Reliability of diagnosis of somatic dysfunction among osteopathic physicians and medical students

Katrine Bengaard, DO, Richard J. Bogue, PhD, W. Thomas Crow, DO

From Family Medicine Residency, Florida Hospital East Orlando, Orlando, FL

Several studies have assessed interexaminer correlation of diagnosis of somatic dysfunction (SD). This study looks at the simple task of palpating the anterior superior iliac spine (ASIS) of both a live and a fixed plastic model to determine whether examination results are reliable. It is expected that osteopathically trained individuals would be able to do this with reasonable accuracy. However, we tested the results of 151 examiners and found low levels of agreement on diagnosis. Furthermore, the fixed models 'ASIS' were set at equal, yet most examiners (89.2%) chose either left or right. Based on these statistically significant results, we can conclude that palpation for symmetry of two paired structures (such as ASIS') is not an accurate way to assess for SD. It is important to have a standardized approach to diagnosis, because comparing one ASIS with the other does not seem to be the best way to teach students how to diagnose.
© 2012 Published by Elsevier Inc.

This study assessed the reliability of the diagnosis of somatic dysfunction (SD) through palpation in a group of osteopathic practitioners. Somatic dysfunction is defined as "impaired or altered function of related components of the somatic system: skeletal, arthrodial, and myofascial structures, and related vasculature, lymphatic and neural elements."[1] Specific components used to identify the presence of SD include one or more of the following: **t**enderness, **a**symmetry, **r**estriction of range of motion, and **t**issue texture change (TART).[1,2] Osteopathic medical students are taught early in their training how to diagnose SD. Despite these definitions, there is often a discrepancy between two practitioners on what the SD diagnosis actually is. Several studies have investigated this inconsistency, and each has shown that there is some degree of disagreement among practitioners.

The palpation of the anterior superior iliac spine (ASIS) is a basic skill that is taught early in osteopathic medical schools. All osteopathic students and physicians are taught to compare the position of one ASIS with the other. If there is a discrepancy in the level of the two ASIS', a SD is considered present (checking for the asymmetry in TART). However, even an apparently simple diagnosis to determine which ASIS is more superior is not always agreed upon among a group of examiners. This study was designed to investigate the degree of concordance among a larger group of examiners regarding their diagnosis of ASIS symmetry and location.

Materials and methods

To investigate the concordance of diagnosis of SD among a group of individuals, 151 osteopathic students, residents, and attendings were recruited to palpate the ASIS' of three different models. The study design was to recruit 50 students, 50 residents, and 50 attendings. However, the actual

Corresponding author: Katrine Bengaard, DO, Florida Hospital East Orlando, Family Medicine Residency, 7975 Lake Underhill Road, Suite 200, Orlando, FL 32822.
E-mail address: kbengaarddo@gmail.com

breakdown of volunteer examiners was 95 students, 26 residents,18 attendings, and 12 unspecified. The examiners then indicated whether they believed that one ASIS was more superior when compared with the other, or whether they were equal.

Data were collected over several evenings at the American Academy of Osteopathy (AAO) Convocation in Little Rock, Arkansas, March 25-29, 2009. Twelve conference participants were recruited to serve as live models. Each completed and signed an informed consent allowing the examiners to palpate their ASIS'. One fixed, plastic model was also used as an independent type of check on interexaminer diagnostic agreement. Volunteer examiners were recruited to participate in this study either by word of mouth or with a poster set up next to the study area. This study was approved by the Florida Hospital Institutional Review Board.

Two live models per shift were positioned lying supine on examination tables. The live models were asked to lie as still as possible, being conscious not to shift their position, thereby potentially changing the position of their ASIS'. There were 25 volunteer examiners palpating each live model at any given time before the models were switched for two new, unpalpated models (12 live models, 6 shifts, up to 25 examiners per live model).

The fixed model was constructed using a plastic pelvis model (American Anatomical Skeleton Torso Model #VC126). This model was fixed to have even innominates using a screw, and the pelvis was then covered with a piece of cotton material and set up on an examination table. Figure 1 shows the actual examination setup.

