



**HAL**  
open science

## Diagnostic accuracy of upper limb neurodynamic tests in the diagnosis of cervical radiculopathy

Francis Grondin, Chad Cook, Toby Hall, Olivier Maillard, Yannick Perdrix,  
Sebastien Freppel

► **To cite this version:**

Francis Grondin, Chad Cook, Toby Hall, Olivier Maillard, Yannick Perdrix, et al.. Diagnostic accuracy of upper limb neurodynamic tests in the diagnosis of cervical radiculopathy. *Musculoskeletal Science and Practice*, 2021, 55, pp.102427. 10.1016/j.msksp.2021.102427 . hal-03375066

**HAL Id: hal-03375066**

<https://hal.univ-reunion.fr/hal-03375066v1>

Submitted on 2 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## **Diagnostic accuracy of upper limb neurodynamic tests in the diagnosis of cervical radiculopathy**

Francis Grondin<sup>1,2</sup>, PhD-candidate, PT, Chad Cook<sup>3</sup>, PhD, PT, FAPTA, Toby Hall<sup>4</sup>, PhD, PT, Olivier Maillard<sup>5</sup>, MSc, PharmD, Yannick Perdrix<sup>6</sup>, MSc, PT, <sup>1</sup>Sebastien Freppel<sup>1</sup>, PhD, MD.

<sup>1</sup>Neurosurgery Department, Centre Hospitalier Universitaire de La Réunion, Reunion, France.

<sup>2</sup>School of Physiotherapy (IFMK), Centre Hospitalier Universitaire de La Réunion, Reunion, France.

<sup>3</sup>Professor. Duke Department of Orthopaedics, Duke Clinical Research Institute, Duke University, 311 Trent Drive, Durham, NC, USA. [chad.cook@duke.edu](mailto:chad.cook@duke.edu)

<sup>4</sup>Adjunct Associate Professor, School of Physiotherapy and Exercise Science, Curtin University, Kent Street, Bentley, Perth, Australia. [halltm@netspace.net.au](mailto:halltm@netspace.net.au)

<sup>5</sup>Centre Hospitalier Universitaire de La Réunion, INSERM, CIC1410, 97410, Saint Pierre, France. [olivier.maillard@chu-reunion.fr](mailto:olivier.maillard@chu-reunion.fr)

<sup>6</sup>Educational Manager, School of Physiotherapy (IFMK), Centre Hospitalier Universitaire de la Réunion, Reunion, France. [yannick.perdrix@ies-reunion.fr](mailto:yannick.perdrix@ies-reunion.fr)

### **Declaration of Conflict of Interests**

The authors declare no conflict of interest for this article.

### **Corresponding Author**

[grondin.fran@yahoo.fr](mailto:grondin.fran@yahoo.fr)

## 1 **Introduction**

2           Cervical radiculopathy (CR) is a relatively common neurological disorder caused by  
3 mechanical compression from a disc or other space-occupying lesion or from inflammation to  
4 the nerve root (Anekstein et al, 2012) The annual incidence of CR is 107.3 per 100,000 for  
5 men and 63.5 per 100,000 for women (Radhakrishnan et al, 1990). The clinical manifestations  
6 of cervical radiculopathy may include pain, sensory deficits, motor deficits, diminished  
7 reflexes, or combinations of these. Cervical radiculopathy typically is self-limiting with 75%–  
8 90% of patients achieving symptomatic improvement or resolution within a year with  
9 conservative care (Woods et al, 2015)

10           Although no gold standard exists as a reference standard for cervical radiculopathy,  
11 magnetic resonance imaging (MRI) is the preferred diagnostic method (Mink et al, 2003),  
12 since it can differentiate tumors, inflammation, visualize trauma, and the extensiveness of  
13 disc, arthritic, neural, and vascular cervical pathologies. Electrodiagnostic tests are capable of  
14 detecting clinically significant problems in many patients as well, although they are operator  
15 dependent and variable methods and normative values are used in practice (Reza Soltani et al,  
16 2014). Furthermore it may be negative if performed before denervation has occurred or when  
17 re-innervation is complete (Ashkan et al, 2002). Cervical radiculopathy is considered a  
18 ‘clinical diagnosis with imaging confirmation’, and it is important to match valid clinical  
19 signs with MRI findings and/or electrodiagnostic test results (Carette and Fehlings, 2005 ;  
20 Kuijper et al, 2009).

