

# How accurate are manual assessment methods?

When interexaminer palpation skills of manual therapists and practitioners are examined, what emerges often seems less than impressive.

*JBMT*, following normal peer review procedures, requested a detailed review of the research methodology utilized in the palpation study reported on in this issue by McPartland & Goodridge (pp. 173–178).

The reviewer, a senior academic research professional (who is not a bodyworker), suggested a few minor research-related clarifications, which were duly passed to the authors, who reworked their text accordingly. An unrelated comment in the covering letter from that reviewer, however, caused a reflective pause:

Given the major part of the paper concerns two experienced examiners, their results did not strike me as being particularly impressive ... the agreement was only judged as fair, with two results judged as moderate. Maybe this is reasonable, though given the experience of the practitioners I would have expected the agreement to be greater.

When the comments of *JBMT*'s research reviewer were passed to the authors, a response was forthcoming, which included the words 'In fact, the results are far better than most reported in the literature; interrater reliability is notoriously low within the field of manual medicine.'

The results in this study showed that when using Jones's strain-counterstrain tender points palpation on symptomatic patients, the interexaminer reliability of two extremely experienced practitioners was 72.7% (agreement,  $\kappa = 0.45$ ), which in comparison with other palpation

assessments is more than respectable, and better than most.

Nevertheless, our reviewer's comment illustrates just how subjective many palpation methods appear to be. And while a misread assessment involving strain-counterstrain methods of treatment would carry no risks whatever (since the approach is essentially extremely gentle) what repercussions could result from a mis-assessment where more vigorous manipulative attention was subsequently applied? What degree of risk attaches to the relative weakness seemingly inherent even in skilled manual palpation?

In tests traditionally and widely utilized by osteopaths, chiropractors and physical therapists, prior to the employment of high velocity thrust and other mobilization techniques for example, the following results were achieved (by experienced experts) in this same study:

Interexaminer reliability, in symptomatic patients, for restriction of motion (ROM) was 66.7% ( $\kappa = 0.344$ ), for palpation of joint capsule tenderness (JT) the figures were 76.9% ( $\kappa = 0.529$ ) and for assessment of local tissue texture change (TTC) they were 70.4% ( $\kappa = 0.190$ ).

A composite figure for these three assessment methods of 67.5% ( $\kappa = 0.38$ ) was noted (involving symptomatic patients).

While such results might surprise non-bodyworkers the fact is that based on the evidence of previous interexaminer studies such figures may represent the best that can be hoped for even from skilled professionals with many years of practice. So where does this leave less experienced therapists/practitioners? And more importantly where does this leave their patients?

There is possibly some perverse consolation in knowing that a great deal of medical (even hospital-based) diagnosis proves – often at autopsy – to have been wildly inaccurate. However, where a simple question is being asked, such as 'Is this joint restricted in its range of motion?' or 'Is there a trigger point in these tissues?', should we not expect that accuracy and reliability of assessment/diagnosis would prove more impressive than the evidence suggests?

Many would argue that single assessment methods are seldom used in isolation, and that a picture of just what is happening in particular tissues and joints is gained by gathering composite evidence from a range of assessment protocols and methods. Of course, this will be valid only for those practitioners and therapists who assiduously apply themselves to the task of understanding just what has happened to dysfunctional tissues.

Examination of a region could involve all or any of the following: soft tissue and bony palpation; range of motion assessment; evaluation of structural and postural asymmetry and tissue texture changes; identification of areas of tenderness, atrophy, hypertrophy, weakness, shortness – all using functional and structural evaluation methods including palpation, motion palpation, gait analysis, etc. Out of such a sequence, even if individually their interexaminer reliability is less than optimal, a compilation of findings and indications would result in a reasonable degree of accuracy in diagnosis. The lesson is that single assessment methods are almost never adequate on their own to determine dysfunction or to identify appropriate treatment methods.

**Leon Chaitow**  
Editor