Volunteer examiners were briefly instructed on the procedure of palpation. It was explained to them that they were to diagnose using static rather than dynamic palpation of the ASIS of both the live and fixed models. They were also asked to examine the model while standing on the same side of the table as the location of their dominant eye. The definition of SD was defined as a greater than or equal to 1 cm difference in the position of the ASIS. If the ASIS' were

Figure 1   ASIS examination area and setup.

within 1 cm of each other, they were to be designated as equal.

The examiners were given a form on which they identified themselves as either a medical student, resident, or attending physician. Once in the examination area, each examiner had up to one minute to palpate the ASIS' of the model and indicate on their form whether they believed the right or left ASIS was superior, or if they were equal. Each examiner palpated two live models as well as the fixed model and marked their forms appropriately (see Appendix 1).

Data were recorded as anonymous aggregated counts for each of 12 live model/examination period combinations for both the live and the fixed models. Analyses included the calculation of the reliability of examination results and comparisons of students with resident or attending physicians using the kappa statistic. Kappa is a statistic for testing the degree to which different raters or examiners agree with each other—beyond chance. Kappa was used to assess the inter-rater reliability of results between examiners in this study. Fleiss' computation for kappa is useful when the assessments of more than two raters are being assessed for inter-rater reliability.[3-5] Statistics were conducted using IBM Statistics SPSS 19 (SPSS, Inc., Chicago, IL).

Results

Concordance in palpation results across examiners of live subjects

The aim of the concordance analyses was to determine the reliability of palpation on live subjects within a convenience sample of palpation-trained medical personnel (osteopathic students, residents, attending physicians). Palpation results across groups were examined for the reliability of agreement. That is, did these trained examiners agree with each other in determining whether the ASIS was left superior, right superior, or equal? There was no "correct" answer, just the diagnosis of each of the examiners. Twelve different human subjects were palpated by 19 to 25 examiners each, producing a total of 278 palpations for which the level of examiner training was known. Table 1 presents the aggregate results for each of the 12 cases, as well as the total and percent for each result (left superior, right superior, or equal).

A casual review of these data might draw a reader to conclude that, at least for some cases, there was a high degree of agreement among examiners. In Case 5, for example, examiners identified ASIS as left superior 79.2% of the time. From a clinical standpoint, 79.2% may or may not be a sufficiently high level of agreement depending on the condition under examination and the importance of an accurate result.

From a statistical standpoint, however, the standard is more straightforward. By chance, 33.3% of examiners

would select left, 33.3% would select right, and 33.3% would select equal. Because we do not know the "correct" result, the degree of concordance is the degree of systematic variation from random chance. What is being tested using Kappa is whether students and physicians trained to use palpation to come up with a diagnosis will arrive at a finding that is consistent with results attained by their peers using the same procedure with the same patient.

Tests of agreement across students and physicians did not support the idea that palpation produces consistent results when assessing the symmetry of the ASIS'. The Fleiss Kappa of 0.028 suggests a very modest degree of consistency that is well below general guidelines for moderate (0.41-0.60), good (0.61-0.80), or excellent (0.81-1.00) inter-rater reliability. Among this group of students and physicians, examinations did not produce consistent results for the 12 cases. Assessing the 190 student ratings produced slightly higher concordance, but that was still well below any indication that palpations were reliable in determining ASIS superiority (Fleiss' kappa = 0.038).

Palpation accuracy relative to a known fixed result

Live subjects may add unmeasured variation to the assessment of inter-rater agreement.[6] This study also asked trained examiners to palpate a fixed model. The research team used a screw to fix the plastic pelvic model with both ASIS' at equal levels. Because the model was set at equal (neither left nor right as superior), perfect accuracy would find that all 151 of the examiners would choose "equal." However, statistical examination of the distribution of results revealed that this did not occur. Only 15 of 139 of the examinations with the fixed, equal model produced a result of equal, as shown in Table 2. These medical students and physicians did not accurately choose equal after palpation when examining a model fixed to be equal. Palpation results

