Published in final edited form as:

J Manipulative Physiol Ther. 2010 June ; 33(5): 378–385. doi:10.1016/j.jmpt.2010.05.005.

Anterior Superior Iliac Spine Asymmetry Assessment on a Novel, Pelvic Model: an Investigation of Accuracy and Reliability

Bradley A Stovall, Sejong Bae, and Shrawan Kumar

Abstract

Objective—The purpose of this study was to develop a novel pelvic model and determine the accuracy, inter- and intra-examiner reliability of anterior superior iliac spine (ASIS) positional asymmetry assessment from both sides of the model by osteopathic pre-doctoral fellows and osteopathic physicians and to evaluate the effect of training.

Methods—Five osteopathic pre-doctoral fellows and 5 osteopathic physicians assessed 13 settings of varied ASIS asymmetry of a novel pelvic model for superior/inferior positional asymmetry from both sides of the model in a random order. Assessment from the right and left sides of the model occurred on 2 separate days. Fellows were trained for a week and retested.

Results—Average inter-examiner reliability was greatest from the left side of the model for physicians and right side for fellows (physicians $k=0.46$; fellows $k=0.37$ respectively) while intra-examiner reliability was greatest from the right in both groups (physicians $k=0.49$; fellows $k=0.52$). Following training of fellows, inter-examiner reliability remained highest from the right side of the model (right: $k=0.48$; left: $k=0.36$) while intra-examiner reliability was higher from the left side (right: $k=0.53$; left: $k=0.59$). Physicians and fellows before training were more accurate from the right side of the model ($k=0.56$ and $k=0.52$ respectively). Following training of fellows, accuracy increased from both sides of the model (right: $k=0.59$; left: $k=0.53$).

Conclusions—A novel, pelvic model was developed to allow assessment of accuracy and reliability of ASIS asymmetry assessment. Individually, physicians and fellows varied in accuracy and inter-/intra-examiner reliability. Further investigation is warranted to understand the clinical and educational application of these results.

Keywords

Palpation; Models, Anatomic; Osteopathic Manipulative Treatment; Chiropractic

INTRODUCTION

As research has begun to evaluate diagnostic and treatment models in manual medicine, the development of valid and reliable models of musculoskeletal dysfunction has remained elusive. Historically, manipulative therapies in these professions developed clinically and subsequent anatomy and physiology was ascribed to them¹. Bony anatomical landmark positional asymmetry is a commonly used form of musculoskeletal assessment hypothesized to give information regarding the relative positions of the structures in question and has

© 2010 National University of Health Sciences. Published by Mosby, Inc. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

gained widespread acceptance in manual medicine²⁻⁷. Some manipulative therapies from the osteopathic, chiropractic and physiotherapy professions are based upon this method of assessment. Of these, Muscle Energy Technique is the most well known and is currently taught in osteopathic institutions throughout the world.

Bony anatomical landmark positional asymmetry, however, is not the only method of assessment used by practitioners of manual medicine. Pain provocation, point tenderness, palpation for taut bands, range of motion, and motion testing are often used in combination with patient history comprising elements of the clinical evaluation. For motion testing in the spine and pelvis, current literature reviews demonstrate poor reliability⁸⁻¹¹. Most pain provocation tests likewise have inadequate reliability^{12,13}. In a recent literature review, only two pain provocation tests by themselves or multiple tests in a group demonstrated adequate reliability^{13,14}. While acceptable reliability has been demonstrated for some pain provocation testing, this form of palpation shares common confounding elements when assessing for 'tender points' or tenderness. Because of the subjective nature of pain interpretation as well as the potential role of patient expectation with repeated assessments of tenderness or pain in research, the reliability of these methods though adequate must be interpreted with caution.

Positional asymmetry of pelvic landmarks is commonly ascribed to dysfunction of the sacroiliac joint (SIJ)^{2-7,15}. Positional findings are often named based upon motion testing hypothesized to localize the dysfunctional SIJ^{2-7,15}. However, poor reliability and experimental evidence investigating the amount of motion occurring at the SIJ has called into question the validity of bony anatomical landmark positional assessment methodologies^{12,16-19}. While this challenges the current concepts of pelvic dysfunction commonly taught in manual medicine, recent research between pelvic bony asymmetry and gross coupled motion in the lumbar spine suggests the potential clinical relevance of bony pelvic asymmetry²⁰⁻²².

