Clinical Diagnostic Tests versus Medial Branch Blocks for Adults with Persisting Cervical Zygapophyseal Joint Pain: A Systematic Review and Meta-Analysis

Kendra Usunier, MCISc (manipulative therapy), PT, FCAMPT; Mark Hynes, MCISc (manipulative therapy), PT, FCAMPT; James Michael Schuster, MCISc (manipulative therapy), PT, FCAMPT; Annie Cornelio-Jin Suen, MCISc (manipulative therapy), PT, FCAMPT; Jackie Sadi, MSc, PT, FCAMPT; David Walton, PhD, PT

ABSTRACT

Purpose: A systematic review and meta-analysis were performed to identify clinical tests for diagnosing cervical zygapophyseal joint pain (CZP) and to determine their diagnostic accuracy. **Method:** A search strategy was carried out to find relevant evidence published in CINAHL, Embase, MEDLINE, and PEDro from 1980 to January 1, 2015, pertaining to the clinical diagnosis of CZP. Quality assessment was completed using the Quality Assessment of Diagnostic Accuracy Studies–2. Results were analyzed to pool sensitivity and specificity and clarify diagnostic value. **Results:** Seven articles ($n = 463$) were included for data synthesis and review. Intersegmental mobility tests were found to have the highest diagnostic accuracy, with pooled sensitivity of 0.91 (95% CI: 0.85, 0.94) and specificity of 0.74 (95% CI: 0.65, 0.81). The pooled sensitivity for mechanical sensitivity (palpation) was 0.88 (95% CI: 0.78, 0.95), and specificity was 0.61 (95% CI: 0.50, 0.71). **Conclusion:** Limited studies are available that discuss the clinical diagnosis of CZP, and significant heterogeneity is present in the available data. In this review, intersegmental mobility tests were found to be the most accurate. Clustering of tests, agreement on a reference standard, and further exploration of CZP referral patterns are recommended.

Key Words: cervical zygapophyseal joint; diagnosis; medial branch block; meta-analysis; systematic review.

RÉSUMÉ

Objectif: analyse systématique et méta-analyse afin de déterminer les tests cliniques nécessaires pour diagnostiquer des douleurs des articulations zygapophysaires cervicales (AZC) et en établir l'exactitude diagnostique. **Méthodologie:** les chercheurs ont adopté une stratégie de recherche pour extraire les données probantes pertinentes publiées dans CINAHL, Embase, Medline et PEDro entre 1980 et le 1^{er} janvier 2015, à l'égard du diagnostic clinique d'AZC. Ils en ont évalué la qualité au moyen des études d'exactitude diagnostique des évaluations de la qualité—2. Ils ont analysé les résultats pour en regrouper la sensibilité et la spécificité ainsi que pour en confirmer la valeur diagnostique. **Résultats:** les auteurs ont conservé sept articles ($n = 463$) pour la synthèse et l'analyse des données. Ils ont découvert que les tests de mobilité intersegmentaire possédaient la plus grande précision diagnostique. En effet, la sensibilité groupée s'établissait à 0,91 (IC à 95% : 0,85, 0,94) et la spécificité, à 0,74 (IC à 95% : 0,65, 0,81). La sensibilité mécanique groupée (palpation) était de 0,88 (IC à 95% : 0,78, 0,95), et la spécificité, à 0,61 (IC à 95% : 0,50, 0,71). **Conclusion:** peu d'études portent sur le diagnostic clinique des AZC, et les données disponibles sont très hétérogènes. Dans la présente analyse, les tests de mobilité intersegmentaire étaient les plus précis. Il est recommandé de regrouper les tests, de s'entendre sur une norme de référence et d'approfondir l'examen des profils d'orientation des AZC vers d'autres spécialistes.

Neck pain is a common and potentially debilitating problem, ranking among the top 10 conditions affecting disability-adjusted life years in terms of the global burden of disease.¹ The cervical zygapophyseal joints have been implicated as the primary pain generator in 36%–67% of patients with chronic neck pain.^{2–4} Despite this prevalence, conventional radiological imaging has been largely inconclusive or unreliable when diagnosing zygapophyseal

joint pathology.⁵ Cervical zygapophyseal joint pain (CZP) is commonly associated with either unilateral or bilateral axial neck pain, which rarely radiates distal to the shoulder.⁵ It is often differentiated from pain of discogenic or other neuropathic origin on the basis of an absence of neurological symptoms.⁶ Historically, zygapophyseal joint syndromes have been a diagnosis of

From the School of Physical Therapy, Master of Clinical Science program, Western University, London, Ont.

