Title:

The validity of lumbo-pelvic landmark palpation by manual practitioners: a systematic review

Author details:

Natalie Alexander - natalliealexander@eso.ac.uk [a]

Alan Rastelli - alanrastelli@eso.ac.uk [a]

Tamsyn Webb - tamsynwebb@eso.ac.uk [a]

Dévan Rajendran - devanrajendran@eso.ac.uk[a,b]

[a]Author Affiliation:

European School of Osteopathy

Research Department

Boxley House

The Street

Boxley

Kent

ME14 3DZ

[b]Corresponding author

Title

The validity of lumbo-pelvic landmark palpation by manual practitioners: a systematic review

Abstract

Background

Lumbo-pelvic landmark palpation is widely used by manual practitioners as part of a clinical assessment. However, research has not clearly described or established the validity of landmark palpation as a diagnostic tool; we aimed to analyse and synthesise data on the validity of lumbo-pelvic landmark palpation compared to an imaging modality reference standard.

Methods and Results

Nine studies met the inclusion criteria (43 practitioners assessing 364 subjects). Mean palpatory accuracy ranged from 42-71%; four studies reported kappa values, ranging from 0.20 (no agreement) to 0.81 (strong); two studies could not be directly compared. The use of multiple palpation techniques improved palpation accuracy, however, no specific technique or technique combination was more accurate than another. Methodological quality was assessed using the QUADAS-2 check-list tool; eight studies were rated as having a risk of bias and/or concerns regarding applicability.

Conclusion

The use of lumbo-pelvic landmark palpation does not reach clinically acceptable levels of validity; however, heterogeneity in the included studies' methods and data

reporting limited between trial comparisons. These results accord with other reviews and suggest the necessity of a shift away from over-reliance on landmark palpation to align manual therapy with practise guidelines.

Keywords

Accuracy, lumbar spine, lumbo-pelvic, manual practitioner, palpation, pelvis, reference standard, validity

Implications for Practice

- Validity of landmark palpation by manual practitioners is rated as 'clinically unacceptable'
- Despite its prevalence, lumbo-pelvic landmark identification may be irrelevant to the diagnosis and management of non-specific LBP (NSLBP)
- Aligning manual therapy and NSLBP management with practise guidelines requires a shift away from over-reliance on landmark palpation
- Lumbo-pelvic landmark palpation may be appropriate in specific scenarios (e.g. spondylolisthesis detection), and the use of multiple palpation techniques can improve accuracy

BACKGROUND

Low back pain (LBP) affects approximately one-third of the adult UK population [1–3], More than 70% of instances of LBP are attributed to non-specific (musculo-ligamentous) causes [4,5]. European and UK national guidelines recommend manual therapy as a component of acute and chronic non-specific LBP management [1,6,7].

Osteopaths, in common with other manual practitioners (e.g. chiropractors and physiotherapists), use palpation as a component of a lower back or pelvic clinical assessment [8–14]. Indeed, the ability to 'interpret palpation findings' is explicitly required of osteopaths on the statutory register in the UK [15]. Palpation is used clinically as a diagnostic test to identify and evaluate motion, symmetry, location, position of anatomical structures [12,16] and to establish mechanical sensitivity of tissues (pain provocation) [17,18]. Manual practitioners may use results from palpation in additional ways: informing a diagnosis; guiding the selection of treatment techniques; determining the application of those techniques; assessing treatment effects and facilitating inter-professional communications [10,19–23]. Palpation is also used by anesthesiologists to identify vertebral segments before administering spinal anesthesia or lumbar epidural injections [10,24,25]. Using palpation in any of the above-mentioned roles relies on the concept that manual palpation is both a reliable and valid diagnostic test.

Ideally, any clinical diagnostic should be both highly valid and highly reliable; terms that are often incorrectly interchanged [13]. Validity is the measure of how well a test compares to the findings of a reference or 'gold' standard and describes whether the test measures what it intends to measure [26,27]. Validity also encompasses test

accuracy, describing how 'true' the test is, or the degree of precision [28]. Reliability measures the extent of agreement between repeated applications of that test, be it from a measuring instrument or a clinical interpretation of a test such as palpation; i.e. are the results reproducible [27,29]. A reliable test, however, is not necessarily a valid test (Figure 1). Some authors have called for clinical tests that lack the necessary validity or reliability to be regarded as 'not useful' [27].

[Figure 1 around here - Validity and reliability [30]]

Research on the reliability of palpation has revealed some general trends: systematic reviews generally report low intra-examiner and/or inter-examiner palpation reliability, with 'clinically unacceptable' kappa levels ($\kappa < 0.40$) [12,17,18,31,32]. The only exception is using palpation for pain provocation, which demonstrates clinically acceptable reliability with 'moderate agreement' ($\kappa \geq 0.40$) [33] [12,17,18].

Study characteristics such as multiple testing, examiner training or using symptomatic subjects appear to have little influence on reliability [17,18]. Literature reviews on the reliability of palpation note the overall heterogeneity of study design and methods; *e.g.* variable testing procedures, non-standardised terminology and/or statistical analyses. [17,18]. Heterogeneity of these data complicates comparison between studies and may lead to conflicting results [18].

In comparison, the amount of research on the validity of palpation is limited, which is likely due to the requirement of using a gold standard benchmark or comparator [34]. Appropriate 'gold standard' comparators for vertebral column palpation include

radiography, computed tomography and magnetic resonance imaging [35]. However, there are several challenges in using these modalities including equipment availability, expense and potential ethical issues related to ionising radiation exposure [34]. *Sui Generis* mechanical models are an alternate proxy gold-standard for testing palpation validity, however, using mechanical models limits comparability and has poor external validity when applied to human subjects [13].

Reviews on the validity of palpation report 'poor' to 'moderate' evidence for motion palpation and pain diagnostic tests [13,36,37] but 'moderate' to 'strong' evidence for cervical spine palpation and lumbar spondylolisthesis detection [35,38]. Both of these latter studies used imaging as the gold standard.

Threshold values deemed to be 'clinically acceptable' have also altered over time. Historically, Landis and Koch's [33] interpretation of Cohen's kappa was the established criteria for results interpretation, with κ values 0.41– 0.60 described as 'moderate', and 'clinically acceptable'. In 2012, McHugh [39] argued that this was too lenient since it allows for less than 40% agreement amongst raters to be described as 'moderate'; flipping a coin would have a higher probability of agreement (50%). McHugh [39] also argued that an inter-rater kappa below 0.60 was inadequate, and little credence should be given to the study results. Likewise, she highlighted that a percentage agreement of up to 61% agreement was similarly problematic, as it implied that a 40% error was deemed clinically acceptable [39]. McHugh [39] therefore recommended 80% agreement as a minimum threshold for inter-rater agreement, despite being necessarily somewhat arbitrary [31]. Therefore, for this

review, we have adopted a clinically acceptable threshold for kappa as ≥ 0.6, and percentage agreement as ≥ 80%.

As far as we know, this is the first review to assess the validity of lumbo-pelvic landmark palpation, apart from lumbar spondylolisthesis detection [38]. Therefore, this study aimed to identify, review and present data from studies that assessed the validity of lumbo-pelvic landmark palpation by manual practitioners, using imaging as a 'gold standard' reference. We have also adopted McHugh's thresholds for 'clinically acceptable' validity (κ ≥ 0.6; percentage agreement ≥ 80%). Where studies have used Landis and Koch's thresholds, we convert and present McHugh's benchmark alongside to allow readers to make meaningful comparisons.

METHODS

Search Strategy

Two independent reviewers (NA, AR) searched the following databases: BioMed Central, Cochrane Library, Index to Chiropractic Literature (ICL), MEDLINE-PubMed, and Science Direct from 1990 to 2018. Citation tracking was used to identify any other trials. Table 1 details the keywords used to build the search strings and Table 2 details our inclusion and exclusion criteria.

[Table 1 Search keywords]

Study Selection

All studies identified were uploaded to Rayyan (Rayyan, Qatar Computing Research Institute, Qatar), a systematic reviewing web application. Within Rayyan, two authors

(NA and AR) independently performed two screening processes: 1) title and abstract based removal of duplicates and non-English texts; 2) title and abstract screening based on exclusion criteria (Figure 2). The two reviewers discussed disagreement to reach consensus; a third assessor (DR) was available if consensus could not be reached. Full-text of remaining studies were subjected to double-blinded screening based upon the criteria and those studies finally included in this review were then subjected to quality assessment.

