

Validity of the Spurling test in the diagnosis of cervical radiculopathy: A systematic review

Ting Wang (✉ ting.wang.1@hotmail.com)

Wellington Medical Group <https://orcid.org/0000-0002-7204-8412>

Ting Wang

University of Otago

Research Article

Keywords: Spurling test, Provocative test, Cervical radiculopathy, Diagnostic accuracy, Validity

Posted Date: May 31st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1621985/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

BACKGROUND: Cervical radiculopathy (CR) is a common clinical presentation. The Spurling test is a commonly used provocative test for diagnosing cervical radiculopathy. Various diagnostic accuracy studies on the validity of the Spurling test have demonstrated high specificity but variable sensitivity.

OBJECTIVE: This study aimed to find an evidence-based explanation for the variation in sensitivity of the Spurling test to assess the value of the Spurling test for clinical practice.

STUDY DESIGN: A review of studies on the diagnostic validity of the Spurling test was conducted. The study sample comprised eight diagnostic accuracy studies of the Spurling test.

OUTCOME MEASURES: Sensitivity, specificity, and likelihood ratios are presented.

METHODS: A literature search of MEDLINE, EMBAS and PubMed was conducted through April 2022. The methodological quality of the studies was evaluated by Quality Assessment of Diagnostic Accuracy Studies (QUADAS).

RESULTS: Eight diagnostic accuracy studies were selected, showing high specificity ranging from 74% to 98.5%, and the sensitivity varied from 14.7% to 95%. Studies that recruited surgical candidates showed a higher sensitivity, ranging from 92% to 95%. Studies that adopted the modified Spurling test manoeuvres (incorporating neck extension, head rotation, and lateral neck bending) revealed a slightly higher sensitivity than those using only the lateral neck bending manoeuvre. The inherent flaws of the reference standards may have altered the validity of the Spurling test.

CONCLUSIONS: The value of the studies on the validity of the Spurling test has been undermined by multiple areas causing bias, such as patient spectrum difference, index test interpretation controversy, and a lack of gold reference standards. Spurling test use in clinical practice should be combined with patient history, careful selection of different Spurling variants, and other provocative tests to increase the diagnostic accuracy. Large-sample-size and well-validated studies on the accuracy of the Spurling test in diagnosing cervical radiculopathy are needed.

Introduction

CR is defined as a group of neurological deficits, including sensory changes (pain, numbness, pins and needles, paresthesia), motor changes (weakness), and/or reflex changes in the specific dermatome, myotome, or sclerotome of one or more cervical nerve root territories in one or both upper limbs (1–5). Any space-occupying lesions near the exit foramen of the cervical spine, such as intervertebral disc herniation, foraminal stenosis, and spondylolisthesis, may lead to nerve root impingement, conduction block, and CR (5, 6).

The author is based in an urban general practice clinic in the Wellington Region. The annual prevalence of CR from 1st April 2021 to 1st April 2022, across the entire centre was 1.444 per 1,000 patients. In contrast, the annual incidence reported in a population-based epidemiological study by Radhakrishnan et al. was slightly higher at 1.708 per 1,000 patients recruited between 1976 and 1990 in Rochester, Minnesota, USA (6).

There are no standardized diagnostic criteria for CR, making it difficult to establish a diagnosis (7). Diagnostic imaging (myelogram, computed tomography (CT), magnetic resonance imaging (MRI)) or electrodiagnostic testing (EDX) (needle electromyography (EMG) and nerve conduction test (NCT)) are commonly used modalities to confirm the diagnosis of CR. Often, these advanced diagnostic studies are invasive, uncomfortable, expensive, and usually not directly accessible in primary care settings. Most of these may entail lengthy waiting lists (3, 6, 8). Thus, a simple clinical examination test, preferably with high sensitivity and specificity, would be a useful tool for diagnosing CR, particularly in primary care settings.

The Spurling test was first named as the “neck compression test” by two American neurosurgeons, Roy Spurling and William Scoville in 1944 (5). Ever since it was introduced, the Spurling test has been a commonly used provocative test for diagnosing CR. It was described as “tilting the head and neck toward the painful side” to produce the characteristic pain and/or radiculopathy with or without “pressure on the top of the head in this position” (Fig. 1) (5). The purpose of the test is to decrease the dimensions of the intervertebral foramen or intervertebral space to exacerbate compression of the exiting nerve roots (5).

