

UNIVERSITY OF BIRMINGHAM

University of Birmingham
Research at Birmingham

Diagnostic accuracy of upper limb neurodynamic tests for the assessment of peripheral neuropathic pain

Koulidis, Konstantinos; Veremis, Yannis; Anderson, Christina; Heneghan, Nicola

DOI:

[10.1016/j.msksp.2019.01.001](https://doi.org/10.1016/j.msksp.2019.01.001)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Koulidis, K, Veremis, Y, Anderson, C & Heneghan, N 2019, 'Diagnostic accuracy of upper limb neurodynamic tests for the assessment of peripheral neuropathic pain: a systematic review', *Musculoskeletal Science and Practice*, vol. 40, pp. 21-33. <https://doi.org/10.1016/j.msksp.2019.01.001>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

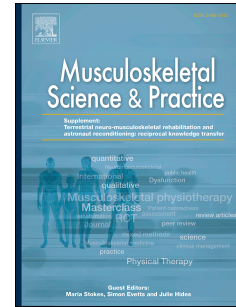
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Accepted Manuscript

Diagnostic accuracy of upper limb neurodynamic tests for the assessment of peripheral neuropathic pain: A systematic review

Konstantinos Koulidis, Yannis Veremis, Christina Anderson, Nicola R. Heneghan



PII: S2468-7812(18)30180-2

DOI: <https://doi.org/10.1016/j.msksp.2019.01.001>

Reference: MSKSP 1970

To appear in: *Musculoskeletal Science and Practice*

Received Date: 16 May 2018

Revised Date: 18 December 2018

Accepted Date: 2 January 2019

Please cite this article as: Koulidis, K., Veremis, Y., Anderson, C., Heneghan, N.R., Diagnostic accuracy of upper limb neurodynamic tests for the assessment of peripheral neuropathic pain: A systematic review, *Musculoskeletal Science and Practice* (2019), doi: <https://doi.org/10.1016/j.msksp.2019.01.001>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

TITLE PAGE

Diagnostic accuracy of upper limb neurodynamic tests for the assessment of peripheral neuropathic pain: A systematic review

Konstantinos Koulidis^a, Yannis Veremis^b, Christina Anderson^c, Nicola R Heneghan^d,

^a *MSc in Advanced Manipulative Physiotherapy, School of Sport, Exercise & Rehabilitation Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK.*

E-mail address: kostaskoul23@gmail.com

^b *MSc in Advanced Manipulative Physiotherapy, School of Sport, Exercise & Rehabilitation Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK.*

E-mail address: veremis.y@gmail.com

^c *Teaching Associate in Physiotherapy, School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK*

E-mail address: c.b.anderson@bham.ac.uk

^d *Lecturer in Physiotherapy, Centre of Precision Rehabilitation for Spinal Pain, School of Sport, Exercise & Rehabilitation Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK.*

E-mail: n.heneghan@bham.ac.uk

Corresponding author: Nicola R Heneghan.

E-mail: n.heneghan@bham.ac.uk

Telephone: +44 121 4158367

Postal address:

Centre of Precision Rehabilitation for Spinal Pain,
School of Sport, Exercise & Rehabilitation Sciences,
University of Birmingham,
Birmingham, B15 2TT, UK.

1. Conflict of Interest: none declared
2. Ethical Approval: None required
3. Funding: none
4. Acknowledgements: None required

ABSTRACT

Background: Upper limb neurodynamic tests (ULNTs) are used to identify a neuropathic pain component in patients' presenting with arm and/or neck pain. Clinical tests with established diagnostic accuracy are required to not only to inform clinical management but also minimise costs associated with expensive medical investigations.

Objective: To evaluate the role of ULNTs in assessment of peripheral neuropathic pain and to inform their value in clinical practice when assessing patients with arm and/or neck symptoms.

Design: Systematic review was undertaken according to published guidelines, and reported in line with PRISMA-DTA.

Method: Key databases were searched up to 21/11/2017. Inclusion criteria: Patient population experiencing arm and/or neck symptoms with suspected peripheral neuropathic involvement, studies that compared ULNT to a reference standard, any study design using primary diagnostic accuracy data. Two reviewers independently assessed risk of bias (ROB) using QUADAS-2. The overall quality of evidence was evaluated using GRADE.

Results: Of eight included studies (n=579), four were assessed as low ROB, although all had concerns regarding applicability. For carpal tunnel syndrome, ULNT1 sensitivity values ranged 0.4-0.93, specificity 0.13-0.93, positive likelihood ratio 0.86-3.67 and negative likelihood ratio 0.5-1.9. For cervical radiculopathy ULNT1 and the combined use of four ULNTs had sensitivity of 0.97 (95%CI 0.85-1.00) whereas the ULNT3 was the most specific (0.87, 95%CI 0.62-0.98). Positive likelihood ratio ranged 0.58-5.68 and negative likelihood ratio 0.12-1.62.

Conclusion: Based on the available evidence ULNTs cannot be utilised as a stand-alone test for the diagnosis of CTS. Limited evidence suggests that ULNTs may be clinically relevant for the diagnosis of CR, but only as a "ruling out" strategy. However, the overall quality of the body of evidence after applying the GRADE approach was low to very low across studies. Further higher quality research is needed to establish firm conclusions.

Key words: entrapment neuropathies, carpal tunnel syndrome, cervical radiculopathy, upper limb neurodynamics, validity

Word count 3685

37

INTRODUCTION

38 Peripheral neuropathic pain (PNP) is a term used to describe pain that results from a lesion or
39 disease affecting the somatosensory nervous system (Finnerup et al., 2016). PNP can arise when a
40 peripheral nerve trunk or a nerve root has been subject to injury, compression, inflammation or
41 ischemia resulting in reduced physical capabilities of the nervous system (Nee and Butler, 2006).
42 Symptoms and signs in neuropathies can be classified as positive (gain of function) or negative (loss
43 of function). Positive symptoms include pain, paresthesia, dysesthesia, hyperalgesia and allodynia
44 and indicate abnormal excitability in the nervous system, whereas negative symptoms, such as
45 hypoesthesia or anesthesia and weakness reflect reduced impulse conduction (Woolf, 2004).

46 The most common conditions affecting the peripheral nervous system are entrapment neuropathies
47 (EN), with carpal tunnel syndrome (CTS), cubital tunnel syndrome and cervical radiculopathy (CR)
48 being examples which contribute considerably to the socioeconomic burden of occupational related
49 musculoskeletal complaints and the associated costs. Individually EN have been associated with
50 severe pain, depression and functional limitations (Fernandez-de-las-Penas et al., 2015). CTS is often
51 observed in activities involving repetitive manual tasks, forceful wrist movements or with direct
52 pressure on the wrist, estimated to affect 2-15% of workers (Atroshi et al., 1999) and costing more
53 than 2 billion dollars each year in the USA (work absenteeism, medical evaluation, treatment) (Saint-
54 Lary et al., 2015). In the case of CR, the data regarding the prevalence and the epidemiology of the
55 condition are very limited. The reported annual incident of CR is 83.2 per 100.000 persons (107.3 for
56 men and 63.5 for women) with a peak incidence in the fifth and sixth decade for both genders
57 (Radhakrishnan et al., 1994).

58 The diagnosis of EN is based on information received during the subjective (history taking) and
59 physical examination, which is then confirmed via diagnostic imaging or electrophysiological studies.
60 Clinical examination of EN encompasses a variety of tests (sensation, muscle strength and reflexes)
61 assessing the integrity and ability of the nervous system to conduct afferent or efferent impulses
62 (loss of function) (Baselgia et al., 2017). In addition, a thorough examination includes evaluation of
63 increased mechanical sensitivity of the nervous system, since PNP can be present without or with
64 minimal loss of nerve conduction (Schmid et al., 2009). Diagnostic imaging and electrophysiological
65 studies are most commonly used to establish a diagnosis of EN (Wainner et al., 2003). For most
66 clinicians, these methods are accessible but given the waiting time for patients and the high cost for
67 the society it would be useful to establish accurate clinical examination tests for the diagnosis of EN.

68 Neurodynamic tests are used by musculoskeletal physiotherapists in order to identify changes of
69 mechanosensitivity in the nervous system, thus assessing gain of function (Baselgia et al., 2017). Due
70 to the interdependence of the mechanical, electrical and chemical properties of the nervous system,
71 changes in one of these features may affect the others (Butler, 2008). Impairments in the
72 surrounding musculoskeletal structures could apply mechanical or chemical stimuli to a nerve,
73 resulting in venous congestion, impaired axoplasmic flow, inflammation and development of
74 mechanosensitive abnormal impulse generating sites (Nee and Butler, 2006).