Table 1 Palpation results (n = 278) across 12 cases

	Counts			Percentages		
	Left	Right	Equal	Left	Right	Equal
Case 1	18	5	1	75.0%	20.8%	4.2%
Case 2	14	9	1	58.3%	37.5%	4.2%
Case 3	15	1	8	62.5%	4.2%	33.3%
Case 4	16	5	3	66.7%	20.8%	12.5%
Case 5	19	3	2	79.2%	12.5%	8.3%
Case 6	7	8	9	29.2%	33.3%	37.5%
Case 7	11	6	2	57.9%	31.6%	10.5%
Case 8	8	9	2	42.1%	47.4%	10.5%
Case 9	9	11	3	39.1%	47.8%	13.0%
Case 10	8	13	2	34.8%	56.5%	8.7%
Case 11	13	5	7	52.0%	20.0%	28.0%
Case 12	10	7	8	40.0%	28.0%	32.0%
Total	148	82	48	52.2%	29.5%	17.3%

Table 2 Observed and expected palpation results with the fixed equal model

	Equal	Left or right
Observed results	15	124
Expected results	139	0

Fisher's exact test, one-tailed $p < .0001$.

were very strongly different from the known fixed expected results. For this group of examiners, palpation appears to not be an accurate method of assessing ASIS symmetry.

Students versus residents/attendings in successes with the fixed result

Disentangling the palpation results of students versus resident or attending physicians might reveal that more experienced practitioners could more accurately assess the model fixed at equal symmetry for ASIS. Table 3 shows that 11 of 95 student assessments succeeded in producing a correct result of equal, as did four of the 44 assessments by resident or attending physicians. There were not enough attending volunteers in the sample to do a separate analysis. Therefore, they were analyzed together with the residents. Fisher's exact test demonstrated no statistically meaningful difference between students and residents/attendings in the proportion of results successfully identifying the fixed model as equal.

Discussion

This study was originally conceived after reviewing the article by O'Haire and Gibbons.[7] The article showed that even a simple palpation of three landmarks (posterior superior iliac spine, sacral sulcus, and sacral inferior lateral angle) could not be agreed upon among 10 examiners.

The results of our study show that among a group of medical students, residents, and attendings, palpation for diagnosis of ASIS superiority is not reliable. This may seem puzzling, because we are all taught how to diagnose the innominates during our first semester of osteopathic medical school. However, 151 examiners were unable to consistently agree on the diagnosis of the three models they

Table 3 Comparing palpation results of students to resident and attending physicians

	Equal	Left or right	Total
Students	11	84	95
Residents/attendings	4	40	44
Total	15	124	139

Fisher's exact test, one-tailed $p = .453$.

palpated. Therefore, although the palpation of landmarks such as the ASIS is taught universally in osteopathic medical schools, the possibility that practitioners cannot do it reliably casts doubts on how it is taught, how it is done, and its value in assessing patients.

Several other studies show similar results, concluding that palpation is not very reliable.[8-10] A recent review article published in the Journal of the American Osteopathic Association found that of the nine studies that met their criteria, there was poor interexaminer reliability without training, whereas reliability only improved slightly after some training for lumbar, sacral, and pelvic anatomic landmarks.[11] Of the studies reviewed that looked at ASIS, the Fleiss kappa measure of inter-rater reliability was, at best, in the "good" range, and that was found in only one of the studies.

The Ohio University Heritage College of Osteopathic Medicine (OUCOM) is now using a Virtual Haptic Back (VHB), which offers feedback on force and touch from a virtual system. OUCOM hopes to allow students to better receive and understand the body's responses to forces and touch.[12] Howell et al. studied this virtual system and showed that students taught with a more standardized system develop more accurate diagnostic skills.[13]

It is interesting, but not unique to this study, that no difference was found between student and physician examiners.[10] Even among examiners, there can be differences in examination technique.[14] It is also important to consider that, although all the examiners in our study were familiar with the ASIS landmark, they may not agree on what part of the bone to palpate.[7] The ASIS may be palpated by an examiner hooking his or her thumb under the ASIS protrusion, finding the center of the ASIS protrusion, or perhaps palpating from above.

Some prior research suggests that reliability of diagnostic palpation improves when examiners receive standardized training before the examination.[11,13] However, Degenhardt et al. argue that this improvement is minimal unless the training is substantial (up to 4 months).[15] In this study, examiners were all trained in osteopathic institutions, which implies they learned a similar style. However, no standard training was given before examinations.