21           There are numerous clinical tests used to diagnose cervical radiculopathy. Upper Limb  
22 neurodynamic tests ((ULNT) 1, 2a, 2b and 3), or also called upper limb tension test (ULTT),  
23 initially described by Elvey (Elvey, 1986), Butler (Butler, 2000) and Shacklock (Shacklock  
24 1996), involve targeted sequences of movement that provoke mechanosensitivity of the nerve.  
25 The tests are performed by placing and releasing progressively more tension on the proposed

26 component of the nervous system that is being tested. A recent systematic review (Thoomes  
27 et al, 2017) concluded that “limited evidence for accuracy of physical examination tests for  
28 the diagnosis of CR” exists.

29         Moreover, neurodynamic test procedures in studies that populated the aforementioned  
30 systematic review used a variety of testing methods and results to determine a positive  
31 finding. Three criteria has been advocated when testing: 1) reproduction of neurogenic pain-  
32 burning or lightning-like pain, tingling sensation in the neck and arm (Apelby-Albrecht et al,  
33 2012), the patient’s symptoms reproduced (Wainner et al, 2003), or reproduction of pain  
34 (Ghasemi et al, 2013); 2) side to side range of motion difference (Wainner et al, 2003) or  
35 side-to-side difference in painful radiation (Apelby-Albrecht et al, 2012); and 3)  
36 increased/decreased symptoms with structural differentiation (Apelby-Albrecht et al, 2012) or  
37 cervical structural differentiation with cervical spine movement alone (Wainner et al, 2003).  
38 Interestingly, there is inconsistency in what is advocated to measure including conflicting  
39 evidence for reproduction of any pain or discomfort (Apelby-Albrecht et al, 2012 ; Ghasemi  
40 et al, 2013); and side to side range of motion comparisons (Nee et al, 2012). Most studies  
41 advocate the use of structural differentiation, which involves directional movement of a  
42 defined body region (e.g., neck side flexion) away from the area assessed to evaluate the  
43 effect of mechanical force on the nervous system and its impact on the patient’s symptoms.

44         In clinical practice, many clinicians have assessed the accuracy of neurodynamic tests  
45 on two criteria as recommended by Nee et al (Nee et al, 2012) : 1) familiar patient’s  
46 pain/symptoms reproduced and 2) increased/decreased symptoms with structural  
47 differentiation. Since we are unfamiliar with any studies that have included both findings in  
48 the assessment of CR, we investigated the accuracy of four ULNTs in comparison against a  
49 reference standard of medical history and MRI confirmation in patients with and without CR.  
50 We hypothesized that the findings may provide insight on the role of ULNTs (e.g., screening

51 or confirmation) and that the more rigid definition of a positive test should improve the  
52 specificity of the test findings. Further, combinations of test findings should result in more  
53 diagnostic accuracy than individual tests alone.

54

## 55 **Materials and Methods**

56 The study was a diagnostic accuracy study (prospective) design in which clinical  
57 testing occurred in a state of diagnostic uncertainty. The study followed the updated 2015  
58 STARD reporting standards (Bossuyt et al, 2015). Patients were informed about the study and  
59 they gave their consent for participation before inclusion. The study was conducted in  
60 accordance with the Ethical principles and the Helsinki Declaration on research involving  
61 human subjects and was approved the French regulatory and ethics rules (n°2212189v0).

62 Participants were recruited from consecutive patients referred to a Neurosurgery  
63 Department by a general practitioner or specialist from September 2017 to September 2019.  
64 Each patient had a suspected neck disorder. Referred patients provided information and  
65 questionnaires about pain intensity and neck disability. To be included patients had to be aged  
66 18 to 65 years, reporting arm pain with or without neck pain of at least 3-months in duration.  
67 In addition, they were required to have a self-reported pain score of at least 30mm and less  
68 than 80 on a 100 mm visual analogue scale (VAS) (Horn et al, 2016) during the previous 24  
69 hours, and had a self-reported score of at least 20% on the Neck Disability Index  
70 questionnaire (NDI) (Masaracchio et al, 2013).

71 Subjects were excluded if they were unable to understand French, had suffered from a  
72 significant neck trauma at the time of the study (i.e., recent cranio-cervical trauma including  
73 cervical spine fracture), had a history of neck or arm surgery, inflammatory joint  
74 condition/arthritis, fibromyalgia, diabetes, pregnancy, cardiovascular, neurological, neoplastic  
75 or psychiatric pathology, cervical myelopathy, pyramidal or extrapyramidal pathology.