Investigation into bony anatomical landmark positional asymmetry assessment in recent years has focused primarily on the pelvic girdle and medial malleoli (MM). MM have commonly been included when evaluating the pelvic girdle because of an assumed association with dysfunction in this region³⁻⁷. Consistently, ASIS and MM have demonstrated higher reliability than other landmarks with the MM always having the highest reliability²³⁻²⁷. Challenges such as examiner fatigue, thumb placement, amount of detectable asymmetry and variation in technique have been discussed as potential factors related to the low inter-examiner reliability observed in evaluation of these methods^{23,24,26,27}. A recent experiment of MM assessment designed to address some of these concerns demonstrated the usual higher reliability for MM and when approximately 4mm of asymmetry was screened prior to assessment, inter-examiner reliability was near perfect²⁶. Thus, the amount of asymmetry appeared to be a significant factor in inter-examiner reliability.

In another study, Fryer et al (2005) evaluated the role of training on the reliability of bony landmark positional assessment and found that training improved reliability for some landmarks while not for others. For example in this study, reliability of ASIS assessment improved significantly while MM reliability did not significantly improve following training²⁷. Another more specific method of training termed 'consensus training' was investigated in the lumbar spine by Degenhardt et al (2005)²⁸. This was designed to address the known low inter-examiner reliability of musculoskeletal palpatory assessment. This training involved examiner discussion following observed differences of assessments during the training phase. Also throughout the experiment, examiners attempted to modify their evaluation procedures until methods were as similar as possible²⁸. From this 'consensus

training', significant improvements were demonstrated indicating the high degree of skill and/or consensus that may be needed for reliable palpatory assessment.

Recent research into these methods of assessment has for the first time allowed practitioners of manual medicine to begin the process of objectively understanding the reliability of assessment methods employed by these professions. Though these previous studies have allowed the first step in understanding the role of human perception in the palpatory process, they have not assessed the accuracy or role of examiner placement commonly advocated in manual medicine texts. With no gold standard of musculoskeletal function, dysfunction or pain, poor reliability testing affords the researcher with little more than an incomplete knowledge of a specific diagnostic method. While this research has demonstrated clear trends, it is the opinion of the authors that understanding the role of human perception in the context of low reliability of palpatory assessment is of significant importance for professions that incorporate the use of these assessment methods. Thus, the goal of the experiment was to develop a novel pelvic model and determine the accuracy, inter- and intra-examiner reliability of ASIS positional asymmetry assessment from both sides of the model by osteopathic pre-doctoral fellows and osteopathic physicians and to evaluate the effect of training on reliability.

METHODS

Participants

Five Osteopathic pre-doctoral fellows and five Osteopathic physicians from the University of North Texas Health Science Center – Texas College of Osteopathic Medicine (UNTHSC-TCOM) department of Osteopathic Manipulative Medicine participated in this study. Pre-doctoral fellows are osteopathic medical students that have shown proficiency in manipulative medicine as observed by faculty and are pursuing an extra year of study to further develop their manipulative medicine skills. These fellows have participated in the 1st and 2nd year of osteopathic undergraduate curriculum that has significant emphasis on anatomical landmark positional asymmetry as guided by the Educational Council on Osteopathic Principles. Recruitment was conducted on a voluntary basis. Three of the physicians were residency trained in Osteopathic Manipulative Medicine/Neuromuscular Medicine (OMM/NMM). All physicians were board certified in family medicine and/or OMM/NMM and in active clinical practice. The physicians had varied amounts of clinical experience: 5, 8, 28, 33, and 40 years. The project was approved by the IRB of the University, and all subjects signed informed consent prior to taking part in the experiment.