[Table 2 around here- Inclusion and exclusion criteria]

Quality Assessment

This review used the Quality Assessment of Diagnostic Accuracy Studies check-list tool (QUADAS-2), and signalling questions were tailored according to assessment guidelines and piloted [40,41]. NA and AR, blinded to the journal, title, and author, independently scored all included studies using QUADAS-2 and consensus obtained as above.

Data Abstraction

NA and TW double abstracted the following data: study type; study aim (landmark assessed); participants; index test procedure; reference standard details (Table 5) as well as *a-priori* outcome measures, palpation accuracy and/or margin of error; accuracy sector; kappa values if reported (Table 6). Any differences were resolved by discussion and DR was a third assessor if necessary.

McHugh's interpretation of kappa is as follows [31,42]: κ values 0–0.20 'no agreement'; 0.21–0.39 'minimal'; 0.40–0.59 'weak'; 0.60– 0.79 'moderate'; 0.80–0.90 'strong'; > 0.90 'almost perfect'. Where studies reported Landis and Koch's [43] interpretation, we reported these data, which allows comparison to other data in published literature (Table 6). Landis and Koch κ values are interpreted as: ≤ 0 'poor'; 0.00–0.20 'slight'; 0.21–0.40 'fair'; 0.41– 0.60 'moderate'; 0.61–0.80 'substantial'; 0.81–1.00 'almost perfect' [33].

RESULTS

Study Selection

NA and AR independently identified 499 studies; 485 were excluded at first pass, full text was obtained for 14 studies. Five were excluded at second pass and citation tracking identified no further studies. Thus, a total of nine studies comprising 43 practitioners assessing 364 subjects were quality scored, evaluated for risk of bias and entered into this review (Figure 2).

[Figure 2 here]

Quality checking

Accuracy and inter-rater reliability of the double abstraction were calculated as κ = 1, 'perfect agreement'. Following discussion of scoring discrepancies and eventual consensus, QUADAS-2 scores were calculated as κ = 1, 'perfect agreement'. QUADAS-2 scores and conclusions for each paper are found in Tables 3 and 4.

Table 3 QUADAS-2 faces

Table 4 QUADAS-2 assessment conclusions

Assessment area

All nine studies assessed the accuracy of lumbar spine landmark identification and two also assessed the accuracy of pelvic landmark identification [34,44]. Study sample sizes were generally small, ranging from one to 81 subjects [44,45]. Practitioner sample sizes ranged from one to 16 practitioners [14,45].

Table 5. Study Data

Methodology and subjects

All studies used a cross-sectional design or did not specify study design; Table 5 outlines study data details. In total the nine studies included 280 symptomatic subjects [11,14,25], nine asymptomatic subjects [16] and the remaining 194 subjects did not have their symptom status described [10,34,44–46]. Forty-three practitioners palpated the subjects; six studies used practitioners with three years or more experience. Two studies [10,45] used undergraduate manual practitioners; one used an osteopath who had re-trained as a spinal physician plus a fourth-year medical student [25].

Outcome Measures

Seven studies defined an 'accuracy sector' (Table 6) to determine correct landmark marker placement by practitioners [10,11,14,16,25,44,46]. Accuracy was thus a binary function, where palpation findings were marked as either correct (marker within the sector) or not correct. Correct findings were reported as a percentage of the total number of palpation procedures performed (percentage accuracy). The two

remaining studies did not define an 'accuracy sector' or report percentage accuracy; only reporting the margin of error or distance between the inferior surface of the relevant landmark and landmark marker [34,45]. Table 6 outlines study outcome measures and results. Further details on the reference standards and methods can be found in Table 5.

Table 6. Study results

Reported percentage accuracy

The reported mean accuracy of lumbo-pelvic landmark identification ranged from 42% [46] to 71% [16] (refer to Table 6). One study reported test procedure or technique combination had no predictive influence on accuracy [11], while others reported technique combination or the use of multiple landmarks enhanced palpation accuracy [10,14,16]. Two studies considered pelvic landmark identification and reported a mean margin of error ranging from 2.79mm to 4.60mm [34] or a reported percentage accuracy ranging from 23-68% [44]. Ferreira et al. (2017) also reported a 'propagation error', related to sequential landmark palpation where inaccuracies in locating landmarks are accumulated to the next target landmark. For example, when using a sequential palpation routine, an error in identifying the posterior superior iliac spines may result in subsequent erroneous identification of the S2 and L5 spinous processes.

DISCUSSION

This review identified and analysed nine cross-sectional studies that included 364 subjects assessed by 43 practitioners. Gold standard comparators included: X-ray; fluoroscopic imaging; ultrasound; CT scan and positional 'Upright' MRI (pMRI). QUADAS-2 scored eight of the nine studies as having a risk of bias and/or having applicability concerns. All nine studies assessed the accuracy of lumbar spine landmark identification; two studies also assessed the accuracy of pelvic landmark identification. The reported mean accuracy of lumbo-pelvic landmark identification ranged from 42% to 71%; only four studies reported kappa values, and these ranged from no agreement to strong evidence (κ 0.20 – 0.81) (according to McHugh's criteria [39]).

None of the studies included in this review reached the clinically acceptable threshold of 80% agreement. One study achieved an upper threshold of 83% accuracy when three techniques agreed [14]; however, the mean percentage accuracy ranged from 42% to 71%. It is clear that at least two of the studies report a clinically unacceptable mean palpation accuracy; 42% and 47%, respectively [11,46]. Studies reporting more than 50% of palpation findings identifying the wrong landmark raises concerns about using palpation to identify vertebrae levels correctly [42]. Our findings accord with the poor validity reported in other landmark palpation reviews [13,35,36,47].

Furthermore, four of the nine studies reported kappa values, varying from 'no agreement' [25] to 'strong' evidence [10]. Reporting results without a corresponding kappa statistic to correct for chance-agreement can be potentially misleading [48]. Three studies [11,16,44] only reported percentage accuracy and did not calculate a kappa statistic, which may not represent the true palpation accuracy.

Factors potentially influencing palpation accuracy include study subject and practitioner characteristics, and methodology. Six of the studies reviewed investigated subject characteristics (e.g. gender, BMI, and anatomical anomalies) and their influence on study outcomes. Only two reported subject characteristics as having a statistically significant influence on the accuracy of landmark identification [10,44]; however, the effect size was not reported, making it difficult to determine whether the results are clinically meaningful. Snider *et al.* [10] also reported that practitioner experience had a statistically significant ($p \leq .03$) association with palpation accuracy; however, they compared three qualified practitioners each trained to teach palpation to one undergraduate practitioner. Arguably this represents two extremes of practitioner experience, and when combined with such a small sample size and without a reported effect size, this finding may not be generalizable. Chakraverty *et al.* [25] found that there was a statistically significant learning effect ($p < 0.05$) taking place during the study: the experienced practitioner improved their accuracy on the next round of palpation compared with the novice practitioner. Chakraverty *et al.* [25] attributed this learning effect to poor methodology, as both practitioners saw the results of the fluoroscopic reference standard performed following their palpation.

Comparability of study results is necessary for valid and clinically useful generalizations [17]. However, the levels of heterogeneity in this review limit the comparability of results, an issue also reported in other validity reviews [13,35,38]. Bossuyt *et al.* [49] describe the interpretation of findings based on heterogeneous investigations suffering several limitations, including problems inferring causal

relationships, and difficulties with confounding between characteristics. The small sample of studies in this review further compound these effects as the statistical power of comparison is dependent on the number of studies with a common finding, and estimate precision [49].

Sources of heterogeneity in this review included differences in study methods and outcome measures. Using a larger accuracy sector [25,46], and variation in the size of markers used to mark palpation findings (from 20 mm [44] to 2.5 mm [10], or not reported) could distort reported findings, resulting in an overestimation of palpation accuracy. Additionally, two studies used a different outcome measure [34,45]; therefore, only their general study characteristics and conclusions could be compared. This review did not specify outcome measures in the inclusion criteria for this review, due to the paucity of research on the subject.

