowed simulation of a PIVM task in humans. Although such a model may be constructed to possess some of the physical properties of a human cervical spine, the complexity of the human anatomy and the accompanying biomechanics cannot be precisely simulated. Further, the interaction of the examiner and individual being examined is multi-dimensional and cannot be reduced to the relatively simple interpretation of force-displacement curves. Examiners usually observe for many indicators

clinical syndromes and overall patient presentation, both subtle and overt, in the course of a palpatory examination. Replicating those with a mechanical model is not possible.

In this study, the PIVM activity is completely extracted from all other examination findings and patient interchanges. The examiners in this study were essentially required to use PIVM as a simplified diagnostic task. Under the conditions used in this investigation, examiners could not consider concurrent pain responses, which have been found to improve psychometric values under experimental conditions.²⁷ The experienced clinicians, in particular, may not have had the opportunity to use other observations upon which they routinely depend for assistance in decision-making, including other communication results, observations, specific examination findings, and well-developed skills in pattern recognition. While perhaps better isolating this particular task and its interpretation, this investigation also suggests the important nature of the clinician's complete appreciation and assessment of a patient's presentation.

Another important limitation of this study concerns the circumstances of data collection. The students were in a relatively quiet environment. Data collection for the clinicians, novice and experienced, occurred at the sites of their clinical practice and often during the courses of their patient care days. The procedure occurred between patient visits with some subjects. Potential distractions from immediately surrounding patients, other clinicians, and clinical staff were present. Given the importance of attention to task along with the required retention of observations and interpretations for short-term recollection, the ability of clinicians to devote full focus to the requested undertaking may have been compromised on multiple occasions. Yet, this is the environment of assessment for clinicians utilizing PIVM as part of their regular exam process.

The availability of clinicians for testing purposes in adequate numbers representative of the physical therapy community was largely dependent upon data collection occurring at their sites of practice. Thus, access to clinician subjects largely governed this limitation being present. Similarly, the distribution of genders across the subjects in both student and clinician groups is a limitation. The subject groups are not representative of the population of students and practicing clinicians in this regard.

Conclusions

Multiple studies have shown the assessment of PIVM to have suspect validity and reliability. The lack of agreement between clinicians in their palpatory assessments has allowed questioning of many of the

constructs of traditional manual therapy practices. The concern for the validity of manual therapy approaches is further reinforced by the widely varied performances of the magnitude of force applications of PIVMs, which was again demonstrated in this study.

The suggestion by this study is of the performance and interpretation of cervical PIVMs is, at least in large part, a function of native sensory discriminatory ability. Some novice students appear to grasp the concept and to have interpretive skill level performances comparable to experienced clinicians. Additionally, the data also imply that experience alone does not automatically advance a practitioner's skill level. How much of these skills are within the realm of refinement by instruction, practice, and feedback are questions for further investigation. Select experienced clinicians, perhaps those with greater native ability and refined from clinical practice, appear to demonstrate exceptional abilities to discriminate subtle changes in simulated PIVM. Whether this superior ability equates to better patient outcomes has yet been determined.

A portion of the data from this study indicates that lightly applied pressures, early in the force-displacement curve of cervical PIVM, may increase discriminatory ability. The evidence is not sufficient, however, to validate this as a clearly superior performance attribute. Until more is known as to whether a particular method of PIVM performance and assessment is superior, the nature of teaching and learning of PIVMs remains largely subjective and imprecise.

References

- Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner MR, Whitman JM, *et al.* Neck pain. *J Orthop Sports Phys Ther.* 2008;38:A1-34.
- Hurwitz EL, Carragee EJ, van der Velde G, Carroll LJ, Nordin M, Guzman J, *et al.* Treatment of neck pain: noninvasive interventions: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine (Phila Pa 1976).* 2008;33:S123-52.
- Jull G, Trott P, Potter H, Zito G, Niere K, Shirley D, *et al.* A randomized controlled trial of exercise and manipulative therapy for cervicogenic headache. *Spine (Phila Pa 1976).* 2002;27:1835-43; discussion 1843.
- Hurwitz EL, Morgenstern H, Harber P, Kominski GF, Yu F, Adams AH. A randomized trial of chiropractic manipulation and mobilization for patients with neck pain: clinical outcomes from the UCLA neck-pain study. *Am J Public Health.* 2002;92:1634-41.
- Walker MJ, Boyles RE, Young BA, Strunce JB, Garber MB, Whitman JM, *et al.* The effectiveness of manual physical therapy and exercise for mechanical neck pain: a randomized clinical trial. *Spine (Phila Pa 1976).* 2008;33:2371-8.
- Tseng YL, Wang WT, Chen WY, Hou TJ, Chen TC, Lieu FK. Predictors for the immediate responders to cervical manipulation in patients with neck pain. *Man Ther.* 2006;11:306-15.
- Thiel HW, Bolton JE. Predictors for immediate and global responses to chiropractic manipulation of the cervical spine. *J Manipulative Physiol Ther.* 2008;31:172-83.
- Rubinstein SM, Knol DL, Leboeuf-Yde C, de Koekkoek TE, Pfeifle CE, van Tulder MW. Predictors of a favorable outcome