75 For disorders affecting the upper limbs four different neurodynamic tests have been proposed to
76 assess mechanosensitivity of the brachial plexus, medial, radial and ulnar nerve (Elvey, 1980)(Table
77 1). Where symptoms are not related to central pain mechanisms (broader distribution of symptoms
78 due to central sensitization e.g. in case of persistent pain) a positive test response could be
79 associated with neural or non-neural tissue sensitivity. A neurodynamic test is considered positive if
80 it can reproduce the patient's own symptoms and if those symptoms can be altered through
81 structural differentiation (Butler, 2000). Schmid and colleagues (2009) assessed the reliability of
82 ULNTs and found that those tests have moderate reliability. Wainner et al. (2003, 2005) reported
83 substantial to almost perfect reliability for the interpretation of the ULNT1 (median) and ULNT2b
84 (radial).

85
86 Although used by clinicians the diagnostic accuracy of upper limb neurodynamic tests (ULNTs) has
87 not yet been fully established and is important to optimise patient care. A recent systematic review
88 has summarized the evidence on diagnostic performance of tests (including ULNTs) which are
89 utilized for the identification of CR and concluded that when consistent with patient history, a
90 combined result of four negative ULNTs (high sensitivity) and a negative Arm Squeeze test could be
91 used to rule out the disorder (Thoomes et al., 2017). Likewise an earlier systematic review,
92 concluded that a positive Spurling's, traction/neck distraction, and Valsalva's test might be indicative
93 of CR, while a negative ULNT1 might be used to rule it out (high sensitivity)(Rubinstein et al., 2007).
94 Of the eight included studies in this systematic review only two had assessed the diagnostic accuracy
95 of ULNTs. Finally in a previous clinical commentary the authors attempted to summarise the
96 available evidence in regard to the diagnostic usefulness of neurodynamic tests (Nee et al., 2012).
97 The authors, based on biomechanical and experimental studies, concluded that ULNTs can
98 potentially distinguish pain related to neural mechanosensitivity from pain arising from other
99 tissues, and therefore could detect PNP. In the view of the growing body of evidence, a systematic
100 review is required to evaluate the quality and synthesis the available current evidence of the
101 diagnostic accuracy of ULNTs and to inform clinical practice. The aim therefore of this study was to

102 examine the intended role of ULNTs in assessment of PNP, by answering the following research
 103 question: What is the diagnostic accuracy of ULNTs when compared to diagnostic imaging or
 104 electrophysiologic studies, and how results from ULNTs can be interpreted when assessing patients
 105 with arm and/or neck symptoms?
 106

Order of movements	ULNT1 (median)	ULNT2a (median)	ULNT2b (radial)	ULNT3 (ulnar)
1	Shoulder depression	Shoulder depression	Shoulder depression	Shoulder depression
2	Shoulder abduction 110°	Elbow extension	Elbow extension	Shoulder abduction 100°
3	Wrist and fingers extension	Lateral rotation of the arm	Medial rotation arm	Lateral rotation arm
4	Forearm supination	Wrist and finger extension	Wrist and fingers flexion	Forearm pronation
5	Shoulder lateral rotation	Shoulder abduction 10°	Shoulder abduction	Elbow flexion
6	Elbow extension	Contralateral lateral flexion of the cervical spine	Contralateral lateral flexion of the cervical spine	Wrist and fingers extension
7	Contralateral lateral flexion of the cervical spine			Contralateral lateral flexion of the cervical spine

107
 108
 109

Table 1. ULNT procedure

110

DESIGN AND METHODS

111 This systematic review was conducted according to a pre-defined protocol based on the Cochrane
 112 Handbook for Diagnostic Test Accuracy studies (Deeks, Wisniewski and Davenport, 2013) and the
 113 Centre for Reviews and Dissemination (CRD, 2009). In addition, the study is reported according to
 114 Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy
 115 Studies (PRISMA-DTA) (McInnes et al., 2018). (Appendix 1)

Search strategy

117 Informed by subject (NH, KK, YV) and methodological experts (NH, CA) key bibliographic databases
 118 were searched independently by two reviewers (KK, YV). The search employed sensitive topic-based
 119 strategies designed for each database from inception to 21st November 2017. Databases of interest
 120 were: PEDro, MEDLINE (through PubMed), AMED, CINAHL, Cochrane Library, and EMBASE. The
 121 search strategy, informed by scoping search included MeSH terms and text words, as well as a

122 combination of both for a comprehensive search. The following keywords and combination of them
 123 were used: upper limb neurodynamic test, neural provocation test, upper limb tension test,
 124 diagnosis, peripheral neuropathic pain, peripheral entrapment neuropathy, radicular pain, cervical
 125 radiculopathy, brachial plexus, carpal tunnel syndrome, cubital tunnel syndrome, accuracy,
 126 specificity, sensitivity, validity.

127 The search was augmented using reference lists of included studies, as well as searching the grey
 128 literature. Box 1 details the MEDLINE search strategy.

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

1. peripheral neuropathic pain.mp or exp Neuralgia/
2. radicular pain.mp or exp Hereditary Sensory and Autonomic Neuropathies/
3. peripheral entrapment neuropathy.mp
4. cervical radiculopathy.mp or exp Radiculopathy/
5. carpal tunnel syndrome.mp or exp Carpal tunnel syndrome/
6. cubital tunnel syndrome.mp or exp Cubital tunnel syndrome/
7. brachial plexus neuropathies.mp or exp Brachial plexus neuropathies/
8. exp Nerve compression syndromes/
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. upper limb neurodynamic test.mp
11. upper limb tension test.mp
12. neural provocation test.mp
13. exp Diagnosis/
14. exp Pain measurements/
15. exp Neurologic examination/
16. exp Physical examination/
17. 10 or 11 or 12 or 13 or 14 or 15 or 16
18. diagnostic accuracy.mp
19. sensitivity and specificity.mp or exp Sensitivity and specificity/
20. validity.mp
21. exp Reproducibility of results/
22. exp Predictive value of tests/
23. 18 or 19 or 20 or 21 or 22
24. 9 and 17 and 23

Box 1. MEDLINE search strategy

155

156

157 Eligibility criteria

158 Eligibility criteria were established following the recommendations of The Cochrane Handbook for
 159 Diagnostic Test Accuracy studies (Bossuyt and Leeflang, 2008) and informed using the SPIDER search
 160 concept (Cooke, Smith and Booth, 2012). Titles and abstract of the identified studies were screened
 161 by two independent reviewers (KK, YV) for eligibility using pre-specified inclusion criteria.

162 Inclusion criteria (based on SPIDER) included that the sample (S) comprised populations aged > 18
163 years with arm and/or neck symptoms with suspected peripheral neuropathic involvement (signs
164 and symptoms suggesting excitability in the nervous system such as pain, paresthesia, dysesthesia,
165 spasm or reduced impulse conduction such as hypoesthesia or anesthesia and weakness)(Nee and
166 Butler, 2006); the phenomenon of interest (PI) was the diagnostic accuracy of ULNTs; investigated
167 using a diagnostic accuracy study design (D); with comparison of the index test (ULNTs) to a
168 reference standard, such as, electrophysiologic examination (electromyography and nerve
169 conduction studies) or advanced imaging (e.g. Magnetic Resonance Imaging (MRI), CT, myelography)
170 (E). Although not perfect, these tests are considered to be the most accurate diagnostic tests
171 available (Wainner, et al., 2003; Jablecki et al., 1993, 2002; Kuijper et al., 2009;).

172 Exclusion criteria: case series, case reports, surgical or cadaveric studies; publications for which full
173 text not available.

174 Quality assessment

175 Two reviewers (KK, YV) independently conducted the risk of bias (ROB) assessment using the Quality
176 Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) - tool, a development of the original tool
177 (Whiting et al., 2011). It consists of four key domains: patient selection, index test, reference
178 standard and, flow and timing. All key areas are assessed for ROB, whereas the first three are also
179 assessed in terms of applicability to the review question. Each domain is judged as “high risk”, “low
180 risk” or “unclear risk” based on signaling questions aiming to assist judgment (Whiting et al., 2011).
181 Overall, a study can be judged as having “low risk of bias” if every domain has been ranked as “low
182 risk”. Assessment of applicability is based on the first three domains and whether they are in line
183 with the review question. The study is judged as having “no concerns” regarding applicability if these
184 domains are in line with the review question and “with concerns” if deviates from the review
185 objective. The QUADAS-2 has been used in recent systematic reviews (Grørdahl et al., 2016; Hegedus
186 et al., 2012) and is recommended by the Cochrane Collaboration and the U.K National Institute for
187 Health and Clinical Excellence (Reitsma et al., 2009).

188

189 Data extraction

190 Diagnostic accuracy data and study characteristics were extracted by one reviewer (KK) using a pre-
191 designed data extraction sheet which covered five areas. The data were audited by a second
192 reviewer (YV) for accuracy. The following data were extracted: authors and publication details,
193 studies’ methods (aim of study, study design, method of recruitment, eligibility criteria, and ethical

194 approval), participant details, diagnostic test data (sensitivity, specificity, predictive values,
195 likelihood ratios and other). Finally, the fifth section was 2x2 contingency tables for the diagnostic
196 tests.