Instrumentation

The Pelvic Model—The pelvis of a skeleton, bisected along the mid-sagittal plane was mounted on two closely placed and aligned separate wooden boards. One of these boards moved with precision on a standardized graduated scale in translation in the coronal and transverse planes. The model was placed on a table 2 ft off of the ground on a wooden plank. The model itself had 1 inch wide curved metal arches at each end that allowed latex skin to be placed over it to simulate abdominal and pelvic skin and velcroed on the underside. The model was covered with foam and a latex skin to simulate patient skin and subcutaneous tissue. A schematic of the model invented by Kumar (2009) can be found in Figure 1.

Calibration—Calibration was conducted first by leveling the ASIS with a Black and Decker digital level. Following calibration of the model for levelness, zeroing of the model occurred by setting the model level for superior/inferior translational displacement. This process was done by one of the authors palpating the most inferior slope of the ASIS on the

surface of the landmark. From the combination of palpation and visual analysis, the ASIS was then labeled with a watermark pen. A wooden block was placed exactly perpendicular to the model, and individual paper cut-outs were adhered to the block for assessment. A T square ruler was then placed at a right angle to the block and used to visually calibrate the marked points. An initial mark was placed on the cut-out to mark the initial position of the moveable ASIS. The ASIS was then adjusted so the movable ASIS was visibly level with the stationary counterpart with the aid of the T square without moving the wooden block. A subsequent mark was placed on the square cut-out indicating the calibrated zero displacement position. This was repeated 9 times on separate square cut-outs. With the use of a Vernier caliper with an electrical read out, displacement between lines was measured and recorded for each cut-out. The average displacement and standard deviation was represented as the calibrated, leveled (0mm) position.

The final phase of calibration was carried out by advancing one half of the pelvic model on the other in increments of 2 mm and measuring displacement of the marks on the wooden blocks with a Vernier caliper. This was done for a nominal distance of 24 mm around and including the zero position representing asymmetry and levelness of the left and right ASIS. Ten measurements with ruler adjusted displacement and Vernier caliper measurement were made of the 13 settings in this manner. The data of the ruler adjusted displacement was plotted against the Vernier caliper reading. A linear fit of data represented a satisfactory calibration as seen in Figure 2.

Procedure

Initially, the procedure was conducted over two examination sessions. There were 13 settings which were evaluated three times during each session. The mean of the 10 readings during calibration represented the value at that setting. The settings ranged from right ASIS inferior by 12mm to right ASIS superior by 12mm in 2mm increments including 0mm displacement. The order of presentation was randomized for each evaluation session, and examiners had no knowledge of the setting variation or repeated assessment. Examiners were instructed to determine which ASIS was more superior to the other or if the landmarks were level. Examiner instructions were to palpate the inferior slope of the ASIS through palpation with their thumbs. Following each assessment, examiners also estimated the displacement in millimeters and recorded both of these findings on a data collection sheet and then placed it in a box next to the model after each examination. Examiners palpated from right side of the model during the first session and the left side the second session. For fellows, evaluation in a group of 5 from the left occurred two days after evaluation from the right. Because of physicians' schedules, examination from the left occurred approximately one week after evaluation from the right. Physicians performed assessments in pairs or individually. Except during the period of feedback training, fellows were instructed not to discuss their findings or opinions about the model until completion of the experiment. During examination of each setting, one minute was allowed for evaluation and recording of findings. Subjects were out of the examination room as the model was adjusted for the next test condition. For pre-doctoral fellows, the entire experiment was then repeated after five consecutive days of feedback training. Because of scheduling conflicts during this week of training, two fellows were only able to attend four of the sessions. For the post training evaluation, fellows evaluated the model first from the right and then the left two days apart in groups of two and three because of time constraints. Following the final examination session for all subjects, perceived eye dominance, handedness, and side of table preference of ASIS evaluation were recorded. Eye dominance was then independently determined with the distance hole-in-the-card test and recorded for pre-doctoral fellows and Osteopathic physicians. The entire study for all subjects took four weeks to complete secondary to physician availability.