Despite its wide use, there are limited studies assessing the diagnostic accuracy of the Spurling test in diagnosing CR (8–15). The general consensus in the literature is that the Spurling test has high specificity but significantly variable sensitivity, ranging from low to high (16–18). Unfortunately, no study has provided a sufficient explanation for the significant sensitivity discrepancy in the Spurling test. The present review aimed to study the validity of the Spurling test and provide evidence-based explanations for the large variation in its sensitivity when diagnosing CR.

Method

1. Search strategy

An electronic search of three online databases (MEDLINE, EMBASE, and PubMed), accessed through the Otago Medical Library portal up to April 2022, was conducted. ‘Spurling test’, ‘Spurling’s test’, ‘cervical radiculopathy’, ‘diagnostic accuracy’ and ‘provocative test’ were used as searching key words in all database searches. The search was limited to the English language and human data. No filter restriction was applied to the year of publication as a small number of articles were expected. Duplicate, irrelevant, low-quality articles were excluded. The flowchart of the search method is shown in Fig. 2.

2. Review of methodological quality

Quality assessment using QUADAS was conducted to evaluate the quality of each primary study and to assess the potential sources of bias that could affect the reliability of each study (19). This tool ensured that high-quality studies were selected and studies of concern due to potential bias were noted (19).

The item was scored as "+" if the criterion was met, as "-" if the criterion was not met, or as "?" if the information was insufficient to make a judgement. A numerical score was obtained based on the total number of criteria met. Studies have shown a score of 10–14 considered as high-quality, a score of 5–9 as moderate-quality, and a score of 4 or less as low-quality (20).

3. Main outcome measurement

The validity of the Spurling test has been studied in terms of its specificity and sensitivity for the diagnosis of CR. Sensitivity or specificity values greater than 80% were considered high, 60%–79% as moderate, and less than 60% as low (11).

Positive likelihood ratios (+LR) and negative likelihood ratios (-LR) were calculated based on sensitivity and specificity to indicate a shift in probability favoring the existence of CR (+LR) or a shift in probability favouring the absence of CR (-LR) (20). Studies with +LR > 10 or -LR < 0.1 suggested large shifts, +LR between 5 to 10 and -LR from 0.1 to 0.2 suggested moderate shifts, and +LR from 2 to 5 or -LR from 0.2 to 0.5 interpreted as small shifts (21).

Result

Study characteristics

Eight primary studies on the diagnostic accuracy of the Spurling test were identified (Fig. 2). A detailed overview of the study's characteristics is provided in Table 1. Three studies focused on the validity and reliability of the Spurling test (9, 10, 15). Five studies also evaluated other clinical examinations for CR, such as the shoulder abduction test, upper limb tension test, cervical distraction test and the Neck Tornado test (8, 11–14).

Methodological quality of included studies

The overall methodological quality of the selected studies was moderate to high as per the QUADAS criteria (Table 2). The major forms of bias are summarized as follows.

1. Bias associated with patient selection

Two studies recruited patients who presented to the spinal surgery or neurosurgical unit (9, 10), whereas the remaining studies selected patients who were referred to the electrodiagnostic center, pain clinic, or multidisciplinary clinic (MDT) (7, 11–13, 15). This may suggest patient spectrum bias. Two studies were considered to have a high risk of bias because no clear or insufficient exclusion criteria were applied (9, 15). The selection criteria for all other studies were different, which could have caused a potential bias.

2. Bias associated with index test interpretation

Ghasemi et al. described the manoeuvre as "the examiner pushed participant's head downward while the head was laterally flexed on the affected side" (12), which is the most accurate interpretation of Spurling and Scoville's original description. Other studies incorporated various additional manoeuvres such as neck extension, and/or slight rotation to the contralateral side (9–14). To be more precise, Wainner et al. named the original Spurling test as "Spurling A," and the one with concurrent neck extension and head rotation as "Spurling B" (8).