Correspondence to: Jackie Sadi, Rm. 1588, School of Physical Therapy, Elborn College, Western University, London, ON N6G 1H1; jsadi2@uwo.ca.

Contributors: All authors designed the study; or collected, analyzed, or interpreted the data; and drafted or critically revised the article and approved the final draft.

Competing interests: None declared.

Physiotherapy Canada 2018; 70(2);179–187; doi:10.3138/ptc.2016-89.mt

exclusion (after ruling out other causes) rather than a recognized clinical diagnosis.⁷

Comparative medial branch blocks (MBBs) are currently the accepted reference standard for diagnosing CZP.⁸ However, we would refrain from claiming MBB to be the gold standard because its diagnostic accuracy has been questioned; even in the earlier work leading to its status as reference standard, the false-positive rate was 27%.⁹ Diagnostic precision was further investigated in a recent systematic review,¹⁰ and MBBs were deemed to be adequately specific for diagnosing CZP only when a high threshold of at least 75% pain relief was required, sacrificing sensitivity in favour of reducing false positives. True diagnostic accuracy has been difficult to ascertain because a criterion standard is lacking, but false-positive rates of 27%–63% have been reported.⁵

MBBs are an invasive procedure, there are substantial wait times for them in Canada, and they require interventional radiologists, specialized equipment, and acceptance of some risk of adverse reaction.¹¹ They are also associated with significant costs, estimated at up to \$10,000 per procedure.^{12,13} For these reasons, several authors have attempted to derive objective, non-invasive clinical measures that can accurately diagnose or screen for cervical zygapophyseal joint dysfunctions. These attempts have included establishing maps of CZP referral in healthy¹⁴ and clinical samples^{4,15} and manual evaluation of joint mobility.¹⁶ Jull and colleagues¹⁶ found perfect diagnostic accuracy (100% sensitivity and specificity) when manual mobility testing was performed by an expert clinician, but these results have not held up on independent investigation.¹⁷ Other clinical tests or clusters thereof have been explored for their diagnostic accuracy with varying degrees of success.^{11,18} Feasible and low-burden, non-invasive clinical screening tests hold the potential to minimize the need for MBB by identifying those patients who are less likely to present with CZP before conducting the invasive reference standard.

To our knowledge, no higher order analysis has been conducted to systematically identify, appraise, and synthesize the diagnostic accuracy of non-invasive clinical tests verified by the reference standard (MBB) for CZP (C2–C7) in adults with persistent neck pain. If non-invasive clinical proxies can be found, this may lead to the creation of a clinical decision aid before a patient is sent for MBB, or, if these proxies are particularly accurate, they may even replace the need for MBB altogether. The purpose of this article is to investigate the accuracy of these clinical index tests as a replacement or a triage for the current test regime (i.e., MBB).

METHODS

Search strategy

The search strategy was formulated by a group of researchers and experienced clinical physical therapists.

It was conducted and reported in accordance with the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) group;¹⁹ it was not registered online because this was not routine at the time.

The databases searched were CINAHL, Embase, MEDLINE, and PEDro. The search terms used were (neck OR cervical) AND (facet* OR zygoapophyseal OR zygapophyseal OR zygapophysial OR z-joint* OR articular* OR joint*) AND (Diagnos* OR discriminat* OR valid* OR sensitivity OR specificity OR likelihood OR predictive value) AND (block). *Block* was included because our goal was to identify and synthesize only the evidence that compared clinical tests with MBB as a reference standard. These databases were searched from inception to January 1, 2015. Inclusion criteria were adults aged 18 years and older, English language, peer-reviewed published evidence, human test subjects, primary sources, and neck pain, as defined within the boundaries of the occiput to T4 and the lateral ends of the acromia. Grey literature such as clinical practice guidelines, abstracts, magazine articles, Web sites, and non-peer-reviewed articles was excluded. A hand search was performed on a previous systematic review's reference list to seek out any further relevant citations.