197 Summary measures

198 Sensitivity, specificity, likelihood ratios (LR) and predictive values (PV) were the outcomes for which
199 data were sought. True positive, false positive, true negative and false negative values were
200 summarised. In cases where only incomplete or raw data were presented, a 2x2 contingency table
201 was used to re-estimate these values. Sensitivity and specificity were graded as low (<0.50),
202 low/moderate (0.51-0.64), moderate (0.65-0.74), moderate/high (0.75-0.84) and high (>0.85) in line
203 with previous systematic reviews of diagnostic accuracy studies (Grødahl et al., 2016; Schneiders et
204 al., 2012). Clinical interpretation of likelihood ratios was based on Jaeschke et al. (1994) as follows:
205 conclusive evidence (LR+>10 and LR-<0.1), strong diagnostic evidence (LR+ 5 to 10 and LR- 0.1 to
206 0.2), weak diagnostic evidence (LR+2 to 5 and LR- 0.2 to 0.5) and negligible evidence (LR+ 1 to 2 and
207 LR- 0.5 to 1).

208 Data analysis

209 Homogeneity among studies was explored to evaluate if the studies were suitable for combining in a
210 meta-analysis. Areas of exploration were: study designs, patient population, comparable reference
211 tests and diagnostic data, no differences in diagnostic thresholds (Burgess et al., 2011). In addition,
212 quality assessment of the included studies was conducted, since studies with high ROB often over-
213 estimate the performance of a test (Lijmer et al., 2002). Given the heterogeneity of the included
214 studies a narrative synthesis was undertaken.

215 Quality of evidence across studies

216 Quality of evidence, including risk of bias across studies was evaluated using GRADE (Schunemann et
217 al, 2008) for individual tests. Quality of overall body of evidence is influenced by amongst other
218 factors, study design, patient populations, precision, consistency, directness and as such each
219 outcome was evaluated by both reviewers independently (Schunemann et al, 2008).

220

221

222

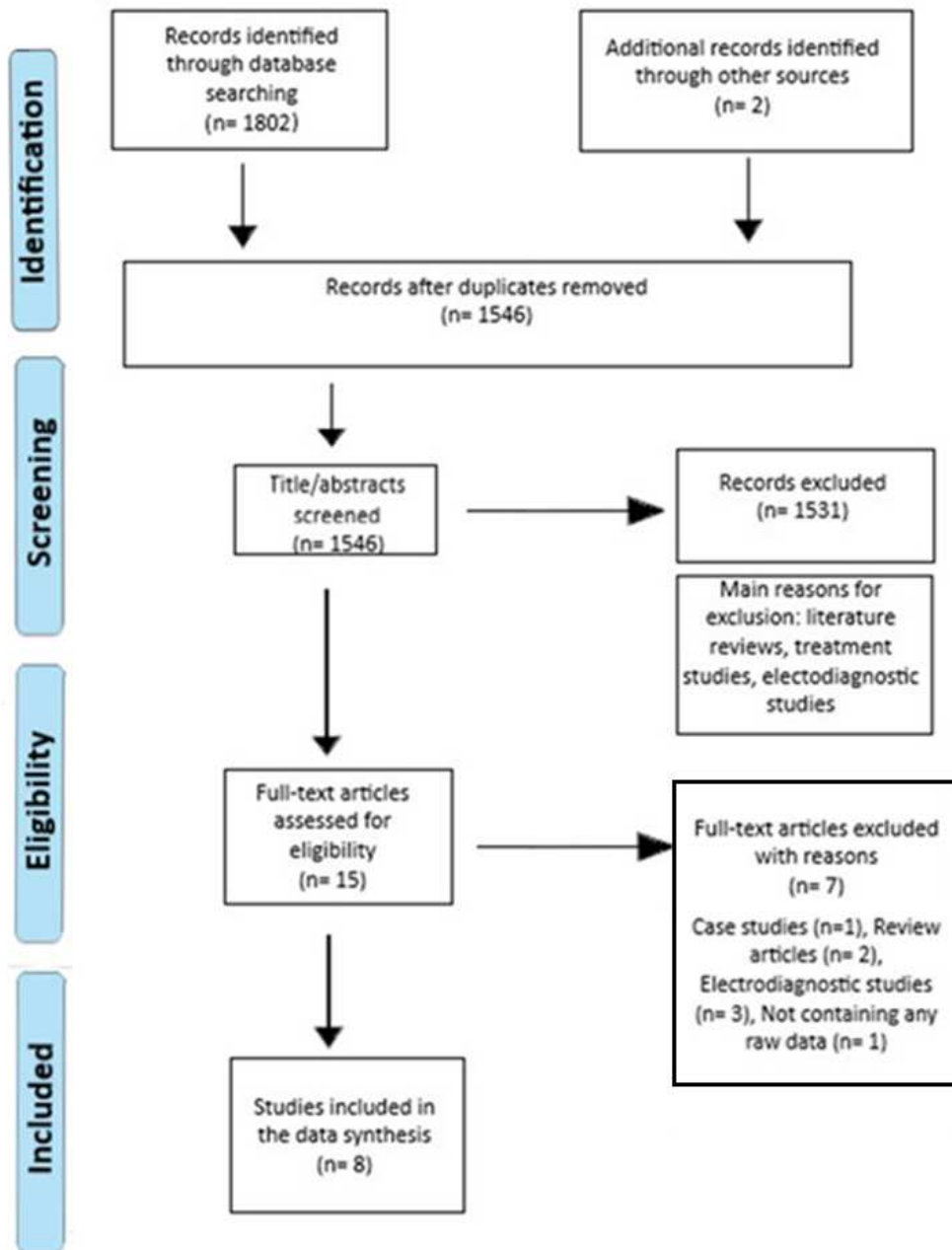
223

224

RESULTS

225 *Study identification*

226 The searches identified 1802 studies with screening of title and abstract resulting in 15 studies that
 227 were retrieved for full-text evaluation and 8 studies (n=579) meeting the eligibility requirements for
 228 inclusion. (Fig.1). There was 100% of agreement between the reviewers on selecting studies.



229

230

Fig.1. PRISMA flow diagram for systematic reviews

231 Study description

232 Table 2 summarises the specific characteristics of all eight studies. Three studies investigated the
233 diagnostic accuracy of ULNTs in individuals with suspected CR (Wainner et al., 2003; Apelby-Albrecht
234 et al., 2013; Ghasemi et al., 2013). Two of the studies used electrophysiologic procedures as the
235 reference standard (Wainner et al., 2003; Ghasemi et al., 2013). One study used MRI, clinical
236 examination and history as a reference standard (Apelby-Albrecht et al., 2013). Five studies
237 investigated the diagnostic accuracy of ULNTs in individuals with suspected CTS with nerve
238 conduction studies as the reference standard (Wainner et al., 2005; Vanti et al., 2011, 2012; Bueno-
239 Gracia et al., 2016; Trillos et al., 2017;).

240

241

242

243

244

245

246

247

248

249

250

251

252

253

Author. (year), country	Type of study	Pathology	Setting	Inclusion/ Exclusion criteria	Population (Number, gender, age)	Outcome measures	Reference Standard	ROB
Apelby-Albrect et al. (2013) Sweden	Prospective cohort	Cervical radiculopathy	Center for spinal surgery	Inclusion: neck/ arm pain Exclusion: History of multitrauma, malignant, system disease with possible neuropathy, or patients whose general condition (physically or/and psychologically) could influence the results.	N= 51 Women n=27 Men n=24 Mean age: 51 (25-67) years	ULNT (1, 2a, 2b, 3) Combined and individually	MRI, Clinical examination, Patient history	At risk
Ghasemi et al. (2013) Iran	Cross-sectional	Cervical radiculopathy	Electrodiagnostic center (hospital)	Inclusion: Aged > 20 years, symptoms of neck/ radicular pain > 3 weeks Exclusion: History of neck trauma, prior surgery, tumors or congenital abnormality of cervical spine, any systemic situation known to cause peripheral neuropathies and known cases of rheumatoid arthritis	N= 97 Women n=72 Men n=25 Mean age: Women 46.14 ±11.45 Men 46.32 ±13.97 years	ULNT 1 (median)	NCS	At risk
Wainner et al. (2003)	Prospective cohort	Cervical radiculopathy	University of Pittsburgh, Wilford Hall USAF Medical Center, Brooke Army Medical Center, and Blanchfield Army Community Hospital	Inclusion: signs and symptoms compatible with CR or CTS Exclusion: systemic disease, primary report of bilateral radiating arm pain, history of conditions involving the affected upper extremity or surgical procedures for pathologies giving rise to neck pain or CTS, discontinuation of work > 6 months, previous EMG and NCS testing the symptomatic limb for CR, CTS, or both	N= 82 Women n=41 Men n=41 Mean age: 45 ± 12 years	ULNT 1 (median), ULNT 2b (radial)	Needle EMG and NCS	Low risk