Limitations of diagnostic validity studies and this review

Validity studies suffer from several general limitations. Compared to randomised control trials, diagnostic validity studies have greater variability associated with the smaller sample sizes that are often employed [50,51]. Stochkendahl *et al.* [17] described studies of reproducibility (or validity, in this review) facing a quandary between high internal validity and clinical applicability. Studies cannot both optimise accuracy results (e.g. blinding, standardising the procedure, limiting the techniques used) and simultaneously reflect the test's performance in the clinical setting (e.g. flexible procedure, a mix of symptomatic and asymptomatic patients); therefore, these two aspects may need to be investigated separately.

A limitation of our review was the narrow inclusion criteria. Specifying 'manual practitioners' and palpation of the 'lumbo-pelvic' region resulted in this review including only nine studies. However, limiting the study characteristics minimised study heterogeneity and clarified findings, relatively. Measures were taken to minimise review bias, such as the searching of multiple databases by two reviewers, a hand-search of references, and blinding of reviewers and abstractors.

Clinical Utility of Palpation

Despite the widespread use of landmark palpation in LBP diagnosis, our review suggests that lumbo-pelvic landmark palpation by manual practitioners lacks the validity requisite of a clinical tool; we also contend it may not be diagnostically relevant. Due to a failure to identify anatomic or physiologic changes that clearly explain symptoms, non-specific LBP (NSLBP) is the most appropriate diagnosis in the majority of LBP cases [5,52]. As NSLBP is by its very definition "non-specific", specific landmark identification will not contribute to the diagnosis. Furthermore, due to the neurological overlap and complexity of the lumbo-pelvic region and the potential role of biochemical and inflammatory factors in NSLBP [53], valid landmark identification may still inaccurately describe the tissue causing symptoms.

The only valid use of lumbo-pelvic landmark palpation may be in diagnosing specific (rather than non-specific) causes of LBP, such as spondylolisthesis. Alqarni *et al.* [38] report that lumbar spinous process palpation is the optimal diagnostic test for lumbar spondylolisthesis, with high specificity (87%–100%) and moderate to high sensitivity [60–88]. This is not a contradiction in terms, as spondylolisthesis detection combines

both landmark location and identification of a prominent spinous process “step” [38]. However, this is only relevant to a minority of LBP cases. Further research is required to determine if landmark palpation has a role in the management of other specific causes of LBP, such as correlating abnormal imaging reports with anatomical findings.

It is important to note that this review’s conclusions are specifically regarding the subject at hand – i.e. lumbo-pelvic landmark palpation by manual practitioners.Palpation Validity in Anesthesiology

This review did not consider the literature investigating the validity of landmark palpation by anesthesiologists, due to differences in palpation training, methods and access to resources (such as fluoroscopy) compared to manual practitioners. However, the broader scope of research on the subject, including the anesthesiology literature, is of interest and some relevance.

The validity of lumbo-pelvic palpation by anesthesiologists prior to a lumbar puncture varies from 29% to 64% [63–68], with one study achieving 94.9% palpation accuracy in a control group, compared to 58.5% accuracy in a pregnant, obese participant group [69]. Overall, the validity of lumbo-pelvic landmark palpation by anesthesiologists is sufficiently low that anatomical identification using imaging (such as ultrasound) is recommended prior to lumbar puncture [66,69–71].

Interestingly, several studies did not find a statistically significant difference in the accuracy of vertebral interspace identification by manual palpation versus ultrasound imaging [67,68]. However, a systematic review on the subject reported robust

evidence of ultrasound imaging before lumbar punctures increasing the success rate and ease of performance and reducing the risk of traumatic procedures, compared to landmark palpation alone [70].

Implications for manual therapy

Underpinning the prevalence of diagnostic landmark palpation in practice, are a set of profession-wide assumptions and paradigms. Traditional patho-anatomical and biomechanical explanations for NSLBP incompletely describe this condition's aetiology, reinforcing paternalistic, practitioner-centred approaches that are insufficient at best and iatrogenic at worst [52,54,55]. The use of landmark palpation does not imply exclusively biomechanical practice, yet a predominantly biomechanical paradigm prevails, despite robust evidence undermining it [56].

Evidence strongly suggests that NSLBP is associated with a combination of neurophysiological, physical, cognitive factors, and lifestyle behaviours [57]. Therefore, a predominantly biomechanical approach is inappropriate; management should instead reflect the multi-dimensional nature of NSLBP. According to the NICE healthcare guidelines [7], the most rigorous evidence available, manual therapy should be considered in the management of NSLBP, as part of a larger 'treatment package' including exercise and/or psychological therapy. This incorporates biopsychosocial factors known to influence management of NSLBP.

Future Research

It is our opinion that further research into the validity of lumbo-pelvic landmark palpation by manual practitioners will not materially alter these findings, presenting an opportunity and challenge to shift manual therapy away from biomedical and mechanistic paradigms, towards a more integrative management of NSLBP [61].

Implications for practice and education

Accuracy of landmark palpation is particularly important for anesthesiologists before medical procedures, such as spinal anesthesia or lumbar puncture, to prevent spinal cord damage [10,24,25]. The same degree of precision is not required of manual practitioners; however, erroneous landmark palpation can result in inaccurate records, complicating inter-practitioner referral, poor correlation between medical imaging reports and palpation examination findings and potentially inappropriate interventions [62].

A recommendation arising from the results of this review is the use of multiple palpation techniques to improve the accuracy of lumbo-pelvic landmark identification, when used [10,14,16,46] (e.g. spondylolisthesis detection [38]). This accords with recent attempts to optimize the accuracy of individual landmark palpation techniques [62]. This review's recommendation concurs with the results of a previous review [17]; however, no specific technique or technique combination was more accurate than another [14].

CONCLUSION

The research investigated in this review suggests that the validity of landmark palpation by manual practitioners is clinically unacceptable; however, heterogeneity and quality concerns in the literature undermine a definitive conclusion. Nonetheless, even if lumbo-pelvic landmark identification were valid, it could still be considered irrelevant to the diagnosis and management of non-specific LBP (NSLBP). Aligning manual therapy and NSLBP management with practise guidelines requires a shift away from over-reliance on landmark palpation.

REFERENCES

[1] National Collaborating Centre for Primary Care. Low back pain: early management of persistent non-specific low back pain. NICE; 2009.

[2] Ramond-Roquin A, Bouton C, Bègue C, Petit A, Roquelaure Y, Huez J-F. Psychosocial Risk Factors, Interventions, and Comorbidity in Patients with Non-Specific Low Back Pain in Primary Care: Need for Comprehensive and Patient-Centered Care. Front Med 2015;2:73. doi:10.3389/fmed.2015.00073.

[3] Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. Lancet 2012;379:482–91. doi:10.1016/S0140-6736(11)60610-7.

[4] Deyo RA, Weinstein JN. Low back pain. N Engl J Med 2001;344:363–70. doi:10.1056/NEJM200102013440508.

[5] Krismer M, van Tulder M. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best Pract Res Clin Rheumatol 2007;21:77–91. doi:10.1016/j.berh.2006.08.004.

[6] Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 2006;15 Suppl 2:S192-300. doi:10.1007/s00586-006-1072-1.

[7] National Guideline Centre. Low back pain and sciatica in over 16s: assessment and management. 2016.

[8] Nicholas AS, Nicholas EA. Atlas of Osteopathic Techniques. 2nd ed. Lippincott Williams & Wilkins; 2012.

[9] Hertling D, Kessler RM. Management of Common Musculoskeletal Disorders: Physical Therapy Principles and Methods. Lippincott Williams & Wilkins; 2006.

[10] Snider KT, Snider EJ, Degenhardt BF, Johnson JC, Kribs JW. Palpatory accuracy of lumbar spinous processes using multiple bony landmarks. J Manip Physiol Ther 2011;34:306–13. doi:10.1016/j.jmpt.2011.04.006.

[11] Harlick JC, Milosavljevic S, Milburn PD. Palpation identification of spinous processes in the lumbar spine. Man Ther 2007;12:56–62. doi:10.1016/j.math.2006.02.008.