Training

Fellows went through a series of feedback training sessions on six different randomized settings of the pelvic model. Examiners were told of the possible settings to maximize training effect to the specific settings. Two were of equal height, two were of 5 mm height difference, and the last two were of 10 mm height difference. For each pair of measurements where applicable, one was more superior on the right and one on the left. The selection of these specific settings was determined to provide examiners a reference for amounts of asymmetry on the model without training subjects to all of the settings specifically used in the methodology. The order of presentation was randomized for each training session; after each assessment, subjects were told of the side and exact difference in ASIS heights on the pelvic model. Subjects were then allowed to re-evaluate the model at the particular setting and discuss amongst themselves reasons for observed differences in assessment. Subjects were also allowed to evaluate the setting from the other side of the model to evaluate perceived differences secondary to subject positioning. Osteopathic physicians only evaluated the model once from each side of the table and did not take part in any training sessions.

Data Collection and Analysis

Data analyses were performed with SAS version 9.1 (SAS, Cary, NC, USA). Cohen's Kappa statistic was calculated for individual agreement with the gold standard of this experiment (calibrated ASIS model setting). Accuracy in percent agreement was also determined for each setting of the model. Fleiss's Kappa statistic was calculated among evaluators to determine inter- and intra-examiner reliability (across the three assessments at each setting). The strength of agreement for the Kappa statistic was interpreted as small (0.00-0.20); fair (0.21-0.40); moderate (0.41-0.60); substantial (0.61-0.80); and almost perfect (0.81-0.99).

RESULTS

Results are summarized in Tables 1-3. Prior to training, the resulting Kappa average inter-examiner reliability was higher from the left side of the model for physicians and right side for fellows (moderate agreement for physicians $k=0.46$; fair agreement for fellows $k=0.37$ respectively) while intra-examiner reliability was higher from the right in both groups (moderate agreement for both physicians $k=0.49$; fellows $k=0.52$). Following training of fellows, inter-examiner reliability remained higher from the right side of the model (right: $k=0.48$; left: $k=0.36$) while intra-examiner was higher from the left side (moderate agreement for both right: $k=0.53$; left: $k=0.59$).

Before training, average accuracy calculations demonstrates that physicians and fellows were more accurate from the right side of the model (moderate agreement for both $k=0.56$ and $k=0.52$ respectively). Following training of fellows, agreement increased from both sides of the model (right: $k=0.59$; left: $k=0.53$). The kappa from the left side evaluation changed from 0.49 to 0.53 while on the right it changed from 0.52 to 0.59 (Table 1). This increase in accuracy was greater than accuracy of physicians without the training sessions from each side of the model.

For the physician and fellow group, average percent agreement for each setting from both sides of the model can be found in Figures 3. As is apparent, accuracy as viewed through percent agreement tended to increase as asymmetry increased. Though this was expected, no obvious threshold of accuracy was apparent. Of note, examiners displayed a higher degree of accuracy when assessing the left ASIS from the right side of the model before training as evident in Fig. 3a; however, this trend did not remain for right ASIS evaluation from the left

side (Fig. 3b) of the model. No specific associations between hand dominance, eye dominance, or examiner preference and reliability or accuracy were observed (Table 1).

DISCUSSION

Palpatory examination in the context of manual medicine often occurs before and after treatment^{2-5,7,15}, but the extent that bony positional asymmetry assessment drives technique selection and implementation has not been well defined. This will remain the case until reliable, standardized methods are established and compared to treatment methods and objective outcomes. For instance if a clinician is performing motion testing along with positional asymmetry assessment of a structure, the extent that one may rely on asymmetry assessment or motion testing when selecting a treatment in the presence of contradictory information represents one of the nuances found in clinical practice.

Inter-examiner reliability without training observed in this experiment was significantly higher than averages from a recent review article for ASIS evaluation ($k=0.13$ vs. $k=0.40$ from the right and left) while intra-examiner reliability was only slightly higher than previous research ($k=0.41$ vs. $k=0.48$ from the right and left)¹⁹. The higher inter-examiner reliability could be related to numerous factors. First, the amount of asymmetry as defined in the methodology could represent asymmetry that is greater than asymmetry found in the patient populations of previous studies. Secondly, one pair of landmarks was evaluated repetitively. In contrast to previous research, 9 up to 32 different subjects were used¹⁹. An adaptation effect of repetitive evaluation of only one pair of ASIS's could have played a role in the higher observed inter-examiner reliability. Another factor that could have led to the higher reliability seen in assessment was specific to the methodology in that the model was evaluated 39 consecutive times for each evaluation session. Though the settings were randomized, the ability of examiners to learn from comparison of previous settings in this experiment cannot be discounted. Intra-examiner reliability that more closely resembles previous studies of this nature suggests that the methodology or the model played a factor in the higher observed inter-examiner reliability.