- in patients treated by chiropractors for neck pain. *Spine (Phila Pa 1976)*. 2008;33:1451–8.
- 9 Tuttle N, Barrett R, Laakso L. Posteroanterior movements in tender and less tender locations of the cervical spine. *Man Ther*. 2009;14:28–35.
- 10 Siegenthaler A, Eichenberger U, Schmidlin K, Arendt-Nielsen L, Curatolo M. What does local tenderness say about the origin of pain? An investigation of cervical zygapophysial joint pain. *Anesth Analg*. 2010;110:923–7.
- 11 Jull G, Bogduk N, Marsland A. The accuracy of manual diagnosis for cervical zygapophysial joint pain syndromes. *Med J Aust*. 1988;148:233–6.
- 12 Bogduk N. In defense of King *et al.*: the validity of manual examination in assessing patients with neck pain. *Spine J*. 2007;7:749–52; author reply 752.
- 13 King W, Lau P, Lees R, Bogduk N. The validity of manual examination in assessing patients with neck pain. *Spine J*. 2007;7:22–6.
- 14 Haas M, Grouppe E, Panzer D, Partna L, Lumsden S, Aickin M. Efficacy of cervical endplay assessment as an indicator for spinal manipulation. *Spine (Phila Pa 1976)*. 2003;28:1091–6; discussion 1096.
- 15 Schomacher J. The effect of an analgesic mobilization technique when applied at symptomatic or asymptomatic levels of the cervical spine in subjects with neck pain: a randomized controlled trial. *J Man Manip Ther*. 2009;17:101–8.
- 16 Aquino RL, Caires PM, Furtado FC, Loureiro AV, Ferreira PH, Ferreira ML. Applying joint mobilization at different cervical vertebral levels does not influence immediate pain reduction in patients with chronic neck pain: a randomized clinical trial. *J Man Manip Ther*. 2009;17:95–100.
- 17 Olson K, Paris S, Spohr C, Gorniak G. Radiographic assessment and reliability study of the craniovertebral side-bending test. *J Man Manip Ther*. 1998;6:87–96.
- 18 Smedmark V, Wallin M, Arvidsson I. Inter-examiner reliability in assessing passive intervertebral motion of the cervical spine. *Man Ther*. 2000;5:97–101.
- 19 Kaltenborn F, Evjenth O, Kaltenborn T, Morgan D, Vollowitz E. Manual mobilization of the joints: the Kaltenborn method of joint examination and treatment. Vol. II. The spine. 5th ed. Oslo: Norli; 2009.
- 20 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 (Phila Pa 1976)*. 2004;29:E413–25.
- 21 Humphreys BK, Delahaye M, Peterson CK. An investigation into the validity of cervical spine motion palpation using subjects with congenital block vertebrae as a 'gold standard'. *BMC Musculoskelet Disord*. 2004;5:19.
- 22 Marcotte J, Normand M. Standardisation de la palpation dynamique en chiropratique: une étude de fiabilité pour la région cervicale. *J Can Chiropr Assoc*. 2001;45:106–12.
- 23 Deboer KF, Harmon R Jr, Tuttle CD, Wallace H. Reliability study of detection of somatic dysfunctions in the cervical spine. *J Manipulative Physiol Ther*. 1985;8:9–16.
- 24 Nansel DD, Peneff AL, Jansen RD, Cooperstein R. Interexaminer concordance in detecting joint-play asymmetries in the cervical spines of otherwise asymptomatic subjects. *J Manipulative Physiol Ther*. 1989;12:428–3.
- 25 Fjellner A, Bexander C, Faleij R, Strender LE. Interexaminer reliability in physical examination of the cervical spine. *J Manipulative Physiol Ther*. 1999;22:511–6.
- 26 Strender LE, Lundin M, Nell K. Interexaminer reliability in physical examination of the neck. *J Manipulative Physiol Ther*. 1997;20:516–20.
- 27 Piva SR, Erhard RE, Childs JD, Browder DA. Inter-tester reliability of passive intervertebral and active movements of the cervical spine. *Man Ther*. 2006;11:321–30.
- 28 Pool JJ, Hoving JL, de Vet HC, van Mameren H, Bouter LM. The interexaminer reproducibility of physical examination of the cervical spine. *J Manipulative Physiol Ther*. 2004;27:84–90.