Four independent clinical physical therapists screened the titles and abstracts of studies retrieved from the database search. Only those articles that met the inclusion criteria and reported at least one statistic related to diagnostic accuracy (e.g., sensitivity, specificity, positive likelihood ratio, negative likelihood ratio) were retained for further review. Any article that reported subjects with significant comorbidities was excluded. Each independent reviewer kept track of his or her own included and excluded articles, and then the reviewers discussed the final, full article selection. If there was any disagreement between the independent reviewers regarding the inclusion or exclusion of a study, the advice of an expert researcher (DW) was sought.

Quality scoring

Quality assessment of the included articles was completed using the Quality Assessment of Diagnostic Accuracy Studies–2 (QUADAS-2).²⁰ This is a revised version of the original QUADAS,²¹ and both are meant to quantify the risk of bias in research studies describing diagnostic accuracy. Four reviewers appraised the included articles after an initial calibration, during which they selected one study at random and reviewed it together to ensure that all reviewers were interpreting the review criteria consistently. When disagreements were observed in scoring, a third independent expert researcher was consulted.

The QUADAS–2 assesses four key domains, each scored as met, not met, or unclear.²⁰ They are (1) patient selection, (2) index test(s), (3) reference standard, and (4) flow and timing.

Table 1 Scoring Test Criteria Using QUADAS–2

Study	Risk of bias				Applicability concerns		
	Patient selection	Index tests	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Aprill et al. ¹⁵		✓		✓		✓	
Dwyer et al. ^{14*}	✓	✓	✓	✓	✓	✓	✓
Jull et al. ¹⁶	✓	✓			✓	✓	
King et al. ¹⁷	✓	✓	✓		✓	✓	✓
Schneider et al. ¹¹	✓	✓	✓	✓	✓	✓	✓
Speldewinde et al. ⁴		✓	✓		✓	✓	✓
Siegenthaler et al. ¹⁸	✓	✓	✓	✓	✓	✓	✓

*The findings in this article were not included in our synthesis, but the article is included here because it is commonly cited despite having only five subjects. QUADAS–2 = Quality Assessment of Diagnostic Accuracy Studies–2.

Using these criteria, the articles were rated for risk of bias (four criteria) and applicability (three criteria). Risk of bias was considered low when at least three of four QUADAS–2 criteria were met, and it was considered high when two or fewer criteria were met. Applicability was considered high when at least two of three applicability criteria were met. The appraisal results are shown in Table 1.

Data extraction and analysis

The independent reviewers extracted data from each of the selected articles. Data collection forms were used to summarize population, setting, geography, test details, number of test subjects, reference standard (single- vs. double-controlled MBB), sensitivity, specificity, and date of publication (see Appendix 1). Extracted data were compared between reviewers to ensure consistency, and disagreements between reviewers were easily resolved through discussion. Raw data were extracted when possible to estimate pooled sensitivity and specificity for any test evaluated in at least two independent cohorts. When data were incomplete or insufficient for meta-analysis, an attempt was made to contact the study's primary author for clarification. If no response was obtained after two attempts, those data were not included in the analysis. The data analysis was performed using Meta-DiSc version 1.4 (Microsoft, Redmond, WA).

The heterogeneity of calculated effect sizes was assessed using χ^2 and I^2 tests. The χ^2 was used to test the hypothesis that two or more population distributions differed from one another at a level greater than chance, with a significant value ($p < 0.05$) indicating heterogeneity. The I^2 value is an estimate of the percentage of total variation that is due to heterogeneity among studies rather than due to chance.²² For the I^2 score, 25% indicates a low level of heterogeneity among effects from primary sources; 50%, moderate; and 75%, high. In all analyses, a more conservative random-effects model was used to account for differences in regional practice, sampling frame, and other uncontrolled confounders.

Moderator analyses

Moderator analyses were planned for any pooled result that revealed statistically significant heterogeneity. This would normally occur by stratifying the sample of articles by the moderator variable to determine the degree of difference in calculated effect sizes by level of moderator. However, in this case the absolute number of primary studies was too low to achieve meaningful moderator analyses. It had been anticipated that age group, single versus double MBB, geography, gender, and setting would be evaluated.