76

77           *Reference Standard:* The diagnosis of CR or a competing diagnosis was made by a  
78 single neurosurgeon with 15 years of experience from the consecutive patients included.  
79 Cervical radiculopathy is a clinical diagnosis that is confirmed through imaging verification,  
80 thus the diagnosis of CR was based on the following criteria: 1) history and presence of  
81 dermatomal radicular pain and/or symptoms (dysesthesia, muscle weakness or altered  
82 reflexes) attributable to a CR and 2) presence of MRI findings. MRI findings were specific  
83 in their confirmation of nerve root compression or irritation by disc herniation or stenosis in  
84 pre- or intra-foraminal space narrowing on the ipsilateral side and at the same or adjacent  
85 level of radicular pain (Kuijper et al, 2008). The reference standard results were interpreted  
86 without knowledge of the results of ULNT.

87           Index tests : Approximately one hour after the reference standard was provided by the  
88 neurosurgeon, a single physiotherapist with 10 years of experience in neck pain management,  
89 and advanced certification for orthopedic assessment evaluated the ULNT on each participant.  
90 No intervention was allowed between the index test(s) and reference standard. The  
91 physiotherapist was blind to the patient history, clinical/MRI findings and the diagnosis. The  
92 index test results were interpreted without knowledge of the results of the reference standard  
93 and the presence of CR. Before the tests, patients were instructed to communicate the onset  
94 of any sensation such as stretch, tingling or pain anywhere in the arm or neck (Schmid et al,  
95 2009). The patient was positioned supine without a pillow (Walsh 2005). The examiner  
96 performed the ULNTs for that are purported for the median (ULNT1 and ULNT2a), radial  
97 (ULNT2b) and ulnar (ULNT3) nerves (Figure 1) in randomized order using randomization  
98 software. Upper limb neurodynamic testing was operated according to the standardized  
99 sequence previously described (Butler, 2000; Nee et al, 2012 ; Schmid et al, 2009), with a 5-  
100 minute break between each test to avoid any pain sensitization by repeating tests (Walsh

101 2005). Passive movements were achieved to the end of range or until symptoms were  
102 produced (Schmid et al, 2009). The non-symptomatic side was tested first for each ULNT for  
103 familiarization with sensation/pain induced by tests. A ULNT was considered as positive if  
104 both of the two following criteria were met:

- 105 - Reproduction of a familiar symptomatic complaint of arm pain and/or neck pain at  
106 least partially (pain or dysesthesia including burning, or lightning-like pain, or  
107 tingling sensation) (Nee et al, 2012);
- 108 - Structural differentiation: Once such a familiar complaint was provoked, structural  
109 differentiation between neurogenic and non-neurogenic sources was performed by  
110 the addition of sensitizing movements at a site distant to the pain: ipsilateral- or  
111 contralateral cervical lateral flexion, elbow or wrist extension/flexion, or shoulder  
112 girdle elevation (Appendix A) (Nee et al, 2012);

113 Tests were considered negative if each failed to meet the positive criteria identified above or  
114 indeterminate if the patient was unable to tolerate the test of position to allow complete  
115 execution of the test.