Bueno-Gracia et al. (2016) Spain	Prospective cohort	Carpal tunnel syndrome	Not reported	Inclusion: patients with hand, wrist or forearm symptoms Exclusion: any ROM limitations of the upper limb, inability to lie supine, any physical contraindications for physical therapy, presence of any cognitive or communicative deficits	N= 58 Women n=42 Men n=16 Mean age: 54.3 ± 14.5 years	ULNT 1 (median)	NCS and clinical presentation	At risk
Trillos et al. (2017) Colombia	Prospective cohort	Carpal tunnel syndrome	Health service institution	Inclusion: age 18-86, referred with a clinical diagnosis of CTS Exclusion: upper limb joint and cervical spine pathologies, patients with history of rheumatoid arthritis, anterior shoulder dislocation, CRPS, Raynaud's syndrome, breast cancer, RC injuries, patients with cervical spinal stenosis, or cognitive deficits	N=118 Women n=98 Men n=20 Mean age: 50.51 ±11.1 years	ULNT 1 (median)	NCS	Low risk
Vanti et al. (2011) Italy	Prospective cohort	Carpal tunnel syndrome	Clinic of Occupational Medicine of the University of Bologna (Italy)	Inclusion: individuals with suspected CTS Exclusion: upper limb joint pathologies inflammatory, infective or systemic pathologies, history of surgical procedure for CTS, CR, cognitive deficits	N= 44 Women n=33 Men n=11 Mean age: 46.3 ±10.8 years	ULNT 1 (median)	NCS	At risk

Vanti et al.(2012) Italy	Prospective cohort	Carpal tunnel syndrome	Occupational Medicine of the Department of Internal Medicine, Geriatrics and Nephrology, Alma Mater Studiorum, University of Bologna (Italy)	Inclusion: individuals with suspected CTS Exclusion: upper limb joint pathologies that could significantly limit the ROM of the upper limbs; inflammatory, systemic, or infectious diseases; history of surgical intervention for CTS; CR; and cognitive deficits	N= 47 Women n=35 Men n=12 Mean age: 45.9 ± 10.6 years	ULNT 1 (median)	NCS	Low risk
Wainner et al. (2005)	Prospective cohort	Carpal tunnel syndrome	Multicenter medical center and community hospital	Inclusion: signs and symptoms compatible with CR or CTS Exclusion: systemic disease, primary report of bilateral radiating arm pain, history of conditions involving the affected upper extremity or surgical procedures for pathologies giving rise to neck pain or CTS, discontinuation of work > 6 months, previous EMG and NCS testing the symptomatic limb for CR, CTS, or both	N= 82 Women n=41 Men n=41 Mean age: 45 ±12 years	ULNT 1 (median) ULNT 2b (radial)	NCS and clinical presentation	Low risk

ROM: Range of motion, ULNT: Upper limb neurodynamic test, NCS: Nerve conduction studies, CTS: Carpal tunnel syndrome, CRPS: Complex regional pain syndrome, RC: Rotator cuff, CR: Cervical radiculopathy, EMG: Electromyography

Table 2. Characteristics of included studies

254 *Risk of bias assessment*

255 Agreement of risk of bias following discussion was excellent (100%). Four studies were
 256 assessed as “low risk of bias” (ROB) (Wainner et al., 2003, 2005; Vanti et al., 2012; Trillos et
 257 al., 2017), but all of them had concerns with regards to applicability (Table 3). Patient
 258 selection procedures and poor reporting of flow and timing were the main areas of ROB.
 259 Only two studies were assessed as no concerns for applicability (Fig. 2) (Apelby-Albrecht et
 260 al., 2013; Bueno-Gracia et al., 2016). Interpretation of the index test was the main reason for
 261 concern regarding applicability since it was not in agreement with our review question. In
 262 our study an ULNT is considered positive only when it reproduces the patient’s clinical
 263 symptoms and those symptoms are modified with structural differentiation (Nee et al.,
 264 2012; Butler, 2000; Coppieters et al., 2002).

265

266

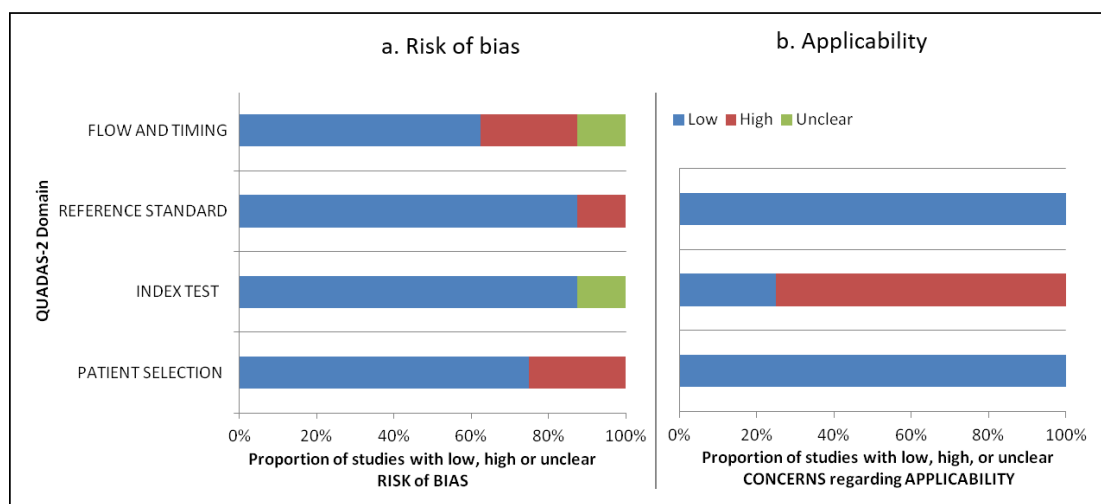
Study	RISK OF BIAS				Summary	APPLICABILITY CONCERNS			Summary
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING		PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	
Apelby-Albrecht et al., 2013					At risk				No concern
Bueno-Gracia et al., 2016					At risk				No concern
Ghasemi et al., 2013					At risk				With concern
Trillos et al., 2017					Low risk				With concern
Vanti et al., 2011					At risk				With concern
Vanti et al., 2012					Low risk				With concern
Wainner et al., 2003					Low risk				With concern
Wainner et al., 2005					Low risk				With concern

267

268

Table 3. Risk of bias assessment of included studies

269



270 **Fig.2.** Proportion of studies assessed as low, high or unclear ROB and/or applicability.
 271
 272

273 Synthesis of results

274 The main limitations for performing a meta-analysis were the heterogeneity in terms of the
 275 reference standard utilised, as well as in the interpretation of the index test and the
 276 methodological quality of the included studies. Since a meta-analysis was not possible,
 277 diagnostic accuracy data (sensitivity, specificity, predictive values and likelihood ratios) are
 278 presented using a narrative approach. The overall body of the evidence in terms of ROB,
 279 inconsistency, indirectness, imprecision, and the presence of potential reported bias after
 280 applying the GRADE approach was low to very low across studies and across outcomes.
 281 Diagnostic accuracy for all clinical indicators is summarised in Table 4 and 5 and outcome of
 282 GRADE evaluation in Table 6 and 7.

283 Diagnostic accuracy of Upper Limb Neurodynamic tests

284 *Carpal tunnel syndrome*

285 Five studies examined the diagnostic accuracy of ULNTs in patients with suspected CTS
 286 (Wainner et al., 2005; Vanti et al., 2011, 2012; Bueno-Gracia et al., 2016; Trillos et al., 2017).
 287 From these studies two were at ROB (Vanti et al., 2011; Bueno-Gracia et al., 2016) and four
 288 had concerns regarding applicability (Wainner et al., 2005; Vanti et al., 2011, 2012; Trillos et
 289 al., 2017). Those at ROB had limitations related to patient selection and flow and timing. The
 290 study of Vanti et al. (2011) was at ROB because the number of patients enrolled in the study
 291 was different from the number of patients that were included in the analysis (Whiting et al.,
 292 2011), whereas in the study by Bueno-Gracia et al. (2016) the authors provided limited
 293 information in regards to the methods used for the enrollment of the sample (consecutive or

294 random sample). The studies that had concerns regarding applicability used a definition for a
295 positive ULNT that differs from that being used in this review.

296 Three studies assessed the validity of ULNT1 (median) considering the test positive in the
297 presence of only one of the following criteria: 1) reproduction of patient's symptoms; 2) side
298 to side differences ($>10^\circ$) in elbow extension; 3) contralateral neck side-flexion increased
299 symptoms or ipsilateral side-flexion decreased symptoms (Wainner et al., 2005; Vanti et al.,
300 2011; Trillos et al., 2017). Sensitivity was moderate/high 0.75 (95%CI 0.58-0.92) (Wainner et
301 al., 2005) to high 0.91 (95%CI 0.74-0.98) (Vanti et al., 2011) and 0.93 (95%CI 0.88-0.96)
302 (Trillos et al., 2011). Specificity was low in all 3 studies: 0.13 (95%CI 0.04-0.22) (Wainner et
303 al., 2005), 0.15 (95%CI 0.05-0.36) (Vanti et al., 2011) and 0.06 (95%CI 0.0-0.33) (Trillos et al.,
304 2017). In the study by Vanti et al. (2011) the authors conducted a second analysis in which
305 "reproduction of patient's symptoms" changed to "reproduction of symptoms in the first,
306 second or third digit", but again only one of the three criteria was required for a positive
307 ULNT1. The second analysis revealed low to moderate sensitivity (0.54, 95%CI 0.35-0.72) and
308 moderate specificity (0.70, 95%CI 0.48-0.85). Overall, none of the interpretations of ULNT1
309 was capable of ruling in or ruling out a diagnosis of CTS because LRs were between 0.5 and
310 2.0.