3. Bias associated with reference standard

Five studies used advanced imaging (myelogram, MRI, or CT) as reference standards (9–11, 13, 14), whereas three studies used EDX testing (EMG, NCS) (8, 12, 15). To the best of our knowledge, none of the available reference standards can be considered the gold standard for diagnosing CR owing to their inherent weaknesses (10, 11, 17, 22). False-positive or false-negative findings of the reference standards would inevitably undermine the validity of the Spurling test. None of the eight selected studies met the third QUADAS criteria.

Overall findings

All eight primary studies agreed that the Spurling test has a high specificity in diagnosing CR, ranging from 74–98.5%. By contrast, their findings on sensitivity varied significantly, ranging from 14.7–95% (8–15). Two high-quality studies based on the QUADAS assessment showed relatively consistent findings of sensitivity ranging from 50–59%, and specificity ranging from 74–86% (8, 11).

Studies investigating patients who were referred to surgical services revealed a significantly high sensitivity (92%–95%) and specificity (94%–95%) (9, 10). Coincidentally, those studies were also the ones that selected CT or MRI as their reference standards. Studies that recruited patients who were referred for further testing or MDT input reported a low sensitivity between 14.7% and 59% (8, 11–15), with EMG groups representing the two lowest groups of 14.7% (12) and 30% (15), respectively.

The modified Spurling test (incorporating neck extension and rotation manoeuvres) seemed to have higher sensitivity than the original Spurling test. Ghasemi et al., who used a manoeuvre similar to the original Spurling test, reported the lowest sensitivity of 14.7% in the chronic CR group and 46.5% in the acute CR group (12).

Discussion

This study reviewed eight primary diagnostic accuracy studies that used the Spurling test. The primary goal of this study was to determine whether the Spurling test can be used as a reliable diagnostic test in primary care settings to avoid over dependence on invasive or expensive modalities. Unfortunately, the goal was limited by three factors: 1) all eight studies identified were designed in tertiary settings (spinal surgical or neurosurgical units, chronic pain clinics, MDT clinics, or electrodiagnostic centers in hospitals); 2) multiple areas with potential risks for bias were identified in almost all studies, which made it impossible to arrive at a straightforward conclusion; 3) despite its wide use in clinical practice, the Spurling test has not been sufficiently studied. The lack of studies with large sample sizes and insufficient utilization of statistical methods have significantly undermined the conclusions of most studies.

The current consensus, based on limited research findings, is that the Spurling test has high specificity but significantly variable sensitivity (16, 17). The discrepancy in the sensitivity of the Spurling test has attracted attention and attempts at explanation. The major areas that may have altered the validity of the Spurling test in diagnosing CR are discussed in subsequent sections.

1) Patient selection variation

The sensitivity of the Spurling test in surgical candidates (9, 10) is significantly higher than that in non-surgical patients. Generally, surgery is considered the last resort if conservative management fails or if the symptoms are severe. It is impossible to compare the selection bias across all studies regarding the severity of symptoms suggesting CR. However, it is not surprising that patients who were referred to the surgeon were more bothered by symptoms or had been locally or generally sensitized. In these cases, patients could be more sensitive to any nerve provocation test, including the Spurling test.

Shah et al. (10) also noted that patients with acute cervical disc prolapse were more likely to have positive findings in the Spurling test than those in the chronic phase. In addition to mechanical compression of the nerve root in both acute and chronic groups confirmed by MRI or surgery in Shah's study (10), herniated disc material (nucleus pulposus) in the acute phase could also escalate the acute inflammatory process (3). This process is likely to be mediated by inflammatory markers such as interleukin factor-6 (IL-6), tumor necrosis factor- α (TNF- α), and matrix metalloproteinases (MMPs), which could lead to further sensitization and pain (3). In the chronic phase, when inflammation is largely regressed, the nerve root becomes less sensitive to induced compression in the Spurling test (10). This theory was supported by Ghasemi et al., who reported an extremely low sensitivity of 14.7% in the chronic CR group but a much higher sensitivity of 46.5% in the acute CR group (12).