116

117

118

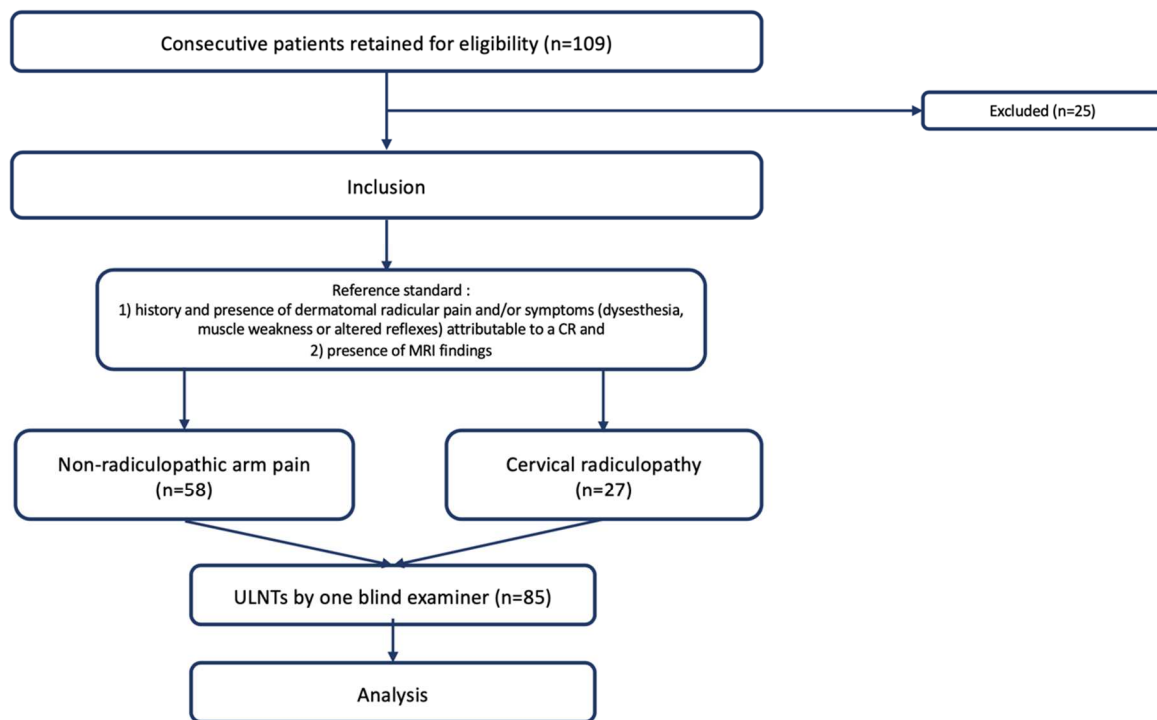
119

120

121

122

123 **Figure 1. Flow chart**



124

125

## 126 Results

127

128 Statistical analysis were carried using SPSS (IBM SPSS, Version 26.0. IBM Corp.

129 Armonk, NY). Variables normality was tested with Shapiro-Wilk test for continuous data.

130 Nonparametric continuous variables were described as medians and interquartile ranges and

131 differences were tested using the Wilcoxon Mann Whitney test. Gaussian variables were

132 described with means and standard deviations and differences were tested using the Student's

133 t-test. Categorical variables were described as numbers and percentages. They were compared

134 using Chi square test or Fisher's exact test, as appropriate. P-value significance level was set

135 at .05 and all tests were bilateral. Two by two tables were created for each ULNT measure.

136 ULNT test performances were analyzed using calculations for sensitivity, specificity, positive

137 likelihood ratios (LR+), and negative likelihood ratios (LR-). Sensitivity is the percentage of

138 people whose test is positive for a specific disease among a group of people who have the

139 disease (Cook et al, 2020). Specificity is the percentage of people whose test is negative for a



140 specific disease among a group of people who do not have the disease [22]. LR+ are the  
141 probability of a person with the disease testing positive divided by the probability of a person  
142 without the disease testing positive (Cook et al, 2020). LR- are the probability of a patient  
143 who has the condition of testing negative divided by the probability of a patient without the  
144 disease, testing negative (Cook et al, 2020).

145 We also calculated pretest probability, which is the probability of the condition being  
146 present before the diagnostic result is known and is sample specific for those enrolled in our  
147 study, and post-test probability with a positive and a negative finding on the ULNTs. Post-test  
148 probability is the percentage chance of the condition being present after a positive or negative  
149 finding for a ULNT. Generally, a positive test will increase the post-test probability of  
150 diagnosing the condition (otherwise known as ruling in the diagnosis). In contrast, a negative  
151 finding will generally decrease the post-test probability of diagnosing the condition  
152 (otherwise known as ruling out the condition) (Cook et al, 2020).

153 We calculated sensitivity/specificity, LR+/LR-, and post-test probabilities with a  
154 positive and a negative finding for the four individual ULNT tests and combinations of these  
155 tests. When calculating combinations of findings, the clusters of tests were placed in  
156 “conditions” (e.g., 1 of 4 is positive, 2 of 4 is positive, etc.) and evaluated for their abilities to  
157 influence post-test probability change with each defined condition. For all analyses, we also  
158 evaluated post-test probability change, which is the difference between the pre-test prevalence  
159 and the post-test finding with a positive or a negative result (Cook et al, 2020). Since the  
160 purpose of a test is to change the post-test probability of an accurate diagnosis, larger post-test  
161 probability changes were considered to have the highest clinical utility. 95% confidence  
162 intervals (95%CI) were calculated for all of these features.

163

164 Between September 2016 to December 2018, 85 participants, from 109 individuals

165 who were screened, were enrolled in the study. Of the 85 participants, 27 (31.7%) were  
 166 diagnosed with CR, 42 with neck and non-radiculopathic arm pain, 12 with peripheral nerve  
 167 entrapment, and 4 with diffuse shoulder pain (Table 1). All participants received the same  
 168 reference standard and were included for analysis (Figure 1). Diagnostic accuracy of the  
 169 four individual ULNTs are presented in Table 2. All four of the tests were more specific,  
 170 than sensitive, with the ULNT3 demonstrating the highest specificity. None of the four tests  
 171 markedly influenced post-test probability with a positive or a negative finding, with post-test  
 172 probability changes from baseline prevalence ranging from 41.58% with a positive for  
 173 ULNT3 to 15.72% with a negative for ULNT 2a.