RESULTS

Study selection

Figure 1 displays the search strategy and results. The initial database search yielded 206 relevant articles; duplicates were removed, and titles were evaluated for applicability to our research question, which narrowed the results to 158 articles. The remaining articles underwent abstract review and application of inclusion and exclusion criteria, and this filtered the results to 31 articles that underwent full review. After full review, 7 articles were retained for data extraction. An additional 2 articles did not provide sufficient data for pooling, and the primary authors either did not have the research details still available ($n = 1$) or did not respond to two contact attempts ($n = 1$). A third article did not meet the criteria for meta-analysis because it did not compare its findings with a reference standard test. These 3 articles were included as part of the discussion and as supporting evidence for pooled findings but did not contribute data to the statistical analysis.

Four clinical tests were identified in the seven studies. Two of the tests had sufficient data and at least two independent cohorts allowing statistical pooling. The four tests were passive intersegmental motion testing,^{11,16,17} mechanical sensitivity,^{11,18} cervical zygapophyseal joint referral patterns,^{4,15} and extension rotation test.¹¹ A description of these clinical tests can be found in Appendix 2.

Figure 1 Search strategy and article selection.

The pooled sensitivity for passive intersegmental motion testing (three articles; $n = 318$) for identifying zygapophyseal joint pain was 0.90 (95% CI: 0.85, 0.94), and the pooled specificity was 0.73 (95% CI: 0.65, 0.81) (see Figure 2). All three studies were at low risk of bias, and there were no identified applicability concerns. Heterogeneity was not statistically significant for either sensitivity ($p = 0.47$) or specificity ($p = 0.53$), with $I^2 < 0.1\%$.

For mechanical sensitivity ($n = 158$; low risk of bias), the pooled sensitivity was 0.88 (95% CI: 0.78, 0.95), and the pooled specificity was 0.61 (95% CI: 0.50, 0.71). The χ^2 indicated significant heterogeneity ($p < 0.01$), and I^2 was 86.8%. Heterogeneity may be due to the difference in collection methods (digital algometer using interval-level scaling¹¹ vs. manual pressure using nominal-level scaling¹⁸).

Two articles evaluated the diagnostic utility of cervical joint referral patterns, or “pain maps.” Aprill and colleagues¹⁵ ($n = 10$; high risk of bias) reported strong agreement between pain maps and localization of CZP, with 9 of 10 participants having MBB-confirmed CZP at predicted segments. Speldewinde and colleagues⁴ ($n = 97$; low risk of bias) completed a retrospective audit and found that, of the 36% of patients confirmed to have CZP, 83% were predicted at the correct segment by following the pain maps described by Dwyer and colleagues.¹⁴ Neither provided sufficient evidence (e.g., false

positives, false negatives, true negatives) to allow statistical pooling, and we were unable to obtain these from the primary authors.

Schneider and colleagues¹¹ ($n = 125$; low risk of bias) described an extension rotation test that was not described in other primary evidence in this systematic review, thereby preventing statistical pooling. They reported sensitivity to be 0.83 (95% CI: 0.76, 0.89) and specificity to be 0.59 (95% CI: 0.50, 0.68).

DISCUSSION

The purpose of this article was to systematically identify, appraise, and synthesize evidence on the diagnostic accuracy of clinical tests for diagnosing CZP, a common practice among spinal rehabilitation professionals. The systematic review identified seven articles for inclusion, four of which provided sufficient data for analysis of four clinical tests. The a priori decision to include only those articles that were referenced against a criterion standard of an MBB limited the studies available but was an acceptable trade-off to improve the homogeneity of the primary sources.

Passive intersegmental motion testing was found to have the largest body of evidence available to support its use in diagnosing CZP against MBB, and it had the highest pooled sensitivity and specificity among the tests identified. Mechanical sensitivity was found to be less

Figure 2 Forest plot for sensitivity and specificity of passive intersegmental motion testing of cervical zygapophyseal joint pain.

accurate, but the pooled results were influenced by significant heterogeneity in the study design and by the availability of only two primary sources.