When comparing the results of physicians with significant clinical experience in manipulative medicine to pre-doctoral fellows, inconsistent accuracy and reliability were demonstrated. For some evaluations, physicians were more reliable while for others, fellows were more reliable. In the context of existing literature, the lack of clear delineation between clinicians and advanced students is not surprising^{10,24}. The variation in accuracy as well as inter-examiner reliability of groups with different levels of experience further supports previous findings that clinical experience itself is not directly associated with examiner skill of bony anatomical landmark positional asymmetry assessment.

The concept of detection thresholds has been suggested to be useful in skill development and is based on the idea that at some threshold of decreasing asymmetry, reliability approaches chance and accuracy of assessment decreases significantly²⁹. Factors involved in this could be related to elements such as: examiner palpation, visual perception, and skill of assessment. From gross observation of individual percent agreement, some subjects in this study demonstrated narrow windows of decreased reliability and overall high accuracy while others were accurate only with a larger degree of asymmetry. Some subjects demonstrated significant differences in accuracy and inter-examiner reliability of assessment from different sides of the model while others showed little variation when evaluating the model from different sides. The absence of a consistent trend, however, does not rule out the existence of quantifiable detection thresholds of assessment. Subjects were given the options of right ASIS greater than left, equal or left greater than right to more closely resemble

clinical decision making. In future studies to more specifically address the concept of detection thresholds, removal of the option for ASIS equal would be useful.

In manual medicine texts advocating anatomical landmark asymmetry assessment as an important component of evaluation, eye dominance associated examiner positioning has been suggested as an important factor to improve accuracy by limiting visual parallax errors^{2-4,15}. This common instruction appears to be based on expert clinical opinion with no supporting data. No consistent associations were observed with eye dominance and examiner position (Table 1), though a larger sample size would be required to make any conclusions regarding the relevance of eye dominance associated positioning. No other specific associations of hand dominance or side preference were observed (Table 1).

Determining what constitutes acceptable levels of inter-examiner reliability has been given qualitative descriptors by Landis and Koch (1977): slight (0.00-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80), and almost perfect (0.81-0.99)³⁰. More recently in physical medicine, thresholds of kappa at 0.40 and 0.60 have been discussed as relevant thresholds for clinical significance³¹. Using this classification, results from this experiment demonstrate fair to moderate reliability as seen in Table 2. From both sides of the model, physicians had moderate reliability while before training fellows had fair reliability from each side. Following training, reliability increased to moderate from the right side while remained fair from the left (Table 2).

Limitations

The use of any model presents inherent limitations to the practical application of experimental results. However, the use of a pelvic model was chosen to better control the variables while retaining some element of realism and allow for evaluation of accuracy and reliability in ASIS asymmetry assessment. All prior studies of pelvic positional asymmetry have demonstrated low inter-examiner reliability and were conducted in vivo, therefore lack known accuracy of asymmetry. Considering the results of previous in vivo experimentation, the authors felt that it was appropriate to evaluate the accuracy of human perception in the context of bony anatomical landmark evaluation with a model. However, it is conceded that a model cannot entirely simulate a patient. Another limitation was that the model was a simulation of only the pelvis and there was no head, torso, or upper and lower limbs. Further, the model was not placed on a traditional evaluation table that would allow height adjustment. The significant limitation of this experiment was the unforeseen challenge of subject scheduling as explained in the Procedure of the Methods section. When performing direct comparison between pre- and post-training as well as comparison of physicians and fellows, one must keep in mind the variation in size of groupings and number of days between assessments. Future experiments of this nature should address subject scheduling when enrolling subjects with limited schedules.