174

175 **Table 1:** Baseline characteristics of the subjects (n=85)

	Cervical radiculopathy	Non-radiculopathic arm pain	
Age*	43.96 (8.94)	45.27 (9.74)	p = 0.61
Height (m)	1.66 (0.09)	1.67 (0.08)	p = 0.54
Body Mass Index *	24.73 (3.91)	24.88 (4.97)	p = 0.74
Duration (self-report) in months *	93.25 (98.41)	70.51 (62.31)	p = 0.69
Visual Analogue Scale for Pain*	5.14 (1.58)	5.03 (1.53)	p = 0.73
Neck Disability Index (%)	38.16 (14.14)	43.07 (13.90)	p= 0.19

176

177

178

179 **Table 2:** Diagnostic Accuracy of Individual Upper Limb Neurodynamic Tests. Pre-test  
 180 Prevalence = 31.7%.

Sensitivity	Specificity	LR+ (95% CI)	LR- (95% CI)	Post-test Probability	Post-test Probability with
-------------	-------------	-----------------	-----------------	--------------------------	-------------------------------

					with a Positive Finding (95% CI)	a Negative Finding (95% CI)
<b>ULNT 1</b>	59.26 (38.80, 77.61)	75.86 (62.83, 86.13)	2.46 (1.41-4.27)	0.54 (0.33-0.87)	53.30 (39.55-66.46)	20.04 (13.28-28.76)
<b>ULNT 2a</b>	70.37 (49.82, 86.25)	72.41 (59.10, 83.34)	2.55 (1.57-4.14)	0.41 (0.22-0.75)	54.20 (42.15-65.77)	15.98 (9.26-25.82)
<b>ULNT 2b</b>	55.56 (35.33, 74.52)	75.86 (62.83, 86.13)	2.30 (1.30-4.06)	0.59 (0.38-0.92)	51.63 (37.63-65.33)	21.49 (14.99-29.92)
<b>ULNT 3</b>	40.74 (22.39, 61.20)	93.10 (83.27, 98.09)	5.91 (2.07, 16.87)	0.64 (0.46-0.88)	73.28 (48.99-88.65)	22.95 (17.59-28.99)

181  
182  
183 Diagnostic accuracy of test conditions for combinations of ULNTs are presented in  
184 Table 3. Characteristically, with lower conditions (e.g., 1 of 4 is positive) values exhibit  
185 high sensitivity and low specificity, whereas higher conditions (4 of 4 are positive) values  
186 exhibit low sensitivity and high specificity. As expected, the condition of 1 out of 4 ULNT  
187 tests positive was the most sensitive combination whereas the condition of 4 out of 4 ULNT  
188 tests was the most specific. The condition of 1 out of 4 tests positive has the ability to “rule  
189 out” CR (LR-=0.08), exhibiting a post-test probability change of 28.12% with a negative  
190 finding. The condition of 4 of 4 tests positive had an infinite LR+ but there were only 3  
191 cases in which all four tests were positive. The condition of 3 of 4 tests positive occurred in  
192 12 of the 27 patients with CR and provided a LR+ of 12.89 and a post-test probability of  
193 85.71 (post-test probability change of 54.01%). No adverse events from performing the  
194 index test or the reference standard were observed.

195  
196  
197  
198  
199  
200 **Table 3.** Diagnostic Accuracy of Clustered Upper Limb Neurodynamic Test findings  
201 (Conditions). Pre-test Prevalence = 31.7%.  
202

Sensitivity	Specificity	LR + (95%	LR- (95% CI)	Post-test Probability	Post-test Probability with
-------------	-------------	-----------	--------------	-----------------------	----------------------------

	CI				with a Positive Finding (95% CI)	a Negative Finding (95% CI)
<b>1 of 4 Positive</b>	96.30 (81.03, 99.91)	46.55 (33.34, 60.13)	1.80 (1.40, 2.32)	0.08 (0.01, 0.56)	45.51 (39.38-51.84)	3.58 (0.46-20.62)
<b>2 of 4 Positive</b>	85.19 (66.27, 95.81)	74.14 (60.96, 84.74)	3.29 (2.07, 5.23)	0.20 (0.08, 0.50)	60.42 (48.99-70.82)	8.49 (3.59-18.83)
<b>3 of 4 Positive</b>	44.44 (25.48-64.67)	96.55 (88.09-99.58)	12.89 (3.10-53.62)	0.58 (0.41, 0.81)	85.71 (59.06-96.14)	23.237 (18.57-28.82)
<b>4 of 4 Positive</b>	11.11 (2.35, 29.16)	100.00 (93.84, 100.00)	Inf.	0.89 (0.78, 1.02)	100	29.23 (26.58-32.13)

203

204

205 **Discussion**

206

207

208

This study sought to determine the diagnostic accuracy of four ULNTs in identifying

209

CR in comparison with a reference standard of clinical diagnosis with MRI confirmation.