Zygapophyseal joint referral patterns have been described previously as a method of identifying zygapophyseal joints as a pain generator in a cervical pain presentation. Three articles were identified that explored the use of joint referral patterns, or pain maps. Dwyer and colleagues¹⁴ assessed cervical zygapophyseal joint referral patterns in healthy individuals, eliciting pain by performing intra-articular joint injections under fluoroscopy; hence, this evidence was not included in our synthesis, which required clinical rather than healthy samples. The article is worth mentioning, however, because it is commonly cited despite including only five subjects. Aprill and colleagues¹⁵ and Speldewinde and colleagues⁴ described referral patterns compared against the reference standard of an MBB in clinical samples; however, insufficient data were reported in the articles to allow statistical analysis and pooling. Of note, Aprill and colleagues¹⁵ used the cervical zygapophyseal joint referral patterns described by Dwyer and colleagues¹⁴ to predict the symptomatic joint level, which was confirmed by diagnostic

block. Speldewinde and colleagues⁴ had three independent examiners complete retrospective chart reviews of people with neck pain that went onto diagnostic zygapophyseal joint blocks. Joint levels were selected for patients, again using the distribution maps of Dwyer and colleagues.¹⁴ In patients confirmed to have pain originating from the zygapophyseal joint, 83% of the ratings identified the correct level on the basis of pain distribution maps, but no results were reported for rate of false positives or false negatives, precluding the calculation of discriminative statistics.

Schneider and colleagues¹¹ described an extension rotation test with moderate sensitivity and specificity for diagnosing CZP; however, no other study was available describing the same test for comparison and pooling of results.¹¹ The results of this single article (sensitivity = 0.83, specificity = 0.59) suggest that it may have limited clinical utility for diagnosing facetogenic neck pain and that it may be more useful as a test of exclusion (high sensitivity) than as a test of inclusion (low specificity), but because it is a simple test with low risk of harm, it may be useful as part of a cluster of tests.

A common issue we encountered was heterogeneity among the studies in their design and inconsistency in the use of the reference standard, even when sources were selected specifically because they had used MBB. Although the current generally accepted standard for using MBB is abolishing at least 80% of pain, this threshold was not consistent among the studies.¹¹ For the clinician community to develop optimal clinical guidelines and examination clusters that may replace other, costlier strategies, there is a clear need for a consistent reference standard. In a recent systematic review on the topic, Falco and colleagues¹⁰ indicated that diagnostic nerve blocks were safe, valid, and reliable.²³ The strength of evidence for diagnostic facet joint nerve blocks is generally considered good, based on high specificity, when a diagnostic threshold of at least 80% pain relief is used.¹¹ Although this criterion sacrifices sensitivity (leading to more false negatives), comparing clinical tests against such a reference standard would increase confidence in the results of knowledge syntheses such as these.

Another, related consideration is that of generalizability to primary care; the most obvious clinical translation of these findings would be in allowing primary care providers to conduct a triage assessment of their patients before sending them to secondary or tertiary environments for interventional assessment or management. However, readers should be aware that all the subjects in the included studies were already part of a selected and somewhat homogeneous group because they were in the secondary or tertiary care environment. The effects of this is that diagnostic accuracy will likely suffer to an unknown degree when applied to the more heterogeneous population seen in primary care.

One concern about the rote endorsement of passive intersegmental motion assessment is that success in this technique may be influenced by clinician experience. The same could arguably be said for all four tests described here, but clinical experience and history would suggest that this particular test requires a large experiential background with different body types, sizes, and compositions to have confidence in reporting findings as normal or abnormal. As described in our pool of evidence, the test was at least partly based on patient self-report of pain, which is presumably independent of operator experience. Regardless, theory and practice indicate that clinician experience is important, but the evidence did not allow this comparison to be conducted.

The limitations to this study are similar to those found in many systematic reviews. The first is the low number of primary sources from which we were able to extract data. While keeping our inclusion criteria strict resulted in overall greater confidence (less heterogeneity) in our results, it does also limit the conclusions that can be drawn. Similarly, conclusions in knowledge syntheses are only as good as the quality of evidence entered. The primary evidence included five sources rated as having a low risk of bias and two rated as having a high risk of

bias. Considering the low number of primary resources, two of which had a high risk of bias, a degree of caution is required when interpreting the results.

Another limitation to interpretation is the difference in application of the clinical tests across study cohorts (see Appendix 2). This is especially true in the mechanical sensitivity test, for which one article used a linear scale of pain sensitivity and another used patient reports of *painful* or *not painful* in response to set pressure. Although both ultimately dichotomized the test as *sensitive* or *not sensitive*, there was significant heterogeneity in effect sizes. With only two studies included, we were unable to confidently identify the source of the heterogeneity, but the difference in measurement seems a likely contributor. Readers should interpret the pooled effect sizes with extreme caution.