- 29 Bertilson BC, Grunnesjo M, Strender LE. Reliability of clinical tests in the assessment of patients with neck/shoulder problems—impact of history. *Spine*. 2003;28:2222–31.
- 30 Johnston WL, Hill JL, Elkiss ML, Marino RV. Identification of stable somatic findings in hypertensive subjects by trained examiners using palpatory examination. *J Am Osteopath Assoc*. 1982;81:830–6.
- 31 Johnston W. Interexaminer reliability in palpation. *J Am Osteopath Assoc*. 1976;76:286–7.
- 32 Johnston WL, Hill JL, Sealey JW, Sucher BM. Palpatory findings in the cervicothoracic region: variations in normotensive and hypertensive subjects. A preliminary report. *J Am Osteopath Assoc*. 1980;79:300–8.
- 33 Comeaux Z, Eland D, Chila A, Pheley A, Tate M. Measurement challenges in physical diagnosis: refining inter-rater palpation, perception and communication. *J Bodyw Mov Ther*. 2001;5:245–53.
- 34 Schops P, Pfingsten M, Siebert U. Reliability of manual medical examination techniques of the cervical spine. Study of quality assurance in manual diagnosis. *Z Orthop Ihre Grenzgeb*. 2000;138:2–7. German.
- 35 Hardy G, Napier J. Inter and intratherapist reliability of passive accessory movement technique. *N Z J Physiother*. 1991;19:22–4.
- 36 Simmonds MJ, Kumar S, Lechelt E. Use of a spinal model to quantify the forces and motion that occur during therapists' tests of spinal motion. *Phys Ther*. 1995;75:212–22.
- 37 Snodgrass SJ, Rivett DA, Robertson VJ. Manual forces applied during posterior-to-anterior spinal mobilization: a review of the evidence. *J Manipulative Physiol Ther*. 2006;29:316–29.
- 38 Snodgrass SJ, Rivett DA, Robertson VJ. Manual forces applied during cervical mobilization. *J Manipulative Physiol Ther*. 2007;30:17–25.
- 39 Snodgrass SJ, Rivett DA, Robertson VJ, Stojanovski E. Forces applied to the cervical spine during posteroanterior mobilization. *J Manipulative Physiol Ther*. 2009;32:72–83.
- 40 Snodgrass SJ, Rivett DA, Robertson VJ, Stojanovski E. Cervical spine mobilisation forces applied by physiotherapy students. *Physiotherapy*. 2010;96:120–9.
- 41 Juliano T, Forster A, Drzal P, Weerasooriya T, Moy P, VanLandingham M. Multiscale mechanical characterization of biomimetic physically associating gels. *J Mater Res*. 2006;21:2084–92.
- 42 Juliano T, Moy P, Weerasooriya T, VanLandingham M, Forster A. Compression and instrumented indentation measurements on biomimetic polymers. Proceedings of the Society for Experimental Mechanics Conference; 2006 Jun 3–6; St Louis, MO, USA: Curran Associates.
- 43 Lee RY, McGregor AH, Bull AM, Wragg P. Dynamic response of the cervical spine to posteroanterior mobilisation. *Clin Biomech (Bristol, Avon)*. 2005;20:228–31.
- 44 McGregor A, Bull A, Lee R, Wragg P. Dynamic response of the human spine to anteroposterior mobilisation manual therapy: an interventional magnetic resonance imaging study. *Physiotherapy*. 2004;90:165–6.
- 45 Tuttle N, Barrett R, Laakso L. Postero-anterior movements of the cervical spine: repeatability of force displacement curves. *Man Ther*. 2008;13:341–8.
- 46 Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Erlbaum; 1988.
- 47 Anson E, Cook C, Camacho C, Gwilliam B, Karakostas T. The use of an educational model in the improvement of student reliability in finding R1. *J Man Manip Ther*. 2003;11:204–12.
- 48 Gay RE, Ilharreborde B, Zhao K, Zhao C, An KN. Sagittal plane motion in the human lumbar spine: comparison of the *in vitro* quasistatic neutral zone and dynamic motion parameters. *Clin Biomech (Bristol, Avon)*. 2006;21:914–9.
- 49 Petty NJ, Maher C, Latimer J, Lee M. Manual examination of accessory movements — seeking R1. *Man Ther*. 2002;7:39–43.