

The lost art of the clinical examination: an overemphasis on clinical special tests

Chad Cook

Duke University, USA

Special tests have assumed the role of the double edged sword or the Catch 22 of the clinical examination. Special tests include imaging, clinical special tests (e.g. Hawkins–Kennedy test for shoulder impingement, a straight leg raise for lumbar radiculopathy, etc.), and laboratory tests that are advocated as diagnostic tools in decision making. The lure (siren if you will) of a special test is enhanced by the challenges of time and the benefits of efficiency during a clinical examination. Yet the potential bias in decision making that each provides often outweighs the benefit of the tool.

Over the past several decades, a notable decline in the capacity of selected musculoskeletal clinicians to perform a skilled clinical examination has occurred and a proportional reliance on clinical special tests, laboratory tests, and imaging has emerged.¹ This form of substitution in examination has led to a number of problems, most notably, an incorrect diagnosis and a more costly management strategy. Of particular risk to this trend are our medical physician colleagues.

It is well documented that physicians have lost clinical examination capabilities over the last several decades.^{1–8} Factors such as time constraints, lack of confidence in one's clinical testing capabilities, and improvements in technology, have perpetuated a focus on laboratory, imaging, and occasionally, representative clinical special tests. Part or most of the decline has been associated with training weaknesses, weaknesses associated with patient history taking, and an overemphasis on organ-based diagnoses and an under-emphasis on physical management. Medical physicians are faced with notable time constraints and are in many occasions, asked to be all things for all patients.

Yet these problems are not solely isolated to medical physicians. Chiropractic physicians have long demonstrated expansive use of imaging methods creating unnecessary costs and exposure.⁹ Overuse of imaging is certainly not unique to chiropractic physicians as the intolerance for uncertainty is apparent across multiple medical disciplines and

specialties. Ordering imaging appears to be the primary consequence of the intolerance.¹⁰

To be fair, an emphasis on laboratory and imaging tests is understandable. Numerous meta-analyses exist that outline the poor capacity of history taking and physical examination tests for capturing selected conditions.¹ Further, time constraints increase our reliance on special tests (quick tests) and the potential for litigation emphasizes the necessity of accuracy. When clinicians have these tests at their capacity, regardless of whether each is fallible or not, the assumed value of the additional testing supports the pursuit of certainty in diagnosis.

Worth noting is that many manual therapists, including numerous manual physical therapists, do not have access to imaging or laboratory tests. Consequently, these clinicians likely rely heavier on clinical special tests in their decision making, and whether this population would also frequent the use of laboratory and imaging tests remains unknown. Physical therapists often use clinical special tests at the beginning of an examination to 'rule out' a condition, or near the end of the examination to 'rule in' a finding. Their reliance on these tests is analogous to the reliance of medical and chiropractic physicians on imaging and laboratory tests and the purpose is the same: information gathering and improving one's accuracy. Accuracy matters.

Fortunately, the language of diagnostic accuracy exists that helps define the value of a given clinical special test in a given situation. Terminology such as *sensitivity*, *specificity*, and *positive and negative likelihood ratios* help provide value to a test, help decide when a test should be used, and assist in discriminating which tests actually help during decision making, or add nothing to the final determination. Sensitivity refers to the percentage of people who test positive for a specific disease among a group of people who have the disease. Specificity refers to the percentage of people who test negative for a specific disease among a group of people who do not have the disease. Positive and negative likelihood ratios are calculated using both sensitivity and specificity, and

are values that influence the post-test probability of decision making.¹¹ A solid foundation in this language is necessary when determining when and how to use clinical special tests.

Recently, Rubenstein and van Tulder¹² reviewed all clinical tests for the spine and stated, 'It is quite remarkable that while many named orthopaedic tests of the neck and low back are often illustrated in orthopaedic textbooks, there is little evidence to support their diagnostic accuracy, therefore their use in clinical practice'.¹² The authors evaluated the tests based on sensitivity, specificity, and the ability of each to alter post-test probability. Conspicuously absent from the short list of recommended tests were: (1) any tests for cervical instability; (2) any thoracic outlet syndrome test; (3) any tests of vertebrasilar insufficiency or cervical artery dysfunction, any test for lumbar instability, and any of the many neurological screening tests such as lower and upper extremity deep tendon reflexes, sensibility testing, and strength testing. The truth is that it would have been easier for me to outline those tests recommended than those tests that were absent, and thus not recommended.

