

Studies of quality and impact in clinical diagnosis and decision-making

Eric J. Hegedus

Duke University, USA

First, I'd like to thank Dr Cook for inviting me to write this guest editorial and commend him, as always, on his thought-provoking writing. Although I think that, in many cases, the research on special tests is more to blame than the tests themselves and that these tests should remain part of a skilled clinical examination, I agree with the message: JMMT and other journals need high-quality diagnostic studies since the components of a physical examination influence clinical decision making.¹ High-quality research is vital, but impactful research is equally important and the terms are not synonymous. I would like to add some additional thoughts, if I might be so bold, which may refine submissions so that the research is both of great quality and impact.

Research in the area of components of the physical examination, even when combined in a clinical prediction rule (CPR), is notoriously poor.²⁻⁴ Underpowered and low-quality studies create errors in clinical decision-making and unfortunately, in my personal investigations, I have found mostly underpowered, low-quality studies full of bias.^{5,6} Therefore, there is ample opportunity to investigate the validity of individual tests and CPRs and guidance on how to design and report quality studies has never been more abundant.

More than any previous time, documents exist to help design studies that minimize bias and improve external validity and generalizability.^{1,7-9} These documents are outlined well elsewhere but a brief mention of two is worthwhile. The Standards for Reporting of Diagnostic Accuracy (STARD)^{7,8} initiative produced a 25-item checklist that is the seminal work providing guidance when publishing research about the diagnostic accuracy of individual tests and measures. The 18-item adapted checklist from Beneciuk *et al.*⁴ is actually proposed as a list used to judge the quality of published CPRs but could be effectively used to design a quality CPR study as well. This particular checklist takes into account study design features germane to physical therapy.

However, experts caution that the traditional design of detecting pathology as if it is always a

'have it' and 'don't have it' proposition when compared to a definitive criterion standard that is either 'positive' or 'negative' will need to be adapted.¹⁰ There are many reasons to modify this traditional design. Some pathologies, such as those labelled 'syndrome', do not have a definitive criterion standard. Clinically, many physical examination tests generate non-specific or equivocal results and yet, this is rarely reported in research articles. Finally, in some cases like a degenerative torn meniscus, diagnosis may not even be the most relevant question. Tests that determine functional status, fall potential, or need for surgery may be far more interesting clinical questions.

Beyond diagnosis, tests and measures exist to help predict an outcome (prognosis), and to help focus interventions.¹ A test or tests with one or more of these three qualities has the greatest impact on everyday practice when that test or tests are validated in a low-bias, high-quality study. If a test or measure does not help diagnose more efficiently, help predict an outcome, or help focus treatment, then the test has no utility, no matter whose surname is attached to it, and the resultant research has minimal, if any, impact.

With further regard to impact, I would paraphrase and echo the words of Dinant *et al.*¹¹ and beseech researchers to stop trying to differentiate one form of non-sinister pain from another and, instead, focus on the determinants of success or failure. This statement applies directly to areas of musculoskeletal therapy like low back pain and shoulder pain where, once serious pathology has been ruled out, there is very little relationship between the pathology-based diagnostic label and treatment effectiveness or outcome.¹¹ In other words, let's not waste our time trying to detect a small versus medium versus large rotator cuff tear and focus instead on the variables that predict which of these patients will benefit from physical therapy or surgery.

There is great room for improvement in investigations of clinical decision-making in diagnosis, prognosis, and intervention. Fortunately, there has never been more information on how to successfully

design and conduct quality research. As researchers and clinicians partner to make the necessary improvements in quality, I would ask that we also remember that impact on daily practice is of equal importance.

References

- 1 Childs JD, Cleland JA. Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Phys Ther* 2006;86: 122–31.
- 2 Knottnerus JA, van Weel C, Muris JW. Evidence base of clinical diagnosis: evaluation of diagnostic procedures. *BMJ* 2002;324: 477–80.
- 3 Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA* 1995;274: 645–51.
- 4 Beneciuk JM, Bishop MD, George SZ. Clinical prediction rules for physical therapy interventions: a systematic review. *Phys Ther* 2009;89: 114–24.
- 5 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.
- 6 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* 2007;42: 80–92.
- 7 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. *Ann Intern Med* 2003;138: 40–44.
- 8 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, *et al.* The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003;138: W1–12.
- 9 May S, Rosedale R. Prescriptive clinical prediction rules in back pain research: a systematic review. *J Man Manip Ther* 2009;17: 36–45.
- 10 Feinstein AR. Misguided efforts and future challenges for research on “diagnostic tests”. *J Epidemiol Community Health* 2002;56: 330–2.
- 11 Dinant GJ, Buntinx FF, Butler CC. The necessary shift from diagnostic to prognostic research. *BMC Fam Pract* 2007;8: 53.