[12] Haneline MT, Young M. A Review of Intraexaminer and Interexaminer Reliability of Static Spinal Palpation: A Literature Synthesis. J Manipulative Physiol Ther 2009;32:379–86. doi:10.1016/j.jmpt.2009.04.010.

[13] Najm WI, Seffinger MA, Mishra SI, Dickerson VM, Adams A, Reinsch S, et al. Content validity of manual spinal palpatory exams - A systematic review. BMC Complement Altern Med 2003;3:1. doi:10.1186/1472-6882-3-1.

[14] Merz O, Wolf U, Robert M, Gesing V, Rominger M. Validity of palpation techniques for the identification of the spinous process L5. Man Ther 2013;18:333–8. doi:10.1016/j.math.2012.12.003.

[15] General Osteopathic Council. Osteopathic Practice Standards. 2012.

[16] Cooper K, Alexander L, Hancock E, Smith FW. The use of pMRI to validate the

identification of palpated bony landmarks. Man Ther 2013;18:289–93. doi:10.1016/j.math.2012.10.005.

[17] Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, et al. Manual examination of the spine: a systematic critical literature review of reproducibility. J Manipulative Physiol Ther 2006;29:475–85. doi:10.1016/j.jmpt.2006.06.011.

[18] Seffinger, Michael A. Najm WI, Mishra SI, Adams A, Dickerson VM, Murphy, Linda S, Reinsch S. Reliability of Spinal Palpation for Diagnosis of Back and Neck Pain: A Systematic Review of the Literature. Spine (Phila Pa 1976) 2004;29:E413–25.

[19] Holmgren U, Waling K. Inter-examiner reliability of four static palpation tests used for assessing pelvic dysfunction. Man Ther 2008;13:50–6. doi:10.1016/j.math.2006.09.009.

[20] Bialosky JE, Simon CB, Bishop MD, George SZ. Basis for spinal manipulative therapy: a physical therapist perspective. J Electromyogr Kinesiol Off J Int Soc Electrophysiol Kinesiol 2012;22:643–7. doi:10.1016/j.jelekin.2011.11.014.

[21] Anders H-L, Corrie M, Jan H, Cuno R, Marianne H, Kristian M, et al. Standardized simulated palpation training-development of a palpation trainer and assessment of palpatory skills in experienced and inexperienced clinicians. Man Ther 2010;15:254–60. doi:10.1016/j.math.2010.01.003.

[22] Lakhani E, Nook B, Haas M, Docrat A. Motion Palpation Used as a Postmanipulation Assessment Tool for Monitoring End-Feel Improvement: A Randomized Controlled Trial of Test Responsiveness. J Manipulative Physiol Ther 2009;32:549–55. doi:10.1016/j.jmpt.2009.08.004.

[23] Sahrmann SA. Diagnosis by the physical therapist--a prerequisite for treatment. A special communication. Phys Ther 1988;68:1703–6. doi:https://doi.org/10.1093/ptj/68.11.1703.

[24] Kim HW, Ko YJ, Rhee WI, Lee JS, Lim JE, Lee SJ, et al. Interexaminer reliability and accuracy of posterior superior iliac spine and iliac crest palpation for spinal level estimations. J Manipulative Physiol Ther 2007;30:386–9. doi:10.1016/j.jmpt.2007.04.005.

[25] Chakraverty RC, Pynsent PB, Westwood A, Chakraverty JK. Identification of the correct lumbar level using passive intersegmental motion testing. Anaesthesia 2007;62:1121–5. doi:10.1111/j.1365-2044.2007.05189.x.

[26] Straus SE, Richardson WS, Glasziou P, Haynes RB. Evidence-Based Medicine: How To Practice and Teach EBM. Edinburgh: Elsevier; 2005.

[27] 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;11:16–25. doi:10.1016/j.ijosm.2008.01.003.

[28] Winter G. A comparative discussion of the notion of “validity” in qualitative and quantitative research. Qual Rep 2000;4.

[29] Fleiss J, Levin B, Myunghee CP. Statistical Methods For Rates and Proportions. 3rd ed. John Wiley & Sons, Inc.; 2003.

[30] Dilmen N. Reliability and validity. Wikimedia Commons 2012. https://commons.wikimedia.org/wiki/File:Reliability_and_validity.svg (accessed October 28, 2015).

[31] May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: A systematic review. Aust J Physiother 2006. doi:10.1016/S0004-9514(06)70044-7.

[32] Cooperstein R, Hickey M. The reliability of palpating the posterior superior iliac spine: a systematic review. J Can Chiropr Assoc 2016;60.

[33] Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. vol. 33. 1977.

[34] Kilby J, Heneghan NR, Maybury M. Manual palpation of lumbo-pelvic landmarks: a validity study. Man Ther 2012;17:259–62. doi:10.1016/j.math.2011.08.008.

[35] Póvoa LC, Ferreira APA, Silva JG. Validation of Palpatory Methods for Evaluating Anatomical Bone Landmarks of the Cervical Spine: A Systematic Review. J Manipulative Physiol Ther 2015;38:302–10. doi:10.1016/j.jmpt.2015.04.002.

[36] Hestbaek L, Leboeuf-Yde C. Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review. J Manipulative Physiol Ther 2000;23:258–75. doi:https://doi.org/10.1067/mmt.2000.106097.

[37] van der Wurff P, Meyne W, Hagmeijer RH. Clinical tests of the sacroiliac joint. Man Ther 2000;5:89–96. doi:10.1054/math.1999.0229.

[38] Alqarni AM, Schneiders AG, Cook CE, Hendrick PA. Clinical tests to diagnose lumbar spondylolysis and spondylolisthesis: A systematic review. Phys Ther Sport 2015;16:268–75. doi:10.1016/j.ptsp.2014.12.005.

[39] McHugh ML. Interrater reliability: the kappa statistic. Biochem Medica 2012;22:276–82. doi:https://doi.org/10.11613/BM.2012.031.

[40] Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–36. doi:10.7326/0003-4819-155-8-201110180-00009.

[41] Reitsma J, Rutjes A, Whiting P, Vlassov V, Leeflang M, Deeks J. Chapter 9: Assessing methodological quality. In: Deeks J, Bossuyt P, Gatsonis C, editors. Cochrane Handb. Syst. Rev. Diagnostic Test Accuracy Version 1.0.0, The Cochrane Collaboration; 2009.

[42] McHugh ML. Interrater reliability: the kappa statistic. Biochem Medica 2012;22:276–82.

[43] Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. vol. 33. 1977. doi:https://doi.org/10.2307/2529310.

[44] Ferreira APA, Póvoa LC, Zanier JFC, Machado DC, Ferreira AS. Sensitivity for palpating lumbopelvic soft- tissues and bony landmarks and its associated factors: A single-blinded diagnostic accuracy study. J Back Musculoskelet Rehabil 2017. doi:10.3233/BMR-150356.

[45] Mieritz RM, Kawchuk GN. The Accuracy of Locating Lumbar Vertebrae When Using Palpation Versus Ultrasonography 2016. doi:10.1016/j.jmpt.2016.05.001.

[46] Robinson R, Robinson HS, Bjorke G, Kvale A. Reliability and validity of a palpation technique for identifying the spinous processes of C7 and L5. Man Ther 2009;14:409–14. doi:10.1016/j.math.2008.06.002.

[47] van der Wurff P, Hagmeijer R, Meyne W. Clinical tests of the sacroiliac joint. A systemic methodological review. Part 1: reliability. Man Ther 2000;5:30–6. doi:https://doi.org/10.1054/math.1999.0228.

[48] Kaufman AB, Rosenthal R. Can you believe my eyes? The importance of interobserver reliability statistics in observations of animal behaviour. Anim Behav 2009;78:1487–91. doi:10.1016/j.anbehav.2009.09.014.

[49] Bossuyt P, Davenport C, Deeks J, Hyde C, Leeflang M, Scholten R. Chapter 11: Interpreting results and drawing conclusions. In: Deeks J, Bossuyt P, Gatsonis C, editors. Cochrane Handb. Syst. Rev. Diagnostic Test Accuracy. Version 0., The Cochrane Collaboration; 2013.