210

The study was performed in a situation of diagnostic uncertainty and used a more rigid

211

definition of what constitutes a positive test compared to previous studies [12-14]; the tests

212

also more closely matched how the tests are used in clinical practice. Findings were that

213

ULNTs when used in isolation did not lead to acceptable LR-, LR+ or post-test probability.

214

However, 3 out of 4 tests positive can rule in CR with a LR+ of 12.89. One of four positive

215

tests provided a LR- of 0.08 indicating that CR can be ruled out if no tests are positive. Of

216

the four tests, the ULNT3 influenced post-test probability the most with a positive test

217

(73.28%), whereas the ULNT2a influenced post-test probability the most with a negative

218

test (15.98%).

219

220

Each ULNT provided stronger LR+ values than LR-, thus influencing post-test probability

221

with a positive finding more notability than a negative finding. Our findings are markedly

222

different than those from Wainner et al. who found very low values of LR+ (<1.3) -

223

suggesting that they did not rule in - and moderately low LR- values (>0.12) - suggesting

224 they are better for ruling out (Wainner et al, 2003). Ghasemi and colleagues [14] failed to  
225 report a LR+ (or a LR-) (Ghasemi et al, 2013) and our calculations from their sensitivity and  
226 specificity values yielded LR+ values similar or worse than those of Wainner and associates  
227 (Wainner et al, 2003). The differences in findings compared to those of others (Wainner et  
228 al, 2003 ; Ghasemi et al, 2013) are likely related to the way we defined a positive index test  
229 (familiar compliant that was altered by structural differentiation). In Wainner and  
230 colleagues' study, an ULNT was defined as positive if only one of the following criteria  
231 were present: reproduction of the patient's symptoms, or side to side range of motion deficit,  
232 or structural differentiation using the cervical spine (Wainner et al, 2003). Ghasemi et al.,  
233 reported a positive finding if 'pain' occurred during testing (hasemi et al, 2013). Basing the  
234 test outcome on one criterion alone as identified by those authors could lead to an increase  
235 in false positive findings, thus decreasing specificity, and worsening the LR+ value  
236 (Schiffman et al, 2014). Apelby-Albrecht et al. defined as positive if all the three following  
237 criteria were met: reproduction of neurogenic symptoms according to a dermatomal pattern,  
238 increased or decreased symptoms with structural differentiation, and a difference in painful  
239 radiation between sides (Apelby-Albrecht et al, 2013). Our LR+ values are very similar to  
240 previous findings by a recent systematic review (Thoomes et al, 2017) calculated from data  
241 reported by Apelby-Albrecht et al (Apelby-Albrecht et al, 2013). However, in our study  
242 ULNT was defined as positive according to two criteria and we include a more mixed  
243 control group population (58 neck or shoulder pain, thoracic outlet syndrome and carpal  
244 tunnel syndrome) than Apelby-Albretch (only 18 subjects with neck pain or carpal tunnel  
245 syndrome) (Apelby-Albrecht et al, 2013). These findings highlight the importance to  
246 clinicians of determining a positive ULNT based on symptom reproduction together with the  
247 effects of structural differentiation, at least in diagnosing cervical radiculopathy.

248

249 In their recent systematic review of diagnostic tests for CR, Koulidis et al [25] concluded  
250 that ULNTs could only be used as a “ruling out” strategy (Koulidis et al, 2019) based on  
251 Apelby-Albrecht et al’s data (Apelby-Albrecht et al, 2013). Conversely, in our sample,  
252 ULNT when used in isolation were better at ruling in CR versus ruling out, yet clustering  
253 the ULNT findings produced large changes in post-test probability with either a negative  
254 finding or a positive finding. The condition of one of four positive tests yields a LR- of 0.08  
255 (95%CI=0.01-0.56). This means that when none of the four ULNTs are positive it can rule  
256 out CR with only a 3.58% chance that the patients in this sample had CR. Moreover, using  
257 multiple combination of ULNT demonstrated that the condition of 3 of 4 positive tests  
258 yielded a LR+ of 12.89 (95%CI=3.10-53.62) which means it can rule in CR with a post-test  
259 probability of 85.71%. We recommend the use of 3 of 4 conditions over 4 of 4, since this  
260 finding was uncommon and because the confidence intervals crossed 1.0 for the LR-  
261 analyses.