CONCLUSION

A novel, pelvic model was developed to allow assessment of accuracy and reliability of ASIS evaluation in a controlled setting. Comparison of the physician and fellow groups demonstrated varied accuracy and reliability. The inter-examiner reliability observed for physicians from both sides of the model and for fellows from the right side of the model after training reached levels that can be considered clinically significant. Therefore, training of students and possibly clinicians has the potential enhance diagnostic accuracy and consistency. Continued development of realistic models for training of palpatory diagnosis along with increased understanding of anatomical and physiological models of function and dysfunction are needed to improve training and education in all professions in manipulative

medicine. Future experiments should explore the reasons for examiner variability and compare model to in vivo asymmetry assessment.

PRACTICAL APPLICATIONS

- Training on a pelvic model may potentially improve reliability of ASIS asymmetry assessment.
- Students with similar levels of training can have variable accuracy and skill of palpation assessment.
- Models of palpation assessment should be developed to improve the quality of education in manipulative therapy professions.

Acknowledgments

The project described was supported by a Fellowship Grant from the American Osteopathic Association and also by Award Number 5T35AT004388-02 from the National Center for Complementary & Alternative Medicine. The content is solely the responsibility of the authors and does not necessarily represent the official views of the American Osteopathic Association or the National Center for Complementary & Alternative Medicine or the National Institutes of Health.

REFERENCES

1. Jordan TR. Conceptual and Treatment Models in Osteopathy I. AAOJ 2003;26:29–36.
2. DiGiovanna, EL.; Schiowitz, S.; Dowling, DJ. An osteopathic approach to diagnosis and treatment. 3rd ed.. Lippincott Williams & Wilkins; Philadelphia: 2005.
3. Isaacs, ER.; Bookhout, MR.; Bourdillon, JF. Bourdillon's spinal manipulation. 6th ed.. Butterworth-Heinemann; Woburn, MA: 2002.
4. Greenman, PE. Principles of manual medicine. 3rd ed.. Lippincott Williams & Wilkins; Philadelphia: 2003.
5. Mitchell, FL.; Moran, PS.; Pruzzo, NA. An evaluation and treatment of osteopathic muscle energy procedures. Mitchell, Moran and Pruzzo Associates; Valley Park, Mo: 1979.
6. Mitchell FL. Structural Pelvic Function. AAO Yearbook 1958:71–90.
7. Kuchera, WA.; Kuchera, ML. Osteopathic principles in practice. Kirksville College of Osteopathic Medicine; Kirksville, Mo: 1992. 2, rev ed.1991
8. Hestock L, LeboeufYde C. Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review. J Manipulative Physiol Ther 2000;23(4):258–275. [PubMed: 10820299]
9. Christensen H, Vach W, Vach K, et al. Palpation of the upper thoracic spine: an observer reliability study. J Manipulative Physiol Ther 2002;25:285–92. [PubMed: 12072848]
10. Seffinger M, Najm W, Mishra S, Adams A, Dickerson V, Reinsch S. Reliability of Spinal Palpation for Diagnosis of Back and Neck Pain: A Systematic Review of the Literature. Spine 2004;29(19):E413–E425. [PubMed: 15454722]
11. Boline P, Haas M, Meyer J, Kassek K, Nelson C, Keating J. Interexaminer reliability of eight evaluative dimensions of lumbar segmental abnormality:Part 2. J Manipulative Physiol Ther 1993;16:363–74. [PubMed: 8409784]
12. van der Wurff P, Hagmeijer RHM, Meyne W. Clinical tests of the sacroiliac joint: A systematic methodological review. Part 1: Reliability. Man Ther 2000;5(1):30–36. [PubMed: 10688957]
13. Szadek KM, van der Wurff P, van Tulder MW, Zuurmond WW, Perez RSGM. Diagnostic Validity of Criteria for Sacroiliac Joint Pain: A Systematic Review. J Pain 2009;10(4):354–368. [PubMed: 19101212]
14. Kokmeyer DJ, van der Wurff P, Aufdemkampe G, Fickenscher TCM. The reliability of multitest regimens with sacroiliac pain provocation tests. J Manipulative Physiol Ther 2002;25(1):42–48. [PubMed: 11898017]