Others have recently denigrated the value of peripheral-based clinical special tests such as those associated with meniscal testing of the knee,¹³ shoulder labrum testing,¹⁴ and hip labrum testing,¹⁵ and textbooks are available that outline the diagnostic value of many special clinical tests used commonly in practice.^{11,16} The common finding among each of the works is the incapacity of the clinical special tests to be used as stand alone findings for significantly altering one's clinical decision making. Adding insult to injury is the potential bias that each test presents by virtue of 'when' it is used during the examination. Most clinical special tests are used near the end of an examination and are prone to bias associated with the *recency* effect.¹⁷⁻¹⁹ The recency effect occurs when clinicians place more credence on a finding that is performed at the very end of the examination. Because of recall bias, expectancy bias, and misuse (i.e. when a test with high sensitivity and low negative likelihood ratio is used at the end of the examination), the recency effect is a demonstrably potent form of bias that can lead to very poor decision making.

Our discussion of the overemphasis on special tests is likely old news to many practicing manual therapists. Yet to some, this finding may be a

refreshing or unacceptable surprise. It is my hope that the clinical examination makes a comeback and receives the emphasis it deserves. JMMT is interested in your skilled clinical examination methods that lead to effective decision making and invite authors to explore this area for publication. There is no better method in exposing the weaknesses of clinical special tests than by building the literature on decision making during the clinical examination.

References

- 1 Feddock CA. The lost art of clinical skills. *Am J Med* 2007;120: 374-8.
- 2 Ramani S, Orlander JD, Strunin L, Barber TW. Whither bedside teaching? A focus-group study of clinical teachers. *Acad Med* 2003;78: 384-90.
- 3 Talley RC. The history and physical - what do we teach? *S D J Med* 1986;39: 29-33.
- 4 Reilly BM. Physical examination in the care of medical inpatients: an observational study. *Lancet* 2003;362: 1100-5.
- 5 Pfeiffer C, Madray H, Ardolino A, Willms J. The rise and fall of students' skill in obtaining a medical history. *Med Educ* 1998;32: 283-8.
- 6 Vukanovic-Criley JM, Criley S, Warde CM, Boker JR, Guevara-Matheus L, Churchill WH *et al.* Competency in cardiac examination skills in medical students, trainees, physicians and faculty. *Arch Intern Med* 2006;166: 610-6.
- 7 Ahmed MK. What is happening to bedside clinical teaching? *Med Educ* 2002;36: 1185-8.
- 8 Wilkerson L, Lee M. Assessing physical examination skills of senior medical students: knowing how versus knowing when. *Acad Med* 2003;78(10 Suppl.): S30-2.
- 9 Bussieres AE, Peterson C, Taylor JA. Diagnostic imaging practice guidelines for musculoskeletal complaints in adults - an evidence-based approach: introduction. *J Manipulative Physiol Ther* 2007;30: 617-83.
- 10 Lysdahl KB, Hofmann BM. What causes increasing and unnecessary use of radiological investigations? A survey of radiologists' perceptions. *BMC Health Serv Res* 2009;9: 155.
- 11 Cook C, Hegedus E. *Orthopedic physical examination tests: an evidence-based approach*. Upper Saddle River, NJ: Prentice Hall; 2008.
- 12 Rubinstein SD, van Tulder M. A best-evidence review of diagnostic procedures for neck and low back pain. *Best Pract Res Clin Rheum* 2008;22: 471-82.
- 13 Hegedus EJ, Cook C, Hasselblad V, Goode A, McCrory DC. Physical examination tests for assessing a torn meniscus in the knee: a systematic review with meta-analysis. *J Orthop Sports Phys Ther* 2007;37: 541-50.
- 14 Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT 3rd *et al.* Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. *Br J Sports Med* 2008;42: 80-92.
- 15 Leibold MR, Huijbregts PA, Jensen R. Concurrent criterion-related validity of physical examination tests for hip labral lesions: a systematic review. *J Man Manip Ther* 2008;16: E24-41.
- 16 Cleland J. *Orthopaedic clinical examination: an evidence based approach for physical therapists*. Carlstadt, NJ: Icon Learning Systems; 2005.
- 17 Hogarth R, Einhorn H. Order effects in belief updating. The belief adjustment model. *Cogn Psych* 1992;24: 1-55.
- 18 Summerton N. The medical history as a diagnostic technology. *Br J Gen Pract* 2008;58: 273-6.
- 19 Healy AF, Havas DA, Parkour JT. Comparing serial position effects in semantic and episodic memory using reconstruction of order tasks. *J Memory Language* 2000;42: 147-216.