262 We are also the first to report post-test probability of a positive and negative finding with an  
263 ULNT, an analysis omitted from past works (Apelby-Albrecht et al, 2013; Ghasemi et al,  
264 2013). Post-test probability provides a better understanding of how markedly one’s decision  
265 is influenced by single, or combined, positive or negative test results. This is of particular  
266 importance since the reporting of individual sensitivity and specificity values is not  
267 recommended (Hegedus and Stern, 2009 ; Baeyens et al, 2019) and may yield conflicting  
268 results for ruling in or ruling out conditions.

269

#### 270 *Study limitations*

271 Although there is notable debate on an appropriate sample size for a diagnostic  
272 accuracy study (Hajian-Tilaki, 2014 ; Bujang and Adnan, 2016), we feel compelled to

273 identify our sample of 85 (including 27 CR) as a potential limitation. A smaller sample size  
274 may lead to less precision (e.g., wide confidence intervals). Only one clinician was involved  
275 in determining the reference standard and another was involved in determining the ULNT.  
276 Although the ULNT tester was blinded to the diagnosis of the patient, the transferability of  
277 their findings is unknown, since we did not test interrater agreement. Future research is  
278 needed to assessed the validity of ULNT with a larger sample of patients with CR and a  
279 larger control group with similar symptoms (thoracic outlet syndrome, neck/shoulder pain,  
280 peripheral nerve entrapment, *etc.*), and with more examiner and reference standards  
281 including magnetic resonance neurography and small fiber function (Schmid et al, 2013).

282

283

## 284 **Conclusions**

285 Our results support past findings that the singular use of ULNT to rule in or rule out  
286 CR is not recommended. When combinations are used, findings have higher clinical utility.  
287 When all ULNTs are negative, CR can be ruled out, whereas when 3 of 4 tests are positive,  
288 CR can be ruled in. As such, we recommend the use of ULNT tests as combinations only.  
289 Our study does not test the validity of ULNT tests for specific nerve trunks, which it is  
290 hypothesized to perform.

291

292

293

294

295

296 **Table 1:** Baseline characteristics of the subjects (n=85)

297 \* *Wilcoxon rank sum test*

298

299

300 **Table 2:** Diagnostic Accuracy of Individual Upper Limb Neurodynamic Tests. Pre-test

301 Prevalence = 31.7%.

302

303 95%CI: Confidence interval at 95%

304 LR+: Positive likelihood ratio

305 LR-: Negative likelihood ratio

306 ULNT: Upper limb neurodynamic test

307

308

309 **Table 3.** Diagnostic Accuracy of Clustered Upper Limb Neurodynamic Test findings

310 (Conditions). Pre-test Prevalence = 31.7%.

311

312 95% CI: Confidence interval at 95%

313 LR+: Positive likelihood ratio

314 LR-: Negative likelihood ratio

315

316

317

318

319

320

321

322

323

324

325

326

327

328



329 **Appendix A.** Standard sequence of joint movements and suggested structural differentiation  
330 maneuvers (sensitizing movements at a site distant to the pain) for each ULNT (Nee et al,  
331 2012)

ULNT 1 (median nerve) :

- Shoulder girdle stabilization
- Shoulder abduction
- Wrist/finger extension
- Forearm supination
- Shoulder external rotation
- Elbow extension
- Structural differentiation:  
Cervical side bending or release wrist extension



ULNT 2a (median nerve) :

- Shoulder girdle depression
- Elbow extension
- Shoulder external rotation and forearm supination
- Wrist/finger extension
- Shoulder abduction
- Structural differentiation:  
Cervical side bending, or release shoulder girdle depression or release wrist extension



ULNT 2b (radial nerve) :

- Shoulder girdle depression
- Elbow extension
- Shoulder external rotation and forearm pronation
- Wrist/finger flexion
- Shoulder abduction
- Structural differentiation :  
Release shoulder girdle depression or release wrist flexion



ULNT 3 (ulnar nerve) :

- Wrist/finger extension
- Forearm pronation
- Elbow flexion
- Shoulder external rotation
- Shoulder girdle depression
- Shoulder abduction
- Structural differentiation :  
Cervical side bending, or release shoulder girdle depression or release wrist extension