One possible reason for low reliability of diagnosis among live models may be that they were, in fact, live. Living tissue is in a constant state of flux.[6] Also, although the live models were asked not to readjust their position, they may have done so inadvertently.[7] Our subjects were asked to lie still on the examination table for up to 30 minutes, whereas other studies had subjects lying for up to 2.5 hours.[7]

It is also important to note that repeated palpation of the subjects (up to 25 palpations) may have influenced the location and symmetry of the ASIS. Although instructed not to motion test, several of the live model subjects reported that the ASIS compression test was used. Any motion induced into the pelvis may have, however inadvertently, delivered treatment that influenced the location of the ASIS.

Such dynamic palpation could certainly have changed the outcome of subsequent examiners' diagnoses. In short, the first examiners may have been palpating a very different pelvis than the last examiner palpating the same subject.

One result of special interest was that, even with the fixed model, the predominant diagnosis chosen was left ASIS superior (presumably indicating either a left posterior innominate or a right anterior innominate). Because left was more often chosen in the live models (148 of the 278 palpations, or 52.2%), it is possible that either our models all had a left superior ASIS or left superior ASIS is a more common diagnosis based on examiner perspective.

J. Gordon Zink, DO, observed a similar finding in his research in the 1970s, when he coined the term *common compensatory pattern* (CCP).[16-18] He noted that 80% of healthy subjects had a left/right/left/right (LRLR) rotational pattern. That is to say, 80% of people he examined that were free of illness or injury would, in fact, have a left superior ASIS. Zink and Lawson explain that "when the patient's pelvis and thorax are not found in the physiologic (even) or the common compensatory pattern (left ASIS superior), the findings are identified as being disparent."*,[19]

Student doctor Sepehri examined 40 Guatemalan patients to determine whether those examined would fit into Dr. Zink's LRLR pattern. She found that only 26% fell into that category, and that 66% did not fit any pattern. She suggests that the observed Zink patterns may only represent the population studied, and that examiner reliability should be assessed and documented in future research.[20]

Assuming that the live models used in our study were healthy, our examiners' frequent finding that left ASIS was superior is in agreement with Dr. Zink's LRLR pattern. However, this does not explain why, in the even, fixed model, they still chose left.

The examiners were asked to stand on the side of their dominant eye; however, there was no instruction on how to determine which eye is dominant. It is therefore likely that several of the examiners were unaware of which eye was dominant and, because the tables were set up right to left, stood on the right side, using their left eye as dominant. Their opinion may have changed had they then palpated from the left side of the table. Although it is of possible importance for future studies, this study was not designed to test this issue.

Another limitation of this study is the sample size. Also, as outlined in the Material and Methods section, we did not fulfill our study design objective, which was to recruit an equal number of students, residents, and attendings. It was particularly difficult to find residents and attendings that would participate, although students seemed more than willing to volunteer. Of course, one of the obvious biases was that we performed our study at the AAO Convocation. Attendees of this gathering have a particular interest in osteopathic manipulation, and should, in theory, be more in

---

*Disparent: Unlike, diverse, of various appearance, as defined in the compact edition of the Oxford English Dictionary.

tune to palpatory skills. Therefore, the findings of this study are even more interesting, because even among those who attended Convocation, data still showed quite poor interexaminer reliability.

Conclusion

In both the live and fixed models, there was low reliability in the diagnosis of somatic dysfunction among trained examiners. Because we are taught that SD is present when asymmetry is palpated, it is important to be in agreement about what type of asymmetry exists. This is considered an objective finding, yet "objective" infers that it is true and would not change if another examiner were also to palpate. Because we are unable to do so reliably, it may be true that asymmetry is not a good way to define SD of the innominates.

References

1. Ward RC, Hruby RJ, Jerome JA, et al: *Foundations for Osteopathic Medicine*. 2nd ed. Philadelphia: Lippincott Williams & Wilkins, 2002

2. DiGiovanna EL: Somatic dysfunction. In: Nelson KE, Glonek T, editors. *Somatic Dysfunction in Osteopathic Family Medicine*. Philadelphia: Lippincott Williams & Wilkins, 2006, pp 35