15. Mitchell, FL.; Mitchell, PKG. The muscle energy manual. 1st ed.. MET Press; East Lansing, Mich: 1995.
16. Tullberg T, Blomberg S, Branth B, Johnsson R. Manipulation does not alter the position of the sacroiliac joint. A roentgen stereophotogrammetric analysis. *Spine* [sacroiliac] 1998;23:1124–1128. discussion 1129.
17. Stureson B, Uden A, Vleeming A. A Radiostereometric Analysis of the Movements of the Sacroiliac Joints in the Reciprocal Straddle Position. *Spine* 2000;214–217. [PubMed: 10685486]
18. Haneline MT, Young M. A Review of Intraexaminer and Interexaminer Reliability of Static Spinal Palpation: A Literature Synthesis. *J Manipulative Physiol Ther* 2009;32(5):379–386. [PubMed: 19539121]
19. Stovall B, Kumar S. Anatomical Landmark Asymmetry Assessment in the Lumbar Spine and Pelvis - A Review of Reliability. *PM R* 2010;2:48–56. [PubMed: 20129513]
20. AlEisa E, Egan D, Deluzio K, Wassersug R. Effects of Pelvic Skeletal Asymmetry on Trunk Movement: Three-Dimensional Analysis in Healthy Individuals Versus Patients With Mechanical Low Back Pain. *Spine* 2006;31(3):E71–E79. [PubMed: 16449891]
21. Al-Eisa E, Egan D, Deluzio K, Wassersug R. Effects of Pelvic Asymmetry and Low Back Pain on Trunk Kinematics During Sitting: A Comparison With Standing. *Spine* 2006;31(5):E135–E143. [PubMed: 16508537]
22. Levangie PK. The association between static pelvic asymmetry and low back pain. *Spine* 1999;24(12):1234–1242. [PubMed: 10382251]
23. Holmgren U, Waling K. Inter-examiner reliability of four static palpation tests used for assessing pelvic dysfunction. *Man Ther* 2008;13(1):50–56. [PubMed: 17210266]
24. Kmita A, Lucas NP. Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain. *Int J Osteopath Med* 2008;11(1):16–25.
25. Tong HC, Heyman OG, Lado DA, Isser MM. Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position. *J Am Osteopath Assoc* 2006;106(8):464–468. [PubMed: 16943516]
26. Fryer G. Factors affecting the intra-examiner and inter-examiner reliability of palpation for supine medial malleoli asymmetry. *Int J Osteopath Med* 2006:58–65.
27. Fryer G, McPherson H,C, O’Keefe P. The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation. *Int J Osteopath Med* 2005:131–138.
28. Degenhardt BF, Snider KT, Snider EJ, Johnson JC. Interobserver reliability of osteopathic palpatory diagnostic tests of the lumbar spine: improvements from consensus training. *J Am Osteopath Assoc* 2005;105(10):465–473. [PubMed: 16314679]
29. Fossum C, Snider E, Fryer G, Gillette N, Degenhardt B. The introduction of a novel approach to the teaching and assessment of osteopathic manipulative medicine assessment skills. *Int J Osteopath Med* 2008;11(4):165–165.
30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174. [PubMed: 843571]
31. Bogduk, N.; McGuirk, B. Pain research and clinical management. Elsevier; Edinburgh: 2006. Management of acute and chronic neck pain. An evidence-based approach.

Figure 1. The pelvic model (invented and designed by Dr. Shrawan Kumar). I: Examination Table. II: Mechanism of Movement. III: Stretched Skin. IV: ASIS Distance. V: Measuring Scale. VI: Foam Layer. VII: Measuring Scale.

Figure 2.

Calibration plots – graduated scale vs. vernier caliper. All values are in millimeters. Graduated Scale = mounted ruler. Vernier Caliper with digital read out was used for measuring actual movement of model. Along the x axis, negative displacement = right > left asymmetry; positive displacement = left > right; zero displacement = level.

Figure 3a, b.

Percent agreement from the right and left side of pelvic model without training. Percent agreement along the y axis. Along the x axis, negative displacement = right > left asymmetry; positive displacement = L > R; zero = level.