[50] Bachmann LM, Puhan MA, ter Riet G, Bossuyt PM. Sample sizes of studies on diagnostic accuracy: literature survey. Br Med J (Clin Res Ed) 2006;332:1127–9. doi:10.1136/bmj.38793.637789.2F.

[51] Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM. Systematic reviews of diagnostic test accuracy. Ann Intern Med 2008;149:889–97.

[52] Deyo RA. Diagnostic evaluation of LBP: reaching a specific diagnosis is often impossible. Arch Intern Med 2002;162:1444-7; discussion 1447-8.

[53] Chawla, J., Talavera, F., Berman, S.A., Schneck, M.J., Wheeler AH. Low Back Pain and Sciatica: Overview, Pathophysiology, Characteristics of Pain-Sensitive Structures. Medscape 2018. https://emedicine.medscape.com/article/1144130-overview (accessed March 25, 2020).

[54] Darlow B. Beliefs about back pain: The confluence of client, clinician and community. Int J Osteopath Med 2016;20:53–61. doi:10.1016/j.ijosm.2016.01.005.

[55] Thomson OP, Petty NJ, Moore AP. Reconsidering the patient-centeredness of osteopathy. Int J Osteopath Med 2013. doi:10.1016/j.ijosm.2012.03.001.

[56] Figg-Latham J, Rajendran D. Quiet dissent: The attitudes, beliefs and behaviours of UK osteopaths who reject low back pain guidance – A qualitative study. Musculoskelet Sci Pract 2017;27:97–105. doi:10.1016/J.MATH.2016.10.006.

[57] O'Sullivan P. It's time for change with the management of non-specific chronic low back pain. Br J Sports Med 2012;46:224–7. doi:10.1136/bjsm.2010.081638.

[58] Sanders T, Foster NE, Bishop A, Ong BN. Biopsychosocial care and the physiotherapy encounter: physiotherapists' accounts of back pain consultations. BMC Musculoskelet Disord 2013;14:65. doi:10.1186/1471-2474-14-65.

[59] Abrosimoff M, Rajendran D. 'Tell me your story' - How osteopaths apply the BPS model when managing low back pain - A qualitative study. Int J Osteopath Med 2019;0. doi:10.1016/j.ijosm.2019.11.006.

[60] Fryer G. Integrating osteopathic approaches based on biopsychosocial therapeutic mechanisms. Part 2: Clinical approach. Int J Osteopath Med 2017;26:36–43. doi:10.1016/j.ijosm.2017.05.001.

[61] Buchbinder R, van Tulder M, Öberg B, Costa LM, Woolf A, Schoene M, et al. Low back pain: a call for action. Lancet 2018. doi:10.1016/S0140-6736(18)30488-4.

[62] Cooperstein R, Truong F. Would adopting a revised landmark rule for the spinal level of the iliac crests improve the accuracy of lumbar level identification? J Can Chiropr Assoc 2019;63:26–35.

[63] Broadbent CR, Maxwell WB, Ferrie R, Wilson DJ, Gawne-Cain M, Russell R. Ability of anaesthetists to identify a marked lumbar interspace. Anaesthesia 2000;55:1122–6. doi: 10.1046/j.1365-2044.2000.01547-4.x

[64] Furness G, Reilly MP, Kuchi S. An evaluation of ultrasound imaging for identification of lumbar intervertebral level. Anaesthesia 2002;57:277–80. doi: 10.1046/j.1365-2044.2002.2403_4.x

[65] Whitty R, Moore M, Macarthur A. Identification of the Lumbar Interspinous Spaces: Palpation Versus Ultrasound. Anesth Analg 2008;106:538–40. doi:10.1213/ane.0b013e31816069d9.

[66] L D, P N, D K, T Ł. Anatomical Landmarks Based Assessment of Intravertebral Space Level for Lumbar Puncture Is Misleading in More Than 30%. Anaesthesiol Intensive Ther 2013;45. doi:10.5603/AIT.2013.0001.

[67] Hayes J, Borges B, Armstrong D, Srinivasan I. Accuracy of manual palpation vs ultrasound for identifying the L3-L4 intervertebral space level in children. Paediatr Anaesth 2014;24:510–5. doi:10.1111/pan.12355.

[68] Ambulkar R, Patil V, Doctor JR, Desai M, Shetty N, Agarwal V. Accuracy of ultrasound imaging versus manual palpation for locating the intervertebral level. J Anaesthesiol Clin Pharmacol 2017;33:348–52. doi:10.4103/joacp.JOACP_285_16.

[69] Malik M, Ismail S. Accuracy of Tuffier's Line Identification by Palpation Method: Cross-Sectional Comparative Study Among Obese, Pregnant and Control Groups. Turkish J Anaesthesiol Reanim 2020;48:108–14. doi:10.5152/TJAR.2019.82346.

[70] Perlas A, Chaparro LE, Chin KJ. Lumbar Neuraxial Ultrasound for Spinal and Epidural Anesthesia: A Systematic Review and Meta-Analysis. Reg Anesth Pain Med 2016;41:251–60. doi:10.1097/AAP.0000000000000184.

[71] Narkhede HH, Kane D, Parekh V, Hemantkumar I. A cohort study of anatomical landmark-guided midline versus pre-procedure ultrasound-guided midline technique of spinal anesthesia in elderly patients undergoing orthopedic surgery. J Anaesthesiol Clin Pharmacol 2019;35:522–7. doi:10.4103/joacp.JOACP_396_17.

Table 4 QUADAS-2 assessment

STUDY	Risk of Bias	Applicability Concerns
Chakraverty et al. (2007)	**High risk of bias** -Practitioners not blinded to reference standard results throughout index testing	**Concerns regarding applicability** -Small practitioner sample (n=2) -One examiner not a manual practitioner (4th year med student) with relatively short training; limits generalizability to a manual practitioner population
Cooper et al. (2013)	**Low risk of bias**	**Concerns regarding applicability** -Small subject sample (n=9) and a high ratio of female to male subjects (8:1) -Small practitioner sample (n=2)
Ferreira et al. (2017)	**Unclear risk of bias** -Practitioner blinding to reference standard results not stated; insufficient information	**Concerns regarding applicability** -Small practitioner sample (n=2) -Index test (landmark palpation) was the result of consensual agreement between two practitioners, not an individual practitioner's findings – i.e. not representative of clinical practice
Harlick et al. (2007)	**Unclear risk of bias** -Practitioner blinding to other practitioner's index tests or reference standard not stated; insufficient information	**No concerns regarding applicability**
Kilby et al. (2012)	**High risk of bias** -Convenience sample – non-probability sampling from a health-oriented programme	**Concerns regarding applicability** -Small subject sample (n=3) -Convenience sample – non-probability sampling from specific demographic
Merz et al. (2013)	**Low risk of bias**	**Concerns regarding applicability** -Small practitioner sample (n=1)
Mieritz & Kawchuk (2016)	**Unclear risk of bias** -Practitioner blinding to reference standard results not stated; insufficient information	**Concerns regarding applicability** -Small subject sample (n=1) -All practitioners students; average of 2 years palpation experience; index test not representative of a manual practitioner population
Robinson et al. (2009)	**High risk of bias** -Subject exclusion criteria included inability of practitioners to identify iliac crests or palpate spinous process in question	**Concerns regarding applicability** -Small practitioner sample (n=2)
Snider et al. (2011)	**Low risk of bias**	**No concerns regarding applicability**

Table 1. Search terms

#1: Validity Terms	**#2: Spine + Pelvis Terms**	**#3: Procedure Terms**	**#4: Reference Test Terms**
Accuracy, accurate Valid, Validation Validity	Lumbar Lumbar spine Lumbar vertebra Pelvis Sacrum / sacroiliac joint / sacroiliac	Palpate Palpation Diagnostic palpation Clinical palpation Manual palpation	X-rays, x-ray Ultrasound MRI / magnetic resonance imaging Reference standard Gold standard
Search Strategy: (#1 AND #2 AND #3 AND #4) (accuracy OR accurate OR valid OR validation OR validity) AND (lumbar OR lumbar spine OR lumbar vertebra OR pelvis OR sacrum OR sacroiliac joint OR sacroiliac) AND (palpate OR palpation OR diagnostic palpation OR manual palpation OR clinical palpation) AND (x-rays OR x-ray OR ultrasound OR MRI OR magnetic resonance imaging OR reference standard OR gold standard)			