311 Two studies examined the diagnostic accuracy of ULNT1 using a different interpretation for a
312 positive test. In these studies the test was considered positive if it was able to reproduce
313 patient's symptoms and these symptoms were altered with structural differentiation (Vanti
314 et al., 2012; Bueno-Gracia et al., 2016). Sensitivity ranged from low 0.05 (95%CI 0.02-0.19)
315 (Vanti et al., 2012) to low/moderate 0.58 (95%CI 0.45-0.71) (Bueno-Gracia et al., 2016).
316 Specificity ranged from moderate/high 0.84 (95%CI 0.72-0.96) (Bueno-Gracia et al., 2016) to
317 high 0.93 (95%CI 0.82-0.98) (Vanti et al., 2012). Bueno-Gracia and colleagues (2016)
318 suggested that the ULNT1 may be clinically useful to determine patients with CTS due to
319 high +LR (3.67). However the high number of false negatives results challenges this notion
320 (Table 4).

321

322

323

324

325

326

327

328

329

330
331
332
333

ACCEPTED MANUSCRIPT

Author (Year)	Test (Positive test criteria)	SN (95% CI)	SP (95% CI)	+LR (95% CI)	-LR (95% CI)	PPV (95% CI)	NPV (95% CI)
Bueno-Gracia et al., 2016	ULNT1 Criterion A -Patient's symptoms reproduced and changed with SD	0.58 (0.45-0.71)	0.84 (0.72-0.96)	3.67 (1.70-7.89)	0.50 (0.36-0.70)	0.85 (0.71-92)	0.43 (36-51)
	Criterion B -Reproduction of symptoms in the wrist and first three digits that changed with SD, regardless of the reproduction of patient's clinical symptoms	0.74 (0.61-0.83)	0.50 (0.35-0.65)	1.47 (1.03-2.10)	0.53 (0.31-0.90)	0.69 (61-75)	0.44 (32-45)
Trillos et al., 2017	ULNT1 -Any one of the following: (1) patient's symptoms reproduced; (2) side to side differences (>10°) in elbow extension on completion of all motion sequences; (3) change of symptoms with SD	0.93 (0.88-0.96)	0.06 (0.0-0.33)	1.00	1.05	0.87 (?)	0.12 (?)
Vanti et al., 2011	ULNT1 Criterion A -Any one of the following: (1) reproduction of patient's symptoms; (2) side to side differences (>10°) in elbow extension on completion of all motion sequences; (3) change of symptoms with SD	0.91 (0.74-0.98)	0.15 (0.05-0.36)	1.07 (0.38-3.08)	0.55 (0.19-1.59)	0.56 (?)	0.40 (?)
	Criterion B -Side to side differences (>10°) in elbow extension on completion of all motion sequences, but (1) and (3) positive only in presence of symptoms reproduction in the 1 st , 2 nd and 3 rd digit of the affected arm	0.54 (0.35-0.72)	0.70 (0.48-0.85)	1.8 (1.13-2.88)	0.65 (0.41-1.04)	0.68 (?)	0.44 (?)
Vanti et al., 2012	ULNT1 Criterion A -symptoms in fingers I,II or III	0.4 (0.26-0.56)	0.79 (0.66-0.88)	1.96 (1.27-3.01)	0.75 (0.49-1.16)	0.58 (0.39-0.75)	0.65 (0.52-0.76)
	Criterion B -A + symptoms increased with controlateral cervical side bending	0.28 (0.16-0.45)	0.82 (0.69-0.91)	1.6 (0.93-2.76)	0.86 (0.50-1.49)	0.55 (0.34-0.75)	0.59 (0.47-0.70)
	Criterion C -A + symptoms decreased with ipsilateral cervical side bending	0.05 (0.02-0.19)	0.93 (0.82-0.98)	0.85 (0.22-3.30)	1.01 (0.26-3.89)	0.4 (0.12-0.77)	0.56 (0.45-0.67)
Wainner et al., 2005	ULNT1 -Any one of the following: (1) patient's symptoms reproduced; (2) side to side differences (>10°) in elbow extension on completion of all motion sequences; (3) change of symptoms with SD	0.75 (0.58-0.92)	0.13 (0.04-0.22)	0.86 (0.67-1.0)	1.9 (0.72-5.1)	(?) (?)	(?) (?)
	ULNT2b -Any one of the following: (1) patient's symptoms reproduced; (2) side to side differences (>10°) in elbow extension on completion of all motion sequences; (3) change of symptoms with SD	0.64 (0.45-0.83)	0.30 (0.17-0.42)	0.91 (0.65-1.3)	1.2 (0.62-2.4)		

334 SD: structural differentiation, SN: sensitivity, SP: specificity, +LR: positive likelihood ratio, -LR: negative likelihood ratio, PPV: positive predictive value, NPV: negative
335 predictive value, CI: confidence intervals,?: Data not available, authors have been contacted but did not respond

336

337

Table 4. Diagnostic ULNTs accuracy data for CTS

338 *Cervical radiculopathy*

339 Three studies investigated the concordance of ULNT1 with a reference standard in patients
340 with suspected CR (Wainner et al., 2003; Apelby-Albrecht et al., 2013; Ghasemi et al., 2013).
341 The reference standard in two of these studies was NCS and needle electromyography
342 (Wainner et al., 2003; Ghasemi et al., 2013), whereas in the third study the authors used the
343 combination of patient history, clinical examination and MRI findings as the reference
344 standard (Apelby-Albrecht et al., 2013). In two of these studies ULNT1 showed moderate to
345 high (0.83, 95%CI 0.66-0.93) and high sensitivity (0.97, 95%CI 0.90-1.0) (Apelby-Albrecht et
346 al., 2013; Wainner et al., 2003) whereas in the third study the sensitivity was low 0.35 for
347 chronic CR and low/moderate 0.6 for acute CR (Ghasemi et al., 2013). Specificity ranged
348 from low 0.22 (95%CI 0.12-0.33) (Wainner et al., 2003) and 0.4 (Ghasemi et al., 2013) to
349 moderate/high 0.75 (95%CI 0.48-0.93) (Apelby-Albrecht et al., 2013). Moreover, in the study
350 of Wainner et al. (2003) the ULNT1 demonstrated negative likelihood ratio (LR) of 0.12,
351 meaning that a negative ULNT1 could rule out CR. This study had low ROB, but had concerns
352 regarding applicability related to the different interpretation of the index test from the
353 authors compared with the review question (Whiting et al., 2011). In addition, due to wide
354 95% CI the results of this study should be interpreted cautiously. Wide CIs reduce the
355 strength of evidence by influencing the precision of the pooled estimates.

356 The validity of ULNT2b (radial) was assessed by two studies (Wainner et al., 2003; Apelby-
357 Albrecht et al., 2013). Sensitivity was moderate in both studies: 0.66 (95%CI 0.48-0.81)
358 (Apelby-Albrecht et al., 2013) and 0.72 (95%CI 0.52-0.93) (Wainner et al., 2003). Specificity
359 ranged from low 0.33 (95%CI 0.21-0.45) (Wainner et al., 2003) to moderate/high 0.75 (95%CI
360 0.48-0.93) (Apelby-Albrecht et al., 2013).

361 Apelby-Albrecht and colleagues (2013) also examined the diagnostic accuracy of ULNT2a
362 (median), ULNT3 (ulnar) and ULNTs combined as a single test. This study was assessed as at
363 ROB due to the time lapse between the MRI and the neurodynamic testing (up to six
364 months) (Whiting et al., 2011); however, no concerns regarding applicability were identified.
365 Combined ULNTs showed high sensitivity (0.97, 95%CI 0.85-1.00) and moderate specificity
366 (0.69, 95%CI 0.41-0.89) whereas the ULNT3 (ulnar) was the most specific (0.87, 95%CI 0.62-
367 0.98) (Apelby-Albrecht et al., 2013) (Table 5).