Table 1

Group and Individual Accuracy Compared with Gold Standard - Cohen's Kappa

<i>Group Accuracy</i>				
<i>Pre-Training</i>	<i>Status</i>	<i>n</i>	<i>Mean</i>	<i>Range</i>
Left kappa	fellow	5	0.49	0.40-0.61
	physician	5	0.42	0.31-0.54
Right kappa	fellow	5	0.52	0.32-0.75
	physician	5	0.56	0.47-0.68

<i>Post-Training</i>	<i>Status</i>	<i>n</i>	<i>Mean</i>	<i>Range</i>
Left kappa	fellow	5	0.53	0.36-0.75
Right kappa	fellow	5	0.59	0.44-0.71

<i>Individual Accuracy</i>				
<i>Fellow - pre</i>	<i>Left kappa</i>	<i>p-value</i>	<i>Right kappa</i>	<i>p-value</i>
1 R L L	0.41	<0.05	0.35	<0.05
2 R R R	0.61	0.51	0.75	0.26
3 R R R	0.58	<0.05	0.55	<0.05
4 L L R	0.47	0.16	0.64	0.17
5 R R R	0.40	<0.05	0.32	<0.05

<i>Fellow – post</i>	<i>Left kappa</i>	<i>p-value</i>	<i>Right kappa</i>	<i>p-value</i>
1 R L L	0.48	<0.05	0.52	<0.05
2 R R R	0.75	0.11	0.71	0.23
3 R R R	0.36	<0.05	0.44	<0.05
4 L L R	0.45	<0.05	0.70	0.26
5 R R R	0.63	0.05	0.56	0.15

<i>Physicians</i>	<i>Left kappa</i>	<i>p-value</i>	<i>Right kappa</i>	<i>p-value</i>
1 R R R	0.38	<0.05	0.49	<0.05
2 R R R	0.54	<0.05	0.63	0.05
3 R R R	0.43	<0.05	0.54	0.19
4 R R R	0.31	<0.05	0.47	0.88
5 R L R	0.45	0.61	0.68	0.64

Gold standard =calibrated ASIS model setting. n = number of examiners. Left = left sided evaluation. Right = right sided evaluation. Letters following subject number in order from left to right: hand dominance, eye dominance, and preferred side of evaluation. R and L = right and left respectively.

Table 2

Inter-examiner Reliability of Physicians and Fellows - Fleiss's Kappa

<i>Physician</i>		
Assessment	Left	Right
1	0.49	0.58
2	0.43	0.41
3	0.45	0.30
average	0.46	0.43
range	0.43-0.49	0.30-0.58

<i>Fellow</i>	<i>Pre-Training</i>		<i>Post-Training</i>	
Assessment	Left	Right	Left	Right
1	0.38	0.42	0.29	0.58
2	0.32	0.36	0.44	0.34
3	0.32	0.34	0.34	0.52
average	0.34	0.37	0.36	0.48
range	0.32-0.38	0.34-0.42	0.29-0.44	0.34-0.58

Left = left sided evaluation. Right = right sided evaluation.

Table 3

Intra-examiner Reliability of Physicians and Fellows - Fleiss's Kappa

Fellow - pre	Left kappa	p-value	Right kappa	p-value
1	0.33	<0.05	0.41	<0.05
2	0.34	<0.05	0.63	<0.05
3	0.61	<0.05	0.60	<0.05
4	0.19	0.07	0.63	<0.05
5	0.64	<0.05	0.32	<0.05
Average	0.42		0.52	
Range	0.33-0.64		0.32-0.63	

Fellow - post	Left kappa	p-value	Right kappa	p-value
1	0.53	<0.05	0.66	<0.05
2	0.75	<0.05	0.52	<0.05
3	0.56	<0.05	0.28	<0.05
4	0.58	<0.05	0.51	<0.05
5	0.52	<0.05	0.69	<0.05
Average	0.59		0.53	
Range	0.52-0.75		0.28-0.69	

Physician	Left kappa	p-value	Right kappa	p-value
1	0.52	<0.05	0.61	<0.05
2	0.65	<0.05	0.66	<0.05
3	0.39	<0.05	0.79	<0.05
4	0.19	0.06	0.17	0.10
5	0.59	<0.05	0.24	<0.05
Average	0.47		0.49	
Range	0.19-0.65		0.17 - 0.79	

Left = left sided evaluation. Right = right sided evaluation.