Table 2 - Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
• Primary research studies published in peer-reviewed journals. • Published between 1990 and 2018. • Studies involving regulated manual therapists (e.g. osteopaths, chiropractors, and physiotherapists). • Studies comparing landmarks identified by palpation to a reference standard (e.g. radiologic imaging) and reporting the validity or accuracy of lumbar spine and/or pelvic palpation, where validity is defined as the ability of palpation to assess what it intends to assess, and accuracy as a measure of validity, or the degree of precision of palpation findings. • Studies using direct palpation of structures of the human lumbar spine and pelvis (including sacrum) including associated joints (sacroiliac joints). • N.B. The method used to identify landmarks was not limited, therefore studies using static or motion palpation techniques or other reference landmarks to locate the appropriate anatomical landmark qualified for inclusion.	• Non-English articles • Full text not available;

Table 3. Consensus quality scores using QUADAS-2

STUDY (Date)	Risk of Bias				Applicability Concerns		
	Patient selection	*Index Test*	*Reference Standard*	*Flow and Timing*	*Patient Selection*	*Index Test*	*Reference Standard*
Chakraverty et al. (2007)	☺	☹	☺	☺	☺	☹	☺
Cooper et al. (2013)	☺	☺	☺	☺	☹	☹	☺
Ferreira et al. (2017)	☺	😐	☺	☺	☺	☹	☺
Harlick et al. (2007)	☺	😐	☺	☺	☺	☺	☺
Kilby et al. (2012)	☹	☺	☺	☺	☹	☺	☺
Merz et al. (2013)	☺	☺	☺	☺	☺	☹	☺
Mieritz & Kawchuk (2016)	☹	😐	☺	☺	☹	☹	☺
Robinson et al. (2009)	☹	☺	☺	☺	☺	☹	☺
Snider et al. (2011)	☺	☺	☺	☺	☺	☺	☺

Key: ☺ = low risk; ☹ = high risk; ? = unclear risk.

Table 5 Study Data

Author (date) ***Study type***	**Landmark**	**Participants** (practitioners and subjects)	**Test Procedure** (including subject position)	**Reference Standard** (including position and marker)
Chakraverty et al. (2007) [25] *Study type: (Not reported)*	L5-S1 interspace[1]	Practitioners: -1 "experienced" spinal physician who had undergone formal osteopathic training -1 "novice" practitioner, 4th year medical student who underwent 3-hour training session with experienced practitioner to learn method of landmark identification Subjects: -Patients undergoing fluoroscopically guided lower lumbar spine injections for LBP -35 patients (M=8, F=27) -Mean BMI[2]: 27 kg/m^2 -BMI Range: 19-40 kg/m^2 -Anatomical abnormalities: not reported	-Intersegmental motion testing used to identify interspace between L5/S1 SPs -Practitioner blinded to markings from prior techniques (UV pen), not blinded to actual level identified through fluoroscopy Subject Position: -Sidelying – turned prone to mark identified level	Fluoroscopic imaging Subject Position: -Prone Marker: -Skin marked with UV pen; radio-opaque markers placed on skin overlying marking, marker size not described
Cooper et al. (2013) [16] *Study type: Pilot cross-sectional study*	L4 SP (spinous process)	Practitioner: -2 physiotherapists -Minimum 15 years experience Subjects: Standing: 9 (F:M 8:1) Prone: 8 (F:M 7:1) -Mean BMI: 23.6 kg/m^2 -BMI range: 19.8-28.2 kg/m^2 -Anatomical abnormalities: not reported	Combination of palpation methods: -Palpation in relaxed semi-flexed position -Iliac crests -Sacral base -Practitioner blinded; each practitioner waiting in a separate room whilst the other was locating landmarks and scans were being done. Subject Position: -Standing -Prone	pMRI – positional "Upright" MRI Subject Position: -Standing (for standing palpation assessment) -Prone (for prone palpation assessment) Marker: -12.5mm cod-liver oil capsule with double-sided tape -No change of subject position between palpation and imaging

Ferreira et al. (2017) [44] *Study type: single-blinded diagnostic accuracy study*	L4 SP L5 SP PSIS* (posterior superior iliac spine) Inferior-lateral sacrum angle* Ischial tuberosity* *Left and right Only bony lumbopelvic landmarks from this study were considered – greater trochanter was not included	Practitioners: -2 physiotherapists graduated >25 years, both with a post-graduate Diploma in Osteopathy and subsequently teaching at an osteopathic institution Subjects: -Recruited from patients referred for CT-scan exam -83 patients (M=41, F=42) -Mean BMI 25.9 kg/m^2 -BMI Range: Not reported -Anatomical abnormalities: not reported	-Landmark location was determined by simultaneous, consensual agreement between the two practitioners (i.e. practitioners not blinded to each other's results) -Palpatory procedure used the iliac crests (CI technique) as reference points for sequential location of the anatomical landmarks Subject Position: -Prone	CT scan Two radiologists evaluated images and confirmed by consensual agreement Subject Position: -Supine Marker: -Skin marked with metal ring (internal diameter = 20mm) fixed with adhesive tape
Harlick et al. (2007) [11] *Study type: Quasi-experimental cross-sectional study*	L1, L3, and L5 SP	Practitioner: -5 manipulative physiotherapists -Minimum 10 years experience Subjects: -Recruited from LBP patients referred for lumbar X-ray -75 patients (M=29, F=46) -Each practitioner allocated 15 of the 75 subjects -Mean BMI: not reported -BMI Range: Not reported -Patient-practitioner distribution demonstrated no statistically significant differences (p >0.5) for sex or BMI -Anatomical abnormalities: not reported	Practitioner used choice of technique to locate SPs: -Iliac crests -Posterior superior iliac spines (PSISs) -12th ribs -Passive movement -Practitioner blinding not stated Subject Position: Variable; dependent on practitioner's technique choice	Lateral X-rays Subject Position: -Side-lying; landmarks re-checked after change in position Marker: -Spherical lead radio-opaque markers (7.5mm) fixed by non-allergenic radio-transparent tape (also positioned mid-way between L1 and L3 markers, and L3 and L5 markers, as well as caudally to L5 marker approximately the same distance to allow for geometric analysis)
Kilby et al. (2012) [34] *Study type: Prospective exploratory study*	-Inferior surface of L4 SP -Inferior surfaces of right and left PSIS	Practitioners: -9 physiotherapists -Mean time since physiotherapy qualification <3.5 years recruited from a postgraduate masters in manual therapy programme Subjects:	Practitioners encouraged to palpate caudad to cephalad using thumb to identify inferior surface of L4 SP. Instructed to palpate iliac crests distally until thumbs reached inferior surface of the structure for the PSIS.	Ultrasound Subject Position: -Prone Marker: -Skin marked with UV pen; large paper clip attached to generate acoustic shadow for