2) Multiple variants of the Spurling test

The discrepancy in various interpretations of the Spurling test was possibly caused by the misleading sketch in the original article, which illustrated a maneuver consisting of extension, rotation and lateral bending (Fig. 1) (5). Interestingly, as mentioned earlier, these additional maneuvers may have added value in enhancing the sensitivity of the Spurling test. This presumption is supported by Farmer et al.'s biomechanical study in 1994, which showed increased neuroforaminal pressure and nerve root compression as a result of additional neck extension (23). Similarly, Yoo et al. confirmed a narrower neuroforaminal size with neck extension (24). Park et al. introduced a new provocative test called Neck Tornado Test (NTT), which involves neck rotation in a 180° tornadic pattern with concurrent axial pressure (13). Similarly, Anekstein et al. designed a study that adopted six varieties of Spurling test, including lateral bending, rotation, and extension in various combinations (25). Based on the above studies, the original Spurling test was inferior to the modified variants in provoking CR symptoms.

Surprisingly, *Wainner* et al. reported no difference in sensitivity when extension was incorporated into the original Spurling test, suggesting that other factors might be responsible for the reproduction of CR symptoms (8). Bayoglu et al. reported in their biomechanical study that neck extension decreased intervertebral disc compression force (26). As a result, neck extension is thought to be helpful in relieving nerve root compression caused by bulging discs (26). Large-sample-size studies are needed to investigate complex cervical spine biomechanical changes with various maneuvers.

3) Lack of gold-stand reference

The reference standard must be reliable in order to confirm the validity of the Spurling test. However, the most critical methodological restriction of all primary studies was the lack of a true reference standard for the diagnosis of CR (16, 17, 27, 28).

Myelography cannot distinguish between cervical disc herniation and osteophytes (7). Although CT scans are sensitive to bony changes, they have a limited ability to detect soft-tissue lesions (7). MRI is advantageous because of its high sensitivity in detecting changes in the disc, spinal cord, nerve root, and soft tissue associated with CR. However, the risk of false-positive findings is high, especially in the absence of clinical information because of the requirement for double-blinding. Boden et al. reported positive MRI findings in 19% of asymptomatic patients (29).

The controversy in defining a positive MRI change in the diagnosis of CR has made MRI operator-dependent, which further lowers the reliability of MRI as a reference standard. After a detailed analysis of the original data presented in Shah et al.'s study (10), it was noted that they excluded cases with osteophytes (hard disc) as positive MRI cases. However, if any of these 20 cases (40% of the total) who were deemed to have negative findings on both MRI and the Spurling test were genuine positive CR cases, the sensitivity of the Spurling test based on their study would be much lower than 92%.

Wainner et al. (7) reported that EDX (EMG, NSC) assesses the neurophysiological status of the muscles. Therefore, they naturally have a very high specificity for confirming nerve damage in cases with suspected CR if other causes of neuropathy are excluded. *Wilbourn* et al. (30) claimed that EMG is almost 100% specific for CR diagnosis. However, EDX is highly operator-dependent, and may be insufficient to detect mild symptomatic cases with CR or cases in the acute phase (7, 12). As a result, the validity of EDX can be decreased by a high false-negative rate (12). In *Tong's* study with EMG as a standard reference, the 12 cases with negative EMG results but positive Spurling test results might contain true-positive cases for CR. Therefore, the Spurling test based on data in their study could be higher than 30% (15).

Conclusion

The Spurling test is a widely used diagnostic tool for CR. However, this topic has not yet been sufficiently studied. Multiple areas in which bias could occur inevitably decrease the reliability of the available diagnostic accuracy studies. Large sample sizes and well-validated studies are needed to guide future clinical practice.

Based on current research, CR is largely a clinical diagnosis. Although there is no single gold standard diagnostic test, the Spurling test should still be used as a first-line clinical examination tool, particularly in primary care settings. A detailed history, rational use of other Spurling variants, and other nerve provocative tests could increase the diagnostic accuracy. Cases with high clinical suspicion of CR can be referred for further testing. Cases with negative findings on first-line clinical examination should be monitored in future studies.