368
369
370

371
372
373
374
375

ACCEPTED MANUSCRIPT

Author (Year)	Test (Positive test criteria)	SN (95% CI)	SP (95% CI)	+LR (95% CI)	-LR (95% CI)	PPV (95% CI)	NPV (95% CI)
Apelby-Albert et al., 2013	ULNT1 -(1) Reproducible neurogenic pain in neck and arm, (2) increased/decreased symptoms with SD, (3) difference in painful radiation between right and left sides	0.83 (0.66-0.93)	0.75 (0.48-0.93)	3.32	0.22	0.88 (0.72-0.97)	0.67 (0.41-0.87)
	ULNT2a -(1) Reproducible neurogenic pain in neck and arm, (2) increased/decreased symptoms with SD, (3) difference in painful radiation between right and left sides	0.66 (0.48-0.81)	0.75 (0.48-0.93)	2.64	0.45	0.85 (0.66-0.96)	0.50 (0.29-0.71)
	ULNT2b -(1) Reproducible neurogenic pain in neck and arm, (2) increased/decreased symptoms with SD, (3) difference in painful radiation between right and left sides	0.43 (0.26-0.61)	0.75 (0.48-0.93)	1.72	0.76	0.79 (0.54-0.94)	0.37 (0.21-0.56)
	ULNT3 -(1) Reproducible neurogenic pain in neck and arm, (2) increased/decreased symptoms with SD, (3) difference in painful radiation between right and left sides	0.71 (0.54-0.85)	0.87 (0.62-0.98)	5.68	0.32	0.93 (0.76-0.99)	0.58 (0.37-0.78)
	ULNTcomb. -(1) Reproducible neurogenic pain in neck and arm, (2) increased/decreased symptoms with SD, (3) difference in painful radiation between right and left sides	0.97 (0.85-1.00)	0.69 (0.41-0.89)	3.11	0.04	0.87 (0.73-0.96)	0.92 (0.62-1.00)
Ghasemi et al., 2013	ULNT1 -Reproduction of pain in any step						
	Acute CR Chronic CR	0.6 0.35	0.4 0.4	1.0 0.58	1.0 1.62	0.68 (?) 0.50 (?)	0.32 (?) 0.27 (?)
Wainner et al., 2003	ULNT1 -Any one of the following: (1) patient's symptoms reproduced; (2) side to side differences (>10°) in elbow extension on completion of all motion sequences; (3) change of symptoms with SD	0.97 (0.90-1.0)	0.22 (0.12-0.33)	1.3 (1.1-1.5)	0.12 (0.01-1.9)	(?)	(?)
	ULNT2b -Any one of the following: (1) patient's symptoms reproduced; (2) side to side differences (>10°) in elbow extension on completion of all motion sequences; (3) change of symptoms with SD	0.72 (0.52-0.93)	0.33 (0.21-0.45)	1.1 (0.77-01.5)	0.85 (0.37-1.9)	(?)	(?)

376 SD: structural differentiation, SN: sensitivity, SP: specificity, +LR: positive likelihood ratio, -LR: negative likelihood ratio, PPV: positive predictive value, NPV: negative
377 predictive value, CI: confidence intervals,?: Data not available, authors have been contacted but did not respond

378

379

Table 5. Diagnostic ULNTs accuracy data for CR

380

	No of studies (No of patients)	Accuracy measures	RoB	Indirectness	Inconsistency	Imprecision	Publication bias	Quality of evince
ULNT1	3 studies (n=230) (Apelby-Albert et al., 2013; Ghasemi et al., 2013; Wainner et al., 2003)	Sensitivity	Serious	Serious	Serious	Serious	Undetected	Very low
		Specificity	Serious	Serious	Very serious	Very serious	Undetected	Very low
ULNT2a	1 study (n=51) (Apelby-Albert et al., 2013)	Sensitivity	Serious	No	No	Very serious	Undetected	Very low
		Specificity	Serious	No	No	Very serious	Undetected	Very low
ULNT2b	2 studies (n=133) (Apelby-Albert et al., 2013; Wainner et al., 2003)	Sensitivity	Serious	Serious	Very serious	Very serious	Undetected	Very low
		Specificity	Serious	Serious	Very serious	Serious	Undetected	Very low
ULNT3	1 study (n=51) (Apelby-Albert et al., 2013)	Sensitivity	Serious	No	No	Very serious	Undetected	Very low
		Specificity	Serious	No	No	Very serious	Undetected	Very low
ULNT (combined)	1 study (n=51) (Apelby-Albert et al., 2013)	Sensitivity	Serious	No	No	Serious	Undetected	Low
		Specificity	Serious	No	No	Very serious	Undetected	Very low

381 ULNT: upper limb neurodynamic test, RoB: risk of bias, CR: cervical radiculopathy

382 **Table 6. GRADE assessment of evidence (CR)**

	No of studies (No of patients)	Accuracy measures	RoB	Indirectness	Inconsistency	Imprecision	Publication bias	Quality of evince
ULNT1	5 studies (n=349) (Wainner et al., 2005; Vanti et al., 2011, 2012; Bueno-Gracia et al., 2016; Trillos et al., 2017)	Sensitivity	Serious	Serious	Serious	Very serious	Undetected	Very low
		Specificity	Serious	Serious	Very serious	Very serious	Undetected	Very low
ULNT2b	1 study (n=82) (Wainner et al., 2005)	Sensitivity	No	Serious	Serious	Very serious	Undetected	Very low
		Specificity	No	Serious	Serious	Very serious	Undetected	Very low

383 ULNT: upper limb neurodynamic test, RoB: risk of bias, CTS: carpal tunnel syndrome

384

Table 7. GRADE assessment of evidence (CTS)

385

386

387

388

389

390

391

392 DISCUSSION

393

394 The purpose of this study was to evaluate the role of ULNTs in the assessment of PNP and to
395 reflect on their value in clinical practice in the assessment and diagnosis of patients with arm
396 and/or neck symptoms. Current research suggests that ULNTs cannot be used in isolation for
397 the diagnosis of PNP. Specifically, ULNTs cannot be utilised as a stand-alone test in the
398 clinical setting for the diagnosis of CTS. Limited evidence suggests that ULNTs demonstrate
399 better diagnostic accuracy and may be clinically relevant for the diagnosis of CR, but only as
400 a “ruling out” strategy. However, the overall body of the evidence after applying the GRADE
401 approach was low to very low for all outcomes, therefore any interpretation of these
402 findings should be made cautiously.

403

404 Carpal tunnel syndrome

405 Overall, the five studies that examined the validity of ULNT1 are characterised by diversity in
406 the interpretation of the index test. From these studies only the interpretation by Bueno-
407 Gracia et al. (2016) is in agreement with the review question, that is, the ULNT1 is
408 considered positive only when it reproduces the patient’s clinical symptoms and those
409 symptoms are modified with structural differentiation. This criterion is supported by several
410 authors, who suggest that structural differentiation is necessary in order to distinguish
411 between neuropathic pain and pain that arises from other somatic sources (Nee et al., 2012;
412 Butler, 2000; Coppieters et al., 2002). Using the above definition for a positive test Bueno-
413 Gracia et al. (2016) found that the ULNT1 may have strong ability to identify patients who do
414 not have CTS (high specificity).

415 Using a different definition of a positive test Wainner et al. (2005), Vanti et al. (2011) and
416 Trillos et al. (2017) found that the ULNT1 had moderate/high to high sensitivity. However,
417 the low specificities and LRs that have been obtained in these studies decrease the
418 diagnostic accuracy of ULNT1 and suggest that they cannot be considered adequate for the
419 diagnosis of CTS.

420

421 Cervical radiculopathy

422 The diagnostic accuracy of ULNTs seems more promising for the diagnosis of CR. Apelby-
423 Albrecht et al. (2013) investigated the validity of ULNTs combined and individually, using the
424 same definition for a positive test as this review. Individually, ULNT1 and ULNT3 were the
425 most valid tests for detecting CR. Combining the tests increased the diagnostic accuracy of
426 ULNTs further, giving an accurate diagnosis in 88.2% of patients.

427 Whilst findings by Wainnner et al. (2003) are in agreement with the study by Apelby-
428 Albrecht et al. (2013) the authors used a more liberal definition of a positive test. In their
429 study, the ULNT1 was highly sensitive and had LR- of 0.12 meaning that when the test is
430 negative, CR can be ruled out. In these studies the vast majority of patients with CR
431 presented with nerve root compression at C6-C7 level, therefore the diagnostic properties of
432 ULNTs may be different when the C5 or C8 root level is involved.

433 Overall, following analysis of the available evidence, ULNTs seem to have no diagnostic
434 accuracy to inform clinical practice in patients with suspected CTS. In contrast, ULNTs may
435 be more useful for the diagnosis of CR, but only as a “ruling out” strategy. Nonetheless,
436 these findings should be interpreted cautiously due to the small number of studies
437 investigating the diagnostic accuracy of ULNTs and the differences between them in regards
438 to the interpretation of a positive test.

439 There are a number of concerns that may explain some of the results obtained in these
440 studies. Firstly, electrodiagnostic testing provides information in regards to conduction loss
441 in large myelinated motor neurons and A β fibres (Schmid et al., 2013). Increased
442 mechanosensitivity, however, is related to increased excitability of small-diameter afferents
443 and sensitization of nociceptors in the nervi nervorum and sinuvertebral nerves (Baron et al.,
444 2010). Moreover, recent evidence suggests that damage of small axons is more common in
445 entrapment neuropathies than previously believed (Chien et al., 2008; Schmid et al., 2012)
446 and may occur even before any dysfunction of large axons (Tamburin et al., 2010). Thus, it
447 becomes apparent that the inability of the criterion standard to identify neuropathies
448 related to small axons damage may have led to false-negative results in cases where NCS
449 classified a patient as not having the condition whereas the ULNTs were positive.