		-Convenience sample recruited from postgraduate Masters in Manual Therapy programme -3 models (M=2, F=1) -Mean BMI: 25.3 kg/m^2 -BMI Range: Not reported -Anatomical abnormalities: not reported	-Practitioners blinded -UV pen used for the U/S identified point, washable ink used by practitioner removed between each practitioner landmark assessment Subject Position: -Prone	imaging -No change of subject position between palpation and imaging
Merz et al. (2013) [14] *Study type: Diagnostic validity study*	L5 SP	Practitioners: -1 manipulative physiotherapist (OMT-DVMT®) Subjects: -Recruited from LBP patients referred for X-ray investigation -66 patients (M=34, F=32) >> only 62 included in final analysis -Mean BMI: 27.7 kg/m^2 -BMI Range: Not reported -Anatomical abnormalities: not reported	Four palpation techniques used to identify L5 SP: 1) Iliac crests as reference points - patient standing 2) Inferior margin of PSIS as bony landmarks 3) Motion palpation performed 4) Palpation area circumscribed by marking reference points (iliac crests and PSIS) and then performing motion palpation with visual orientation -Practitioner blinded to markings from prior techniques (UV pen) Subject Position: Varied with technique Standing (1&2) Seated (3&4)	Lateral view X-ray Subject Position: -Standing Marker: -Skin marked with UV pen; 3mm stainless steel pellet radio-opaque markers attached with medical tape -Multiple changes of subject position between palpation and imaging
Mieritz & Kawchuk (2016) [45] *Study type: Validation study*	L3 SP	Practitioners: -16 chiropractic students -Years of palpation experience: 8 students 2nd year 5 students 3rd year 2 students 4th year 1 student 5th year Subjects: -One female patient -BMI: 19.4 kg/m^2 -Anatomical abnormalities: no variation	5 minute instruction video watched by each student before beginning palpation and before beginning ultrasonic imaging on the respective technique -Practitioners were allowed to use any technique they were taught in their training Most commonly Iliac crest was used as a reference landmark. Subject lying prone	Ultrasound Subject Position: -Prone Marker: -Reflective marker used, size not described

		of lumbar segmentation visualised	Subject seated 3) Ultrasonic imaging *Palpation using each method time limited to 1 minute* -Practitioner blinding not stated -Insufficient information regarding practitioner blinding	
Robinson et al. (2009) [46] *Study type: (Not reported)*	C7 and L5 SPs (spinous processes) *only L5 results relevant to this review	Practitioners: -2 Manual Therapists (MTs) -Norwegian postgraduate qualification in manual therapy -Minimum 16 years clinical experience Subjects: -31 patients referred to a particular radiology unit in Oslo for X-ray of lumbar spine (F:M = 21:10) -BMI: 14 x normal BMI 18.5-24.9 kg/m^2 12 x overweight BMI 25-29.9 kg/m^2 5 x obese BMI >30 kg/m^2 -Anatomical abnormalities: not reported	-Practitioners had 3 training sessions with the researcher -Palpation technique was using the iliac crests as a landmark, with L5 identified as the first SP under an imaginary line between the iliac crests (limited to one technique) -Practitioners blinded -UV pen used by both manual therapists when marking the skin, so were blinded to each-others marks on the subjects. Subject Position: -Side-lying	A/P and lateral X-rays Subject Position: -Side-lying Marker: -Skin marked with UV pen; “small magnet” taped to skin for imaging (size not stated) -Subjects changed position between palpation and imaging
Snider et al. (2011) [10] *Study type: Cross-sectional study*	L1, L2, L3, and L4 SP	Practitioners: -3 osteopathic physicians, board-certified specialists in Neuromusculoskeletal Medicine and Osteopathic Manipulative Medicine -1 third year resident in Neuromusculoskeletal Medicine and Osteopathic Manipulative Medicine Subjects: -Recruited from the local community -60 participants (M=27; F=33) -Mean BMI: 25.6 kg/m^2	All practitioners identified L1-4 SPs using the following landmarks: -Sacral base -L5 SP -Tuffier’s line -T12 SP -12th ribs -Practitioners conducted palpation “independently”, markers removed between each practitioner palpation	Posteroanterior X-ray Subject Position: -Prone Marker: -SP marked with 2.5mm SDM-BB25 Spee-D-Mark radiopaque self-adhesive skin markers -No change of subject position between palpation and imaging

-BMI range: 19.2 – 42.7 kg/m^2 (10% of subjects obese) -Anatomical abnormalities present in 23 (38%) of subjects, most commonly a rudimentary 12th rib (unilateral or bilateral) found in 12 (20%) of subjects	Subject Position: -Prone

Key:

[1] While the interspace may not technically be a “landmark”, identification of interspaces requires orientation relative to the lumbar spinous processes, therefore this study was considered applicable to landmark assessment

[2] BMI (body mass index):Normal weight: BMI 18.5 – 24.9; Overweight: BMI 25 – 29.9; Obese: BMI >30

Table 6 - Study results

Study	Landmark	Palpation Accuracy	Accuracy Sector	Margin of Error	Kappa	Kappa Interpretation
Chakraverty et al. (2007) [25]	L5/S1 interspace	Experienced practitioner: 57% Novice practitioner: 54%	Interspace between SP's (If SP identified, nearest interspace recorded)	Not reported	Kappa: 0.20 (95% CI -0.02 – 0.44)	MH* - none LK† - slight
Cooper et al. (2013) [16]	L4 SP	Standing: 72% Prone 69% Mean: 71%	Overlap of any part of the marker with the boundaries of L4 SP	Largest median distance 2.7mm (IQR 0.0-3.6)	Not reported	Not Applicable
Ferreira et al. (2017) [44]	L4 SP L5 SP PSIS *(left and right)* Inferior-lateral sacrum angle *(left and right)* Ischial tuberosity *(left and right)*	-L4 SP: 39% -L5 SP: 37% -PSIS Left: 53% -PSIS Right: 41% -ILA Left: 26% -ILA Right: 33% -IT Left: 68% -IT Right: 61%	If any part of the landmark was located on the images within the area of the metal ring placed over the palpated area, this was classed as accurate identification	Not reported	Kappa not reported CI upper and lower bounds: L4 lower: 28 L4 upper: 49 L5 lower 27 L5 upper 48 PSIS L lower: 42 PSIS L upper: 64 PSIS R lower: 30 PSIS R upper: 52 ILA sacrum L lower: 17 ILA sacrum L upper: 49 ILA sacrum R lower: 1 ILA sacrum R upper: 25 IT L lower: 48 IT L upper: 89	Not Applicable

					IT R lower: 39 IT R upper: 84	
Harlick et al. (2007) [11]	L1 SP L3 SP L5 SP	L1: 49% L3: 41% L5: 49% Mean all levels: 47%	Overlap of any part of the marker with the boundaries of the SP	Mean error for all SP: 19.3mm (SD 18.6)	Not reported	Not Applicable
Kilby et al. (2012) [34]	Inferior border L4 SP Inferior border PSIS (left and right)	Not reported	Participants requested to locate inferior border of each landmark. No boundary provided.	Inferior border L4 SP = 15.63 SD 3.89mm Inferior border Left PSIS 20.07 SD 4.60 mm Inferior border Right PSIS 20.07 SD 4.60 mm	Kappa not reported (ICC = 0.821 p=0.00)	Not Applicable
Merz et al. (2013) [14]	L5 SP	Mean accuracy (all techniques): 68% Palpation techniques: PSIS 61% Manual palpation + visual orientation 61% Manual Palpation: 44% (calculated) Iliac Crest: 47% (calculated)	Marker within the superior and inferior boundaries of the SP	No reported	Kappa: 0.637	MH* - moderate LK† - substantial
Mieritz & Kawchuk (2016) [45]	L3 SP	Not reported	Not reported, though each participant used ultrasound as a standard	Sitting palpation: 32mm (SD 19) Prone Palpation: 26mm (SD 14)	Not reported	Not Applicable

Robinson et al. (2009) [46]	L5 SP	MT‡ 1: 48% MT‡ 2: 36% Mean both MT‡ : 42% (calculated, not reported)	Marker within 'sector' defined by area within lines marked in the middle of each adjacent interspace	Not reported	Kappa: 0.48 (95% CI 0.18-0.78)	MH* - weak LK† - moderate
Snider et al. (2011) [10]	L1 SP L2 SP L3 SP L4 SP	Mean all levels: 69% (Resident examiners accuracy 51%, faculty examiners average accuracy 74%)	Marker within the superior and inferior boundaries of the SP	Not reported	Kappa: 0.81 (95% CI 0.79-0.83) (for mean accuracy all levels)	MH* - strong LK† - almost perfect

Key: 95%CI - 95% Confidence Interval; MH* - McHugh (2012) kappa interpretation; LK†- Landis and Koch (1977) kappa interpretation; MT‡ - Manual Therapist