Declarations

There is no conflict of interest associated with this manuscript

References

1. McCartney S, Blagg S, Baskerville R, McCartney D (2018) Cervical Radiculopathy and Cervical Myelopathy: Diagnosis and Management in Primary Care. *Br J Gen Pract* 68(666):44–46
2. Apelby-Albrecht M, Andersson L, Kleiva IW, Kvåle K, Skillgate E, Josephson A (2013) Concordance of upper limb neurodynamic tests with medical examination and magnetic resonance imaging in patients with cervical radiculopathy: A diagnostic cohort study. *J Manipulative Physiol Ther* [Internet]. ;36(9):626–32. Available from: <http://dx.doi.org/10.1016/j.jmpt.2013.07.007>
3. Iyer S, Kim HJ (2016) Cervical radiculopathy. *Curr Rev Musculoskelet Med* 9(3):272–280
4. Bono CM, Ghiselli G, Gilbert TJ, Kreiner DS, Reitman C, Summers JT et al (2011) An evidence-based clinical guideline for the diagnosis and treatment of cervical radiculopathy from degenerative disorders. *Spine J* [Internet]. ;11(1):64–72. Available from: <http://dx.doi.org/10.1016/j.spinee.2010.10.023>
5. Spurling R, Scoville W (1944) Lateral Rupture of the Cervical Intervertebral Discs: A Common Cause of Shoulder and Arm Pain. *Surg Gynec Obs* 78:350–358
6. Radhakrishnan K, Litchy WJ, O'Fallon WM, Kurland LT (1994) Epidemiology of cervical radiculopathy: A population-based study from rochester, minnesota, 1976 through 1990. *Brain* 117(2):325–335
7. Wainner RS, Gill H (2000) Diagnosis and Nonoperative Management of Cervical Radiculopathy. *J Orthop Sports Phys Ther* 30(12):728–744
8. Wainner RS, Fritz JM, Irrgang JJ, Boninger ML, Delitto A, Allison S (2003) Reliability and Diagnostic Accuracy of the Clinical Examination and Patient Self-Report Measures for Cervical Radiculopathy. *Spine (Phila Pa 1976)* 28(1):52–62
9. Shabat S, Leitner Y, David R, Folman Y (2012) The Correlation between Spurling Test and Imaging Studies in Detecting Cervical Radiculopathy. *J Neuroimaging* 22(4):375–378
10. Shah KC, Rajshekhar V (2004) Reliability of Diagnosis of Soft Cervical Disc Prolapse Using Spurling's Test. *Br J Neurosurg* 18(5):480–483
11. Sleijser-Koehorst MLS, Coppieters MW, Epping R, Rooker S, Verhagen AP, Scholten-Peeters GGM (2021) Diagnostic accuracy of patient interview items and clinical tests for cervical radiculopathy. *Physiotherapy* [Internet]. ;111:74–82. Available from: <https://doi.org/10.1016/j.physio.2020.07.007>
12. Ghasemi M, Golabchi K, Mousavi SA, Asadi B, Rezvani M, Shaygannejad V et al (2013) The Value of Provocative Tests in Diagnosis of Cervical Radiculopathy. *J Res Med Sci.* ;18(SPL. 1):35–8
13. Park J, Park WY, Hong S, An J, Koh JC, Lee YW et al (2017) Diagnostic Accuracy of the Neck Tornado Test as a New Screening Test in Cervical Radiculopathy. *Int J Med Sci* 14(7):662–667
14. Viikari-Juntura E, Porras M, Laasonen EM (1989) Validity of Clinical Tests in the Diagnosis of Root Compression in Cervical Disc Disease. *Spine (Phila Pa 1976)* 14(3):253–257
15. Tong HC, Haig AJ, Yamakawa K (2002) The Spurling Test and Cervical Radiculopathy. *Spine (Phila Pa 1976)* 27(2):156–159
16. Thoomes EJ, van Geest S, van der Windt DA, Falla D, Verhagen AP, Koes BW et al (2018) Value of Physical Tests in Diagnosing Cervical Radiculopathy: A Systematic Review. *Spine J* [Internet]. ;18(1):179–89. Available from: <https://doi.org/10.1016/j.spinee.2017.08.241>
17. Rubinstein SM, Pool JJM, Van Tulder MW, Riphagen II, De Vet HCW (2007) A systematic Review of the Diagnostic Accuracy of Provocative Tests of the Neck for Diagnosing Cervical Radiculopathy. *Eur Spine J* 16(3):307–319
18. Lemeunier N, da Silva-Oolup S, Chow N, Southerst D, Carroll L, Wong JJ et al (2017) Reliability and validity of Clinical Tests to Assess the Anatomical Integrity of the Cervical Spine in Adults with Neck Pain and Its Associated Disorders: Part 1 – A Systematic Review from the Cervical Assessment and Diagnosis Research Evaluation (CADRE). *Eur Spine J* [Internet]. ;26(9):2225–41. Available from: <https://doi.org/10.1007/s00586-017-5153-0>
19. Whiting P, Rutjes AW, Beitsma JB, Bossuyt PM, Kleijnen J (2003) The Development of QUADAS: A Tool for the Quality Assessment of Studies of Diagnostic Accuracy Included in Systemic Reviews. *BMC Med Res Methodol* 3(25):1–13
20. Cleland J, Koppenhaver S, Su J, Ntter FH (2016) *Netter's Orthopaedic Clinical Examination: An Evidence-Based Approach*, 3rd edn. Elsevier, hiladelphia
21. Jaeschke R, Guyatt GH, Sackett DL, Evidence Based Medicine Working Group (1994) How to Use an Article About a Diagnostic Test. *JAMA* 271(5):389–391
22. Rhee JM, Yoon T, Riew KD (2007) Cervical Radiculopathy. *JJAAOS-ournal Am Acad Orthop Surg* 15(8):486–494