Finally, other considerations are the exclusion of grey literature, a common approach intended to improve the robustness of the included sources but at the sacrifice of comprehensiveness.

CONCLUSION

The results of this study suggest that skilled clinicians may be able to diagnose pain of facetogenic origin with a pooled sensitivity of 0.90 (95% CI: 0.85, 0.94) and pooled specificity of 0.74 (95% CI: 0.65, 0.81) using manual application of intersegmental joint pressures. No other tests were identified that could be confidently endorsed. The implications for practice are that skilled clinicians may be able to fill a triage role when a patient is being considered for interventional investigation, with a greater ability to rule out those unlikely to respond to MBB (higher sensitivity) than to identify those who will respond.

A cluster of tests is likely to bear greater results, but clustered tests are currently difficult to find in the pool of empirical evidence.

The implication for research is an ongoing issue related to the inconsistent application of a reference standard to facilitate interpretation of primary findings. The results of this study should be interpreted in light of the quality and size of the primary sources. There is also a consistent problem with reporting results in this field, with author groups appearing to prioritize reporting false negatives in the absence of other data that would facilitate knowledge syntheses. This approach may be partly due to the lack of a true criterion standard, rendering the true diagnostic accuracy of non-invasive clinical tests or MBB open to ongoing skepticism.

KEY MESSAGES

What is already known on this topic

Clinicians routinely evaluate patients for the presence or absence of facet dysfunction as a potential driver of neck pain. Several studies have now explored the discriminative value of many such clinical tests in isolation or as part of a cluster, but there has yet to be a structured

systematic review and higher order analysis to synthesize the findings of such work.

What this study adds

A structured systematic review and higher order analysis identified consistent weaknesses in the pool of literature that can be easily addressed through careful attention to research design. The test with the greatest overall discriminative accuracy was passive intersegmental motion testing, although important caveats are endorsed, including the likely effect of clinician experience when interpreting clinical findings.