450 Secondly, in a recent study Baselgia et al. (2017) found that >54% of patients with CTS had
451 negative ULNT1 despite a clear dysfunction in the median nerve, as proven with NCS. The
452 authors advocated that the non-reproduction of symptoms during neurodynamic testing can
453 be a sign of a more severe neural dysfunction of the unmyelinated fibres (Baselgia et al.,
454 2017). These findings, could explain some of the false-negative results that have been
455 obtained in the included studies in cases where the NCS confirmed a diagnosis but the
456 neurodynamic testing was negative.

457

458

459

460

461 Future direction

462 A reference standard should be comprehensive enough to accurately inform clinicians in
463 regards to the diagnostic accuracy of an index test. Given the insufficiency of
464 electrodiagnostic tests to provide information about the integrity of small-diameter nerve
465 fibres (Schmid et al., 2013), it becomes apparent that diagnostic accuracy studies need a
466 supplementary test that will increase the criterion validity of the reference standard.
467 Quantitative Sensory Testing (QST) provides information for both loss and gain of function,
468 in large myelinated (A β) and thinly myelinated (A δ) or unmyelinated fibres (C-fibres) (Rolke
469 et al., 2006). QST protocols include tests that investigate thermal, mechanical and pain
470 thresholds, and based on the results clinicians could be informed in regards to which type of
471 nerve fibres might be involved. Incorporating QST in protocols, may enhance their ability to
472 correctly classify patients with PNP. Additionally future diagnostic accuracy studies aiming to
473 investigate the validity of ULNTs in patients with CTS could adopt the principle of
474 “neurodynamic sequencing” and alter the order of joint movement. Various studies have
475 shown that the range of motion and the symptoms can be modified by altering the testing
476 sequence during straight leg raise (Boland and Adams, 2000), slump test (Johnson and
477 Chiarello, 1997) and ULNT1 (Coppieters et al., 2001). Moving the wrist to extension first
478 during ULNT1 testing may increase the likelihood of a positive neurodynamic test (Baselgia,
479 2017). Moreover, consensus as to what defines a positive test would be useful.
480 Standardisation of the performance and the interpretation of ULNTs are essential to draw
481 safe inferences for the true diagnostic accuracy of the tests (Nee et al., 2012). Finally, future
482 diagnostic accuracy studies should evaluate the diagnostic utility of ULNTs for ulnar nerve
483 EN, since currently there are limited evidence regarding to the validity of ULNTs in
484 pathologies such cubital syndrome.

485

486 Strengths and limitations

487 The strengths of this review are that provides clear recommendations for future studies and
488 emphasises the importance of precisely reported methodologically robust studies. Among
489 the limitations of this systematic review is that it includes studies only written in English
490 which may have introduced bias (Song et al., 2002). Whilst we have adopted the grading of
491 sensitivity and specificity using parameters based on existing reviews we acknowledge
492 interpretation is context specific; further research is required to validate these categories.

493

494

495

CONCLUSION

496 Based on the available evidence, ULNTs have no diagnostic accuracy to identify patients with
497 CTS when used in isolation. Limited evidence suggests that ULNTs demonstrate better
498 diagnostic accuracy and may be clinically relevant for the diagnosis of CR, but only in a
499 “ruling out” strategy. However, the overall quality of the body of evidence after applying the
500 GRADE approach was low to very low across studies. Further higher quality research is
501 needed to establish firm conclusions regarding to the value of ULNTs in the assessment and
502 diagnosis of patients with arm and/or neck symptoms.

503

Funding statement

505 This research did not receive any specific grant from funding agencies in the public,
506 commercial, or not-for-profit sectors.

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

REFERENCES

- 533
534
535 Apelby-Albrecht, M., Andersson, L., Kleiva, I.W., et al. (2013) Concordance of upper limb
536 neurodynamic tests with medical examination and magnetic resonance imaging in patients
537 with cervical radiculopathy: a diagnostic cohort study. **Journal of Manipulative and**
538 **Physiological Therapeutics**, 36 (9): 626-632.
- 539 Atroshi, I., Gummesson, C., Johnsson, R., Ornstein, E., Ranstam, J., & Rosén, I. (1999).
540 Prevalence of carpal tunnel syndrome in a general population. **Jama**, 282(2), 153-158.
- 541 Baron, R., Binder, A. and Wasner, G. (2010) Neuropathic pain: diagnosis, pathophysiological
542 mechanisms, and treatment. **The Lancet Neurology**, 9 (8): 807-819.
- 543 Baselgia, L.T., Bennett, D.L., Silbiger, R.M., et al. (2017) Negative Neurodynamic Tests Do Not
544 Exclude Neural Dysfunction in Patients With Entrapment Neuropathies. **Archives of Physical**
545 **Medicine and Rehabilitation**, 98 (3): 480-486.
- 546 Birchall, D., Connelly, D., Walker, L., et al. (2003) Evaluation of magnetic resonance
547 myelography in the investigation of cervical spondylotic radiculopathy. **The British Journal of**
548 **Radiology**, 76 (908): 525-531.
- 549 Boland, R.A. and Adams, R.D. (2000) Effects of ankle dorsiflexion on range and reliability of
550 straight leg raising. **The Australian Journal of Physiotherapy**, 46 (3): 191-200.
- 551 Burgess, R. M., Rushton, A., Wright, C., & Daborn, C. (2011). The validity and accuracy of
552 clinical diagnostic tests used to detect labral pathology of the hip: a systematic
553 review. **Manual Therapy**, 16(4), 318-326.
- 554 Butler D. **The neurodynamic techniques. A definitive guide from the Noigroup Team.**
555 Adelaide: Noigroup Publication; 2008.
- 556 Butler, D. S. (2000). **The sensitive nervous system.** Noigroup publications.
- 557 Chien, A., Eliav, E. and Sterling, M. (2008) Whiplash (grade II) and cervical radiculopathy
558 share a similar sensory presentation: an investigation using quantitative sensory testing. **The**
559 **Clinical Journal of Pain**, 24 (7): 595-603.
- 560 Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: the SPIDER tool for qualitative
561 evidence synthesis. *Qualitative health research*, 22(10), 1435-1443.
- 562 Coppieters, M., Stappaerts, K., Janssens, K., et al. (2002) Reliability of detecting 'onset of
563 pain' and 'submaximal pain' during neural provocation testing of the upper
564 quadrant. **Physiotherapy Research International**, 7 (3): 146-156.

- 565 Coppieters, M.W., Stappaerts, K.H., Everaert, D.G., et al. (2001) Addition of test components
566 during neurodynamic testing: effect on range of motion and sensory responses. **The Journal**
567 **of Orthopaedic and Sports Physical Therapy**, 31 (5): 226-237.
- 568 CRD. Systematic reviews: CRD's guidance for undertaking reviews in health care. Centre for
569 Reviews and Dissemination; 2009 [Chapter 2].
- 570 Deeks JJ, Wisniewski S, Davenport C. Chapter 4: Guide to the contents of a Cochrane
571 Diagnostic Test Accuracy Protocol. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane
572 Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The **Cochrane**
573 **Collaboration**, 2013. Available from: <http://srdta.cochrane.org/>.
- 574 Egger M, Smith G, Altman D. (2001). **Systematic Eviews in health care: meta analysis in**
575 **context**. Bodmin: BMJ Publishing Group; 248-281.
- 576 Elvey RL: **Brachial plexus tension tests and the pathoanatomical origin of arm pain**. In
577 Aspects of Manipulative Therapy Edited by: Idczak RM. Melbourne: Lincoln Institute of
578 Health Sciences; 1980:105-110.
- 579 Fernández-de-las-Peñas, C., Fernández-Muñoz, J. J., Palacios-Ceña, M., Navarro-Pardo, E.,
580 Ambite-Quesada, S., & Salom-Moreno, J. (2015). Direct and Indirect effects of function in
581 associated variables such as depression and severity on pain intensity in women with carpal
582 tunnel syndrome. **Pain Medicine**, 16(12), 2405-2411.
- 583 Finnerup, N.B., Haroutounian, S., Kamerman, P., et al. (2016) Neuropathic pain: an updated
584 grading system for research and clinical practice. **Pain**, 157 (8): 1599-1606.
- 585 Ghasemi, M., Golabchi, K., Mousavi, S.A., et al. (2013) The value of provocative tests in
586 diagnosis of cervical radiculopathy. **Journal of Research in Medical Sciences : The Official**
587 **Journal of Isfahan University of Medical Sciences**, 18 (Suppl 1): S35-8.
- 588 Grodahl, L.H., Fawcett, L., Nazareth, M., et al. (2016) Diagnostic utility of patient history and
589 physical examination data to detect spondylolysis and spondylolisthesis in athletes with low
590 back pain: A systematic review. **Manual Therapy**, 24 7-17.
- 591 Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [updated
592 September 2008]. **The Cochrane Collaboration**, 2008.
- 593 Hegedus, E.J., Goode, A.P., Cook, C.E., et al. (2012) Which physical examination tests provide
594 clinicians with the most value when examining the shoulder? Update of a systematic review
595 with meta-analysis of individual tests. **British Journal of Sports Medicine**, 46 (14): 964-978.