23. Farmer JC, Wisneski RJ (1994) Cervical Spine Nerve Root Compression an Analysis of Neuroforaminal Pressures with Varying Head and Arm Positions. Vol. 19, Spine. p. 1850–5
24. Yoo JU, Edwards W, Bayley J, Yuan HA (1992) Effect of Cervical Spine Motion on the Neuroforaminal Dimensions of Human Cervical Spine. Spine (Phila Pa 1976) 17(10):1131–1136
25. Anekstein Y, Blecher R, Smorgick Y, Mirovsky Y (2012) What is the Best Way to Apply the Spurling Test for Cervical Radiculopathy? Clin Orthop Relat Res 470(9):2566–2572
26. Bayoglu R, Galibarov PE, Verdonshot N, Koopman B, Homminga J (2019) Twente Spine Model: A thorough investigation of the spinal loads in a complete and coherent musculoskeletal model of the human spine. Med Eng Phys [Internet]. ;68:35–45. Available from: <https://doi.org/10.1016/j.medengphy.2019.03.015>
27. Jinright H, Kassoﬀ N, Williams C, Hazle C (2021) Spurling’s Test–Inconsistencies in Clinical Practice. J Man Manip Ther 29(1):23–32
28. Yan J, Wang Y, Liu X, Li J, Jin Z, Zheng Z (2010) Vertical weight-bearing MRI provides an innovative method for standardizing Spurling test. Med Hypotheses [Internet]. ;75(6):538–40. Available from: <http://dx.doi.org/10.1016/j.mehy.2010.07.022>
29. Boden SD, McCowin PR, Davis DO, Dina TS, Mark AS, Wiesel S (1990) Abnormal Magnetic-Resonance Scans of the Cervical Spine in Asymptomatic Subjects. A prospective investigation. J Bone Jt Surg - Ser A 72(8):1178–1184
30. Wilbourn J, Aminoff MJ (1988) AAEE Minimonograph #32: The Electrophysiologic Examination in Patients with Radiculopathy. Muscle Nerve Off J Am Assoc Electrodiagn Med 11(11):1099–1114

Tables

Table 1
Overview of literatures which investigated validity of Spurling test in diagnosing cervical radiculopathy