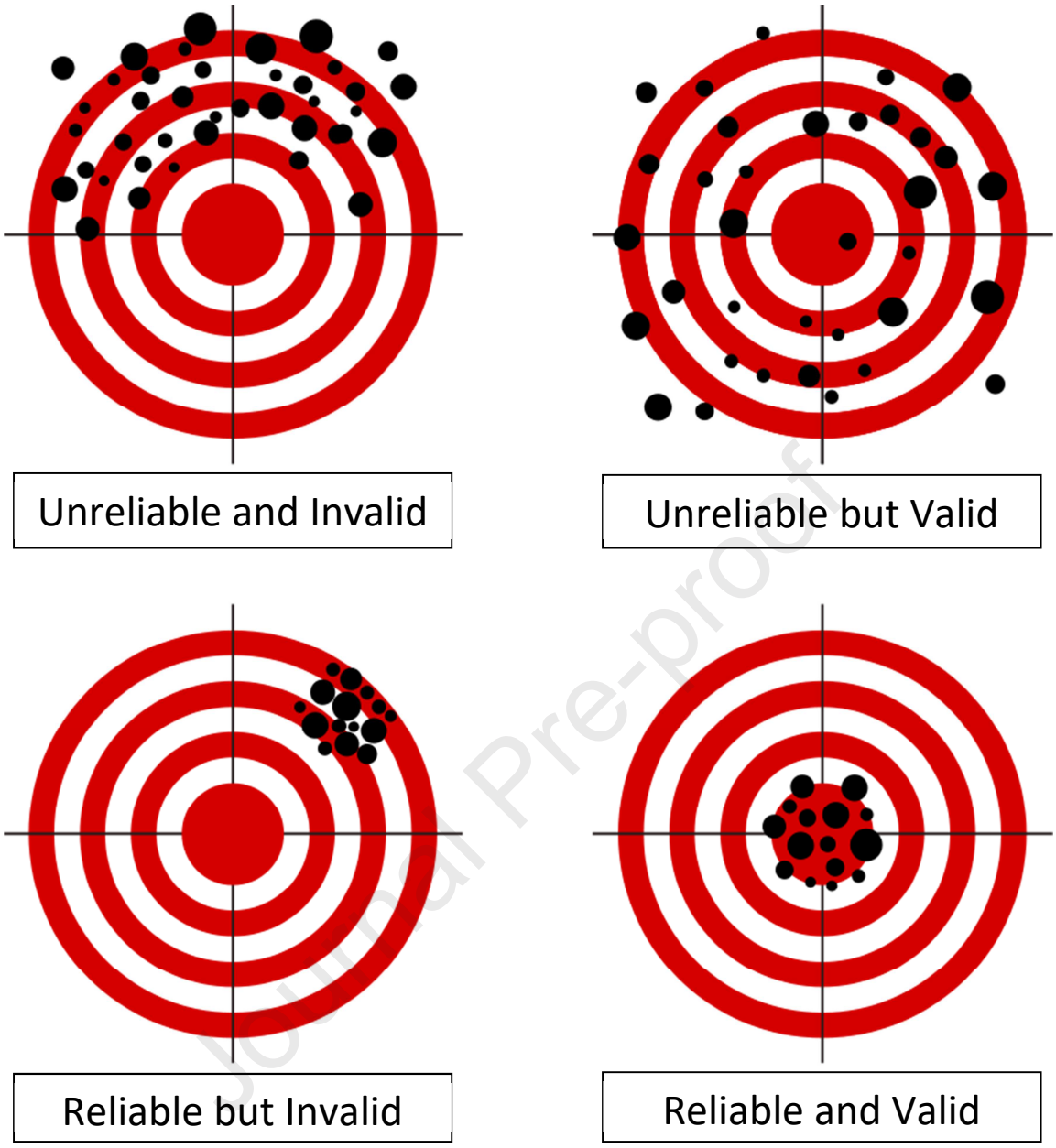

Figure 1 Validity and Reliability - a graphical representation.

[available from: https://commons.wikimedia.org/wiki/File:Reliability_and_validity.svg]

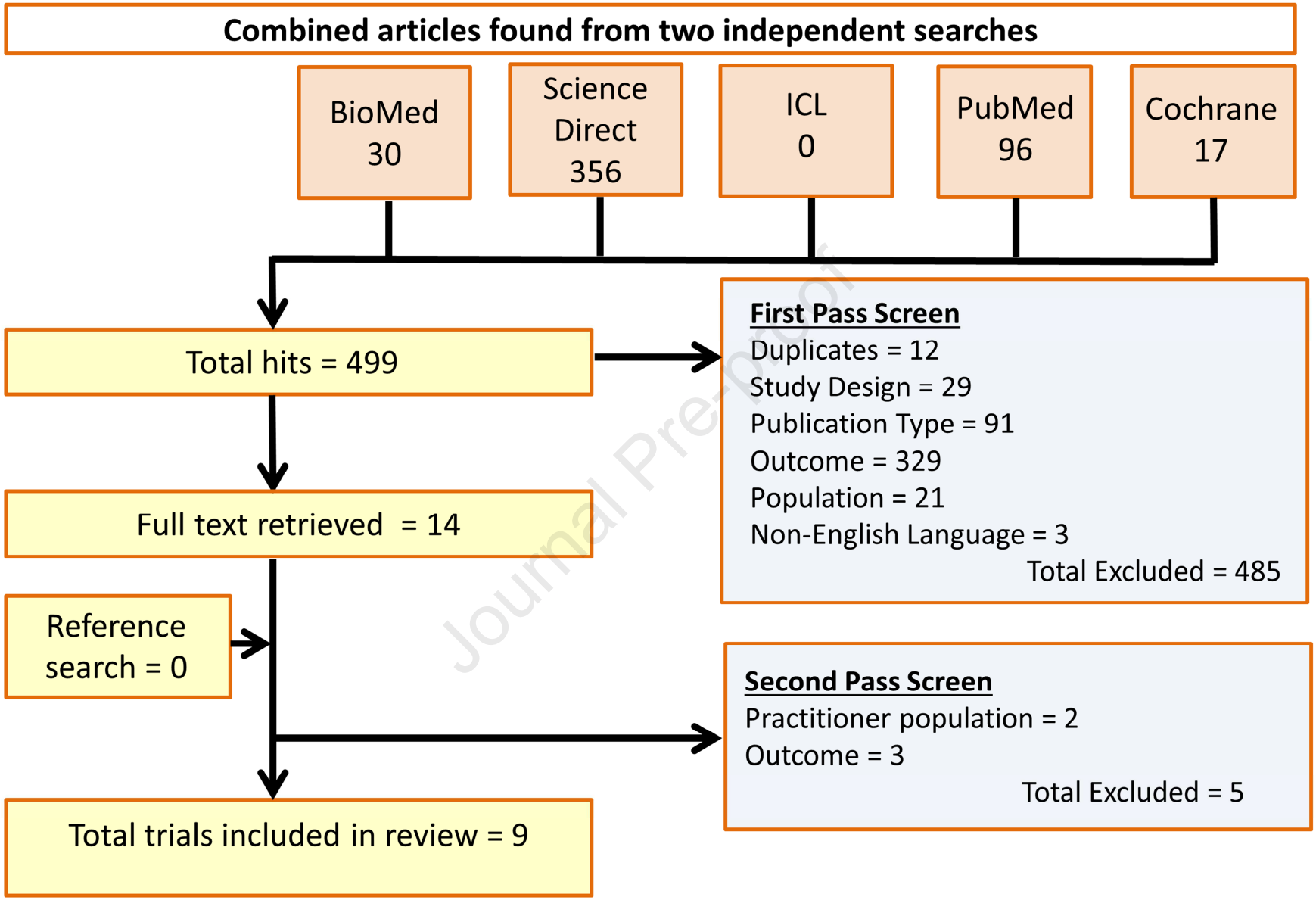

Figure 2 - Search Results

HIGHLIGHTS

- Lumbo-pelvic landmark palpation (LPLP) has sub-threshold validity scores
- These scores indicate that LPLP lacks necessary validity for clinical practice
- Research on LPLP suffers with heterogeneity and methodological issues
- Aligning with nsLBP practice guidelines will reduce practitioners' reliance on LPLP

Declarations of Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

1. Conflict of Interest: DR is a member of the Editorial Board of the International Journal of Osteopathic Medicine. He was not involved in review or editorial decisions regarding this manuscript
2. Funding Sources: DR and TW are employed by the European School of Osteopathy research department.
3. Ethical approval details: the protocol for this review was approved by the European School of Osteopathy Research Ethics Committee.