REFERENCES

- GBD 2013 DALYs and HALE Collaborators, Murray CJ, Barber RM, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: Quantifying the epidemiological transition. *Lancet*. 2015;386(10009):2145–91.
- Lord SM, Barnsley L, Wallis BJ, et al. Chronic cervical zygapophysial joint pain after whiplash. A placebo-controlled prevalence study. *Spine (Phila Pa 1976)*. 1996;21(15):1737–44, discussion 1744–5. <http://dx.doi.org/10.1097/00007632-199608010-00005>. Medline:8855458
- Manchikanti L, Boswell MV, Singh V, et al. Prevalence of facet joint pain in chronic spinal pain of cervical, thoracic, and lumbar regions. *BMC Musculoskelet Disord*. 2004;5(1):15. <http://dx.doi.org/10.1186/1471-2474-5-15>. Medline:15169547
- Speldewinde GC, Bashford GM, Davidson IR. Diagnostic cervical zygapophyseal joint blocks for chronic cervical pain. *Med J Aust*. 2001;174(4):174–6. Medline:11270757.
- Sehgal N, Dunbar EE, Shah RV, et al. Systematic review of diagnostic utility of facet (zygapophysial) joint injections in chronic spinal pain: an update. *Pain Physician*. 2007;10(1):213–28. Medline:17256031.
- Narouze SN, Van Zundert J, Van Kleef M. Cervical facet syndrome: cervical medial branch block and radiofrequency ablation. *Interventional management of head and face pain*. New York: Springer; 2014. p. 83–90.
- Bogduk N, Marsland A. The cervical zygapophysial joints as a source of neck pain. *Spine (Phila Pa 1976)*. 1988;13(6):610–7. <http://dx.doi.org/10.1097/00007632-198813060-00003>. Medline:3175750
- Pauza KJ. Educational guidelines for interventional spinal procedures [Internet]. Rosemont (IL): American Academy of Physical Medicine and Rehabilitation; 2008 [cited 2015 June 5]. Available from: <http://www.aapmr.org/practice/guidelines/documents/edguidelines.pdf>.
- Barnsley L, Lord S, Wallis B, et al. False-positive rates of cervical zygapophysial joint blocks. *Clin J Pain*. 1993;9(2):124–30. <http://dx.doi.org/10.1097/00002508-199306000-00007>. Medline:8358135
- Falco FJ, Manchikanti L, Datta S, et al. Systematic review of the therapeutic effectiveness of cervical facet joint interventions: an update. *Pain Physician*. 2012;15(6):E839–68. Medline:23159978.
- Schneider GM, Jull G, Thomas K, et al. Derivation of a clinical decision guide in the diagnosis of cervical facet joint pain. *Arch Phys Med Rehabil*. 2014;95(9):1695–701. <http://dx.doi.org/10.1016/j.apmr.2014.02.026>. Medline:24662813
- Bogduk N, Holmes S. Controlled zygapophysial joint blocks: the travesty of cost-effectiveness. *Pain Med*. 2000;1(1):24–34. <http://dx.doi.org/10.1046/j.1526-4637.2000.99104.x>. Medline:15101961
- Cohen SP, Williams KA, Kurihara C, et al. Multicenter, randomized, comparative cost-effectiveness study comparing 0, 1, and 2 diagnostic medial branch (facet joint nerve) block treatment paradigms before lumbar facet radiofrequency denervation. *Anesthesiology*. 2010;113(2):395–405. <http://dx.doi.org/10.1097/ALN.0b013e3181e33ae5>. Medline:20613471
- Dwyer A, Aprill C, Bogduk N. Cervical zygapophyseal joint pain patterns. I: A study in normal volunteers. *Spine (Phila Pa 1976)*. 1990;15(6):453–7. <http://dx.doi.org/10.1097/00007632-199006000-00004>. Medline:2402682
- Aprill C, Dwyer A, Bogduk N. Cervical zygapophyseal joint pain patterns. II: A clinical evaluation. *Spine (Phila Pa 1976)*. 1990;15(6):458–61. <http://dx.doi.org/10.1097/00007632-199006000-00005>. Medline:2402683
- Jull G, Bogduk N, Marsland A. The accuracy of manual diagnosis for cervical zygapophysial joint pain syndromes. *Med J Aust*. 1988;148(5):233–6. Medline:3343953.
- King W, Lau P, Lees R, et al. The validity of manual examination in assessing patients with neck pain. *Spine J*. 2007;7(1):22–6. <http://dx.doi.org/10.1016/j.spinee.2006.07.009>. Medline:17197328
- Siegenthaler A, Eichenberger U, Schmidlin K, et al. What does local tenderness say about the origin of pain? An investigation of cervical zygapophysial joint pain. *Anesth Analg*. 2010;110(3):923–7. <http://dx.doi.org/10.1213/ANE.0b013e3181cbd8f4>. Medline:20185669.
- Moher D, Liberati A, Tetzlaff J, et al.; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–9, W64. <http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135>. Medline:19622511.
- Whiting PF, Rutjes AW, Westwood ME, et al.; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529–36. <http://dx.doi.org/10.7326/0003-4819-155-8-201110180-00009>. Medline:22007046.
- Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3(1):25. <http://dx.doi.org/10.1186/1471-2288-3-25>. Medline:14606960.
- Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–60. <http://dx.doi.org/10.1136/bmj.327.7414.557>. Medline:12958120.
- Falco FJ, Erhart S, Wargo BW, et al. Systematic review of diagnostic utility and therapeutic effectiveness of cervical facet joint interventions. *Pain Physician*. 2009;12(2):323–44. Medline:19305483.
- Fritzell P, Berg S, Borgström F, et al. Cost effectiveness of disc prosthesis versus lumbar fusion in patients with chronic low back pain: randomized controlled trial with 2-year follow-up. *Eur Spine J*. 2011;20(7):1001–11. <http://dx.doi.org/10.1007/s00586-010-1607-3>. Medline:21053028.