Studies	1st author, year, country	N (M: F)	Participants description	Selection criteria	Exclusion criteria	Spurling test brief description	Positive test definition	Reference standard	Sn (95% CI)
1	Viikari-Juntura(14) 1989 Finland	43 (27:16)	Consecutive patients presented to physiology department (March 1982-January 1985)	Cervical disc disease such as spondylosis and/or disc herniation	Cervical spine cancer. Cervical spine malformation. RA	Lateral bending Slight rotation compression (7kg)	Radicular pain, paresthesia or numbness (ipsilateral or contralateral side)	Myelography	46.0%
2	Tong(15) 2002 USA	255	Consecutive Patients referred for EMG (1988-1993)	Upper extremity nerve disorder	unclear	Lateral flexion neck extension Axial spine compression	Pain/tingling radiating from shoulder distally to elbow	EMG	30.0%
3	Wainner (8) 2003 USA	82 (41:41)	Consecutive patients from 4 medical facilities enrolled in the study (Dec 1998-April 2000)	Patients suggestive of CR or CTS referred to the electrophysiologic lab	Disease causing PN; Bilateral arm pain; Work disruption > 6months; Surgery that causes neck pain or CTS; Previous EMG/NCS for CR, CTS or bot	Spurling A: Lateral bending & 7kg axial pressure Spurling B: Extension + lateral bending, rotation & 7kg axial pressure	Reproduction of CR symptoms	Needle EMG + NCS	50% 50%
4	Shah(10) 2004 India	50 (37:13)	Patients presented to Neurosurgical units (May 2001-May 2003)	Patients with neck, arm pain / CR; Age 22-60y. Duration 2weeks-36 months	Neck trauma, myelopathy, prior surgery on the cervical spine	Extension, lateral bending & axial compression	If radicular pain or paraesthesia in upper limb reproduced or aggravated	MRI, or surgical findings	92.0%
5	Shabat (9) 2012 Israel	257 (159: 98)	Patients presented to Spinal Surgery unit	Patients with symptoms of unilateral CR lasting ≥ 4 weeks	Patients with upper motor neuron signs or cervical myelopathy	Extension Rotation & axial Compression	Radicular pain to specific dermatome (121), Occiput /scapula (84)	CT (all) ± MRI (21%)	95.0%
6	Ghasemi (12) 2013 Iran	97 (25:72)	Patients referred to Electrodiagnostic center (Jan 2010- Mar 2011)	≥ 20 years old with Neck + arm pain with sensory and/or motor and/or reflex changes of upper limbs ≥ 3 weeks	Patient with history of neck trauma, surgery, tumours or congenital C-spine disease; Disease causing peripheral neuropathies (ie. RA)	Lateral bending & Axial Compression	Reproduction of symptoms	EMG A C	46.5% 14.7%
7	Park(13) 2017 Korea	135 (78:57)	Patients referred to pain clinic with a history neck pain (Sep 2012-Aug 2015)	Neck pain ± CR (Confirmed by experienced pain clinician + MRI findings)	C- spine surgery, inflammatory disease, pregnancy, or previous nerve block for CR	Extension Lateral bending & Axial Compression	Reproduction of radicular pain or paraesthesia	MRI	55.2%

Studies	1st author, year, country	N (M: F)	Participants description	Selection criteria	Exclusion criteria	Spurling test brief description	Positive test definition	Reference standard	Sn (95% CI)
8	Sleijser-Koehorst(11) 2021 Australia	134 (68:66)	Consecutive patients referred to MDT clinic with neck and/or arm pain; Age ≥ 18 years old; Understand Dutch	CR was diagnosed by Neurosurgeon based on clinical symptoms and MRI	Patients with severe neck pathology (cancer, RA, fracture, past surgery, myelopathy), neurological condition (MS, DM, CRPS, polyneuropathy).	Extension, Rotation Lateral bending +/- axial compression.	Presence of paresthesia numbness and weakness	Clinical symptoms; MRI	59.0%

N = number of patients; M: F = male-to-female ratio. Sn: sensitivity; Sp: specificity. +LR: positive likelihood ratio; -LR: negative likelihood ratio.

CR: cervical radiculopathy. PN: peripheral neuropathy; CTS: carpal tunnel syndrome; NCS: nerve conduction study

EMG: electromyography study MDT: multidisciplinary clinic

A: acute cervical radiculopathy; C: chronic cervical radiculopathy.

RA: Rheumatoid arthritis; MS: multiple sclerosis; DM: Diabetes Mellitus. CRPS: complex regional pain syndrome.