- 596 Jablecki, C.K., Andary, M.T., Floeter, M.K., et al. (2002) Practice parameter: Electrodiagnostic
597 studies in carpal tunnel syndrome. Report of the American Association of Electrodiagnostic
598 Medicine, American Academy of Neurology, and the American Academy of Physical
599 Medicine and Rehabilitation. **Neurology**, 58 (11): 1589-1592.
- 600 Jaeschke, R., Guyatt, G. H., Sackett, D. L., Guyatt, G., Bass, E., Brill-Edwards, P., ... & Haynes,
601 B. (1994). Users' Guides to the Medical Literature: III. How to Use an Article About a
602 Diagnostic Test B. What Are the Results and Will They Help Me in Caring for My
603 Patients?. **Jama**, 271(9), 703-707.
- 604 Jaeschke, R., Guyatt, G.H. and Sackett, D.L. (1994) Users' guides to the medical literature. III.
605 How to use an article about a diagnostic test. B. What are the results and will they help me
606 in caring for my patients? The Evidence-Based Medicine Working Group. **Jama**, 271 (9): 703-
607 707.
- 608 Johnson, E.K. and Chiarello, C.M. (1997) The slump test: the effects of head and lower
609 extremity position on knee extension. **The Journal of Orthopaedic and Sports Physical**
610 **Therapy**, 26 (6): 310-317.
- 611 Kuijper, B., Tans, J.T., Schimsheimer, R.J., et al. (2009) Degenerative cervical radiculopathy:
612 diagnosis and conservative treatment. A review. **European Journal of Neurology**, 16 (1): 15-
613 20.
- 614 Lijmer, J. G., Bossuyt, P. M., & Heisterkamp, S. H. (2002). Exploring sources of heterogeneity
615 in systematic reviews of diagnostic tests. **Statistics in Medicine**, 21(11), 1525-1537.
- 616 McInnes, M. D., Moher, D., Thombs, B. D., McGrath, T. A., Bossuyt, P. M., Clifford, T., ... &
617 Hunt, H. A. (2018). Preferred reporting items for a systematic review and meta-analysis of
618 diagnostic test accuracy studies: the PRISMA-DTA statement. **Jama**, 319(4), 388-396.
- 619 Nee, R.J. and Butler, D. (2006) Management of peripheral neuropathic pain: Integrating
620 neurobiology, neurodynamics, and clinical evidence. **Physical Therapy in Sport**, 7 (1): 36-49.
- 621 Nee, R.J., Jull, G.A., Vicenzino, B., et al. (2012) The validity of upper-limb neurodynamic tests
622 for detecting peripheral neuropathic pain. **The Journal of Orthopaedic and Sports Physical**
623 **Therapy**, 42 (5): 413-424.
- 624 Radhakrishnan, K., Litchy, W.J., O'Fallon, W.M., et al. (1994) Epidemiology of cervical
625 radiculopathy. A population-based study from Rochester, Minnesota, 1976 through
626 1990. **Brain : a journal of Neurology**, 117 (Pt 2) (Pt 2): 325-335.

- 627 Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9:
628 Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), **Cochrane**
629 **Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0**. The Cochrane
630 Collaboration, 2009. Available from: <http://srdta.cochrane.org/>
- 631 Rolke, R., Baron, R., Maier, C., et al. (2006) Quantitative sensory testing in the German
632 Research Network on Neuropathic Pain (DFNS): standardized protocol and reference
633 values. **Pain**, 123 (3): 231-243.
- 634 Rubinstein, S. M., Pool, J. J., van Tulder, M. W., Riphagen, I. I., & de Vet, H. C. (2007). A
635 systematic review of the diagnostic accuracy of provocative tests of the neck for diagnosing
636 cervical radiculopathy. **European spine journal**, 16(3), 307-319.
- 637 Saint-Lary, O., Rébois, A., Mediouni, Z., & Descatha, A. (2015). Carpal tunnel syndrome:
638 primary care and occupational factors. **Frontiers in Medicine**, 2, 28.
- 639 Schmid, A.B., Brunner, F., Luomajoki, H., et al. (2009) Reliability of clinical tests to evaluate
640 nerve function and mechanosensitivity of the upper limb peripheral nervous system. **BMC**
641 **Musculoskeletal Disorders**, 10 11-2474-10-11.
- 642 Schmid, A.B., Nee, R.J. and Coppieters, M.W. (2013) Reappraising entrapment neuropathies-
643 -mechanisms, diagnosis and management. **Manual Therapy**, 18 (6): 449-457.
- 644 Schmid, A.B., Soon, B.T., Wasner, G., et al. (2012) Can widespread hypersensitivity in carpal
645 tunnel syndrome be substantiated if neck and arm pain are absent? **European Journal of**
646 **Pain**, 16 (2): 217-228.
- 647 Schneiders, A.G., Sullivan, S.J., Hendrick, P.A., et al. (2012) The ability of clinical tests to
648 diagnose stress fractures: a systematic review and meta-analysis. **The Journal of**
649 **Orthopaedic and Sports Physical Therapy**, 42 (9): 760-771.
- 650 Shacklock, M. Neurodynamics. **Physiotherapy**, 81 (1): 9-16.
- 651 Shacklock, M.O. (2005) **Clinical neurodynamics: a new system of musculoskeletal**
652 **treatment / Michael Shacklock**. Elsevier Butterworth-Heinemann, Edinburgh ; London.
- 653 Song F, Khan K, Dinnes J, Sutton A., (2002) Asymmetrical funnel plots and publication bias in
654 meta-analyses of diagnostic accuracy. **International Journal of Epidemiology**. (31):88e95.
- 655 Tamburin, S., Cacciatori, C., Praitano, M.L., et al. (2011) Median nerve small- and large-fiber
656 damage in carpal tunnel syndrome: a quantitative sensory testing study. **The Journal of**
657 **Pain**, 12 (2): 205-212.

- 658 Thiese, M.S., Gerr, F., Hegmann, K.T., et al. (2014) Effects of varying case definition on carpal
659 tunnel syndrome prevalence estimates in a pooled cohort. **Archives of Physical Medicine
660 and Rehabilitation**, 95 (12): 2320-2326.
- 661 Thoomes, E. J., van Geest, S., van der Windt, D. A., Falla, D., Verhagen, A. P., Koes, B. W., ... &
662 Vleggeert-Lankamp, C. L. (2017). Value of physical tests in diagnosing cervical radiculopathy:
663 a systematic review. **The Spine Journal**.
- 664 Toussaint, C.P., Perry, E.C.,3rd, Pisansky, M.T., et al. (2010) What's new in the diagnosis and
665 treatment of peripheral nerve entrapment neuropathies. **Neurologic clinics**, 28 (4): 979-
666 1004.
- 667 Wainner, R.S., Fritz, J.M., Irrgang, J.J., et al. (2003) Reliability and diagnostic accuracy of the
668 clinical examination and patient self-report measures for cervical radiculopathy. **Spine**, 28
669 (1): 52-62.
- 670 Wainner, R.S., Fritz, J.M., Irrgang, J.J., et al. (2005) Development of a clinical prediction rule
671 for the diagnosis of carpal tunnel syndrome. **Archives of Physical Medicine and
672 Rehabilitation**, 86 (4): 609-618.
- 673 Whiting, P.F., Rutjes, A.W., Westwood, M.E., et al. (2011) QUADAS-2: a revised tool for the
674 quality assessment of diagnostic accuracy studies. **Annals of Internal Medicine**, 155 (8): 529-
675 536.
- 676 Woolf, C.J. (2004) Dissecting out mechanisms responsible for peripheral neuropathic pain:
677 implications for diagnosis and therapy. **Life Sciences**, 74 (21): 2605-2610.

678
679
680
681



PRISMA-DTA Checklist

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE / ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	1
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	2-4
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	2-4
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	N/A
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4-5
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	5
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7-8
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6-7
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	6-7
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	6
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	7
Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	14



PRISMA-DTA Checklist

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	N/A
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	7-8
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	8-12
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	13-14
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	?
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	14-18
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	N/A
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	21-22
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	23
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	24
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	24

Adapted From: McInnes MDF, Moher D, Thoms BD, McGrath TA, Bossuyt PM, The PRISMA-DTA Group (2018). Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163.

For more information, visit: www.prisma-statement.org.

Highlights

- Diagnostic accuracy of ULNT in carpal tunnel syndrome is limited
- Evidence supports ULNTs in cervical radiculopathy only as a “ruling out” strategy
- NCS may not be adequate to determine diagnostic accuracy of ULNTs
- Integrating QST with ULNT may enhance classification of patients with PNP