APPENDIX 1: DESCRIPTION OF CLINICAL TESTS

Study	Test	Sensitivity, %	Specificity, %	Risk-of-bias score	Country	Reference standard	Gender, M:F	Age range, y	No.
Aprill et al. ¹⁵ (prospective case series)	Cervical zygapophyseal joint referral patterns	Not specified	Not specified	High	Australia and United States	1 MBB	Not specified	Not specified	10
Jull et al. ¹⁶ (prospective case series)	Intersegmental motion tests	100	100	High	Australia	1 MBB	7:13	Not specified	20
King et al. ¹⁷ (prospective case series)	Intersegmental motion tests	89	47	Low	Australia	2 MBB	78:95	19–75	173
Schneider et al. ¹¹ (prospective cohort)	Intersegmental motion tests	92	71	Low	Canada	2 MBB	41:84	21–65; avg 49	125
Schneider et al. ¹¹ (prospective cohort)	Mechanical sensitivity	67	16	Low	Canada	2 MBB	41:84	21–65; avg 49	125
Schneider et al. ¹¹ (prospective cohort)	Extension rotation test	83	59	Low	Canada	2 MBB	41:84	21–65; avg 49	125
Siegenthaler et al. ¹⁸ (prospective case series)	Mechanical sensitivity	67	16	Low	Switzerland	2 MBB	13:20	48–68	33
Speldewinde et al. ⁴ (retrospective case series)	Cervical zygapophyseal joint referral patterns	Not specified	Not specified	Low	Australia	2 MBB	52:45	18–62	97

M:F = male:female; MBB = medial branch blocks; avg = average.

APPENDIX 2: DESCRIPTION OF CLINICAL DIAGNOSTIC TESTS FOR CERVICAL ZYGAPOPHYSEAL JOINT PAIN

Intersegmental Motion Tests

Jull et al.¹⁶

Passive accessory movements were performed to determine the suspected cervical zygapophyseal joint that could be contributing to pain symptoms. This was completed by manipulative therapists with postero-anterior glides, which were performed centrally over the spinous processes, and unilaterally over the laminae and zygapophyseal joints of all segmental levels that were putatively relevant to patient's complaint.²⁴ The criteria adopted were abnormal "end feel," abnormal quality of resistance to motion, and reproduction of pain, either local or referred, when passive accessory movements were tested.²⁴ The results were confirmed with diagnostic nerve block.

King et al.¹⁷

Passive accessory movements were performed following the guidelines described by Jull et al.¹⁶ to determine whether cervical symptoms were due to cervical zygapophyseal joint pain. The results were confirmed with a diagnostic nerve block.

Schneider et al.¹¹

A manual segmental examination was performed to determine whether the cervical zygapophyseal joint was the source of pain symptoms.

The subject was positioned prone with the cervical spine in a neutral position. The assessor applied a posterior-anterior directed force over the articular pillars from C2–3 to C6–7 on each side. Any perceived resistance to motion was categorized as normative, slight, moderate or marked.^{6(p.88)}

Comparative medial branch blocks (MBBs) were performed to confirm the level.

Mechanical Sensitivity

Schneider et al.¹¹

"Palpation for segmental tenderness was performed with the subject in a prone position. The assessor palpated the segmental muscles overlying the facet joints bilaterally."^{6(p.88)} "The test was considered positive if the patient reported an increase in the familiar pain, either local or referred at an intensity of $\geq 3/10$."^{6(p.88)} Comparative MBBs were performed to confirm the level.

Siegenthaler et al.¹⁸

The symptomatic cervical zygapophyseal joint was determined by the pressure pain detection threshold, assessed with a digital algometer.

The patients were placed in a side-lying position on their non-painful side, with a pillow below the head and the neck to keep the cervical spine in horizontal position. Each of the zygapophyseal joints, C2–C7, was located by ultrasound.⁸

The pressure was increased continuously from 0 to 1200 kPa at a rate of 30 kPa/s. Patients were instructed to press a button when the stimulus changed from a non-painful to a painful sensation: this was noted as the pain pressure threshold.⁸

The painful joints were confirmed with diagnostic MBB.

Cervical Zygapophyseal Joint Referral Patterns

Aprill et al.¹⁵

This study assessed the validity of cervical zygapophyseal pain distribution using the pain maps first described by Dwyer and colleagues.¹⁴ Ten subjects were reviewed by radiologists to determine the predicted cervical zygapophyseal joint, comparing subjective reports to the pain maps. This was confirmed by cervical MBBs.

Speldewinde et al.⁴

A total of 97 patient records were reviewed of individuals who had undergone diagnostic zygapophyseal

joint blocks. Cervical MBBs were performed at the anticipated symptomatic level using the pain distribution maps of Dwyer and colleagues.¹⁴

Extension Rotation Test: Schneider et al.¹¹

Subjects were asked to actively move into a position of full extension and then full rotation to both left and right. "A positive test was defined as the provocation of familiar cervical spine pain intensity of $\geq 3/10$."^{6(p.88)} Comparative MBBs were performed to confirm the level.