Table 2

Methodological quality evaluation of studies which investigated diagnostic accuracy of Spurling test for patients with cervical radiculopathy

Studies	1st Author, (reference), year	QUADAS criteria number *														Quality score (No. criteria met)
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	
1	Viikari-Juntura (14), 1989	-	+	-	+	+	+	+	+	+	+	+	-	-	-	9 Δ
2	Tong (15), 2002	-	-	-	?	+	+	+	+	+	+	?	+	?	-	7 Δ
3	Wainner (8), 2003	-	+	-	+	+	+	+	+	+	+	+	+	?	+	11 ©
4	Shah (10), 2004	+	+	-	?	+	-	+	+	+	+	+	+	-	-	9 Δ
5	Shabat (9), 2012	+	+	-	?	+	+	+	+	+	+	?	+	-	-	9 Δ
6	Ghasemi (12), 2013	-	+	-	?	+	+	+	+	+	+	?	+	-	+	9 Δ
7	Park (13), 2017	-	+	-	?	+	+	+	-	+	?	-	+	-	-	6 Δ
8	Sleijser-Koehorst (11) 2021	-	+	-	+	+	+	+	+	+	+	+	+	+	+	12 ©

Scoring: [+] met criteria; [-] did not meet criteria; [?] unclear.

© Good quality [+] = 10–14; Δ Fair quality: [+] = 5 to 9. ☹ Poor quality [+] ≤ 4

* QUADAS assessment criteria (19)

1. Was the spectrum of patients representative of the patients who will receive the test in practice?
2. Were selection criteria clearly described?
3. Is the reference standard likely to correctly classify the target condition?
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?
6. Did patients receive the same reference standard regardless of the index test result?
7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
8. Was the execution of the index test described in sufficient detail to permit replication of the test?
9. Was the execution of the reference standard described in sufficient detail to permit its replication?
10. Were the index test results interpreted without knowledge of the results of the reference standard?
11. Were the reference standard results interpreted without knowledge of the results of the index test?
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
13. Were uninterpretable/ intermediate test results reported?
14. Were withdrawals from the study explained?

Adapted from Whiting P, Rutjes AW, Beitsma JB, Bossuyt PM, Kleijnen J. The Development of QUADAS: A Tool for the Quality Assessment of Studies of Diagnostic Accuracy Included in Systemic Reviews. BMC Med Res Methodol. 2003;3(25):1–13.



Figure 1

Spurling test (Spurling R., Scoville W. Lateral rupture of the Cervical Intervertebral Discs: A Common Cause of Shoulder and Arm Pain. Surg Gynec Obs. 1944;78:350-8)(5).

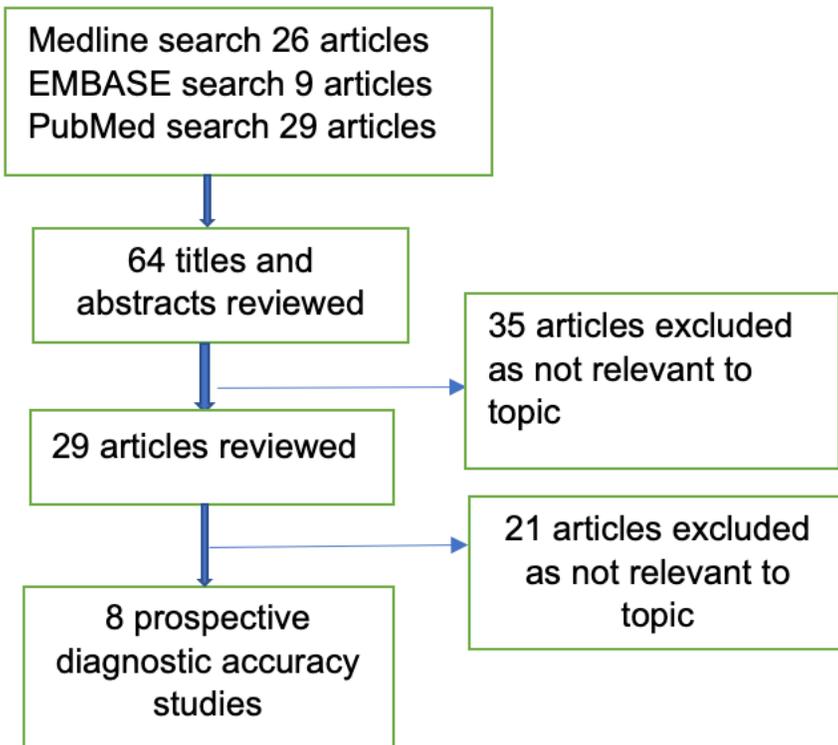


Figure 2

Literature search flow diagram