3. Fleiss JL, Cohen J: The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 33:613-619, 1973

4. Gwet KL: Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 61:29-48, 2008

5. Li MF, Lautenschlager G: Generalizability theory applied to categorical data. Educ Psychol Meas 57:813-822, 1997

6. Galewaler JE: Motion in health and disease. JAOA 69(2):173/101-179/107, 1969

7. O'Haire C, Gibbons P: Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study. Man Ther 5:13-20, 2000

8. Paulet T, Fryer G: Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region. Int J Osteopath Med 12:92-96, 2009

9. Halma K, Degenhardt B, Snider K, et al: Intraobserver reliability of cranial strain patterns as evaluated by osteopathic physicians: a pilot study. J Am Osteopath Assoc 108:493-502, 2008

10. Kmita A, Lucas N: 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 11:16-25, 2008

11. Stovall B, Kumar S: Reliability of bony anatomic landmark asymmetry assessment in the lumbopelvic region: application to osteopathic medical education. JAOA 110:667-674, 2010

12. Bombei B: Students experience virtual reality training to develop palpatory skills. Osteopathic Family Physician News Nov/Dec:24-25, 2008

13. Howell J, Conatser R, Williams R, et al: Palpatory diagnosis training on the virtual haptic back: performance improvement and user evaluations. J Am Osteopath Assoc 108:29-36, 2008

14. McConnell DG, Beal MC, Dinnar U, et al: Low agreement of findings in neuromusculoskeletal examinations by a group of osteopathic physicians using their own procedures. J Am Osteopath Assoc 79:59-68, 1980

15. Degenhardt B, Snider K, Snider E, et al: Interobserver reliability of osteopathic palpatory diagnostic tests of the lumbar spine: improvements from consensus training. J Am Osteopath Assoc 105:465-473, 2005

16. Dalton E: Short right leg syndrome [page 139-153]. Available at: http://www.somatics.de/Dalton/ShortRightLeg1.pdf. Accessed December 28, 2010.

17. Pope RE: The common compensatory pattern: its origin and relationship to the postural model. AAO Journal 13:19-40, 2003

18. Quinn A: The compensatory pattern, as seen in art and osteopathy. AAO Journal 10:21-23, 2000

19. Zink J, Lawson W: An osteopathic structural examination and functional interpretation of the soma. Osteopath Ann 7:12-19, 1979

20. Sepehri DS: Osteopathic survey of somatic dysfunction and Zink compensatory patterns in Sololá, Guatemala [lecture]. Available at: http://www.osteopathic.org/inside-aoa/development/internationl-licensure/Documents/biomea-2009-sepeheri.pdf. Accessed December 28, 2010. [Abstract found in JAOA 110:25, 2010)

Appendix 1. Answer Form

Correlation of Diagnosis of Somatic Dysfunction among Osteopathic Physicians and Medical Students

Dear Physician or Medical Student,

You are invited to be a part of a study that seeks to evaluate whether there is a correlation between the diagnosis of somatic dysfunction in a group of osteopathic practitioners. Osteopathic medical students are taught from their first day how to diagnose SD. However, there is often a discrepancy between two practitioners on what the SD diagnosis actually is. There have been several studies that have investigated this inconsistency, and each time it has been shown that there is some degree of disagreement among practitioners.

Participation in this study will involve palpating the ASIS' of two live models and one fixed model. **Palpation is defined as light touch. Please do not motion test or otherwise move the models**. You will have up to 1 minute to palpate and determine the diagnosis. Please examine the model standing on the side of the table of their dominant eye. The definition of somatic dysfunction will be defined as a greater than or equal to 1 cm difference in the position of the ASIS. If the ASIS' are within 1 cm of each other, please designate them as equal.

The Institutional Review Board at Florida Hospital has approved the study and its procedures and may review records to confirm the study has been conducted with your best interest in mind.

By filling out the form below, you are consenting to participate in this research. If you need additional information, please contact me at 407-303-8683.

Sincerely,

Answer Form

Please mark your answers by circling the most correct response.

I am a: Medical Student   Resident   Practicing Physician

Live Model #1

<u>Diagnosis</u>:   Left ASIS superior   Right ASIS superior   Equal (Neither ASIS superior)

Live Model #2

<u>Diagnosis</u>:   Left ASIS superior   Right ASIS superior   Equal (Neither ASIS superior)

Fixed Model

<u>Diagnosis</u>:   Left ASIS superior   Right ASIS superior   Equal (Neither ASIS superior)

Thank you for volunteering to participate in this research study!