332

333

334

335 **Abbreviations :**

336 CR : Cervical radiculopathy

337 ULNT : Upper limb neurodynamic tests

338 MRI : magnetic resonance imaging

339 LR+ : Positive likelihood ratio

340 LR- : Negative likelihood ratio

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357 **References**

- 358 1. Anekstein Y, Blecher R, Smorgick Y, Mirovsky Y. What is the best way to apply the  
359 spurling test for cervical radiculopathy? *Clinical Orthopaedics and Related Research:*  
360 *The Association of Bone and Joint Surgeons.* 2012; 470(9):2566-2572. Doi:  
361 10.1007/s11999-012-2492-3
- 362 2. Radhakrishnan K, Litchy WJ, O'Fallon WM, Kurland LT. Epidemiology of cervical  
363 radiculopathy. A population-based study from Rochester, Minnesota, 1976 through  
364 1990. *Brain.* 1994;117 ( Pt 2):325-335. doi:10.1093/brain/117.2.325
- 365 3. Woods, Barrett I. MD; Hilibrand, Alan S. MD Cervical Radiculopathy, *Journal of*  
366 *Spinal Disorders and Techniques:* June 2015 - Volume 28 - Issue 5 - p E251-E259 doi:  
367 10.1097/BSD.0000000000000284
- 368 4. Mink JH, Gordon RE, Deutsch AL. The cervical spine: radiologist's perspective. *Phys*  
369 *Med Rehabil Clin N Am.* 2003;14(3):493-vi. doi:10.1016/s1047-9651(03)00035-4
- 370 5. Reza Soltani Z, Sajadi S, Tavana B. A comparison of magnetic resonance imaging  
371 with electrodiagnostic findings in the evaluation of clinical radiculopathy: a cross-  
372 sectional study. *Eur Spine J.* 2014;23(4):916-921. doi:10.1007/s00586-013-3164-z
- 373 6. Ashkan K, Johnston P, Moore AJ. A comparison of magnetic resonance imaging and  
374 neurophysiological studies in the assessment of cervical radiculopathy. *Br J*  
375 *Neurosurg* 2002;16:146–8.
- 376 7. Crette S, Fehlings MG. Clinical practice. Cervical radiculopathy. *N Engl J Med.*  
377 2005;353(4):392-399. doi:10.1056/NEJMcp043887
- 378 8. Kuijper B, Tans JT, Schimsheimer RJ, et al. Degenerative cervical radiculopathy:  
379 diagnosis and conservative treatment. A review. *Eur J Neurol.* 2009;16(1):15-20.  
380 doi:10.1111/j.1468-1331.2008.02365.x

- 381 9. Elvey RL. Treatment of arm pain associated with abnormal brachial plexus tension.  
382 Aust J Physiother . 1986;32(4):225-30. doi: 10.1016/S0004-9514(14)60655
- 383 10. Butler, DS. The Sensitive Nervous System Unley, NOI Group Publications. Adelaide,  
384 Australia 2000
- 385 11. Shacklock, MO. (1996). Positive upper limb tension test in a case of surgically proven  
386 neuropathy: analysis and validity. *Manual Therapy*, 1996;1(3): 154-161. Doi:  
387 10.1054/math.1996.0265
- 388 12. Thoomes EJ, Van Gees S, Van der Windt DA, Falla D, Verhagen AP, Koes BW,  
389 Thoomes-de Graaf M, Kuijper B, Scholten-Peeters WGM, Vleggeert-Lankamp CL.  
390 Value of physical tests in diagnosing cervical radiculopathy: a systematic review.  
391 *Spine Journal*. 2017; 18(1):179-189. doi: 10.1016/j.spinee.2017.08.241
- 392 13. Apelby-Albrecht M, Andersson L, Kleiva IW, Kvale K, Skillgate E, Josephson A.  
393 Concordance of Upper Limb Neurodynamic Tests with medical examination and  
394 magnetic resonance imaging in patients with cervical radiculopathy: a diagnostic  
395 cohort study. *Journal of Manipulative and Physiological Therapeutics*.  
396 2013;36(9):626-632. Doi: 10.1016/j.jmpt.2013.07.007
- 397 14. Wainner RS, Fritz JM, Irrgang JJ, Boninger ML, Delitto A, Allison S. Reliability and  
398 Diagnostic Accuracy of the Clinical Examination and Patient Self-Report Measures  
399 for Cervical Radiculopathy. *Spine*, 2003;28 (1):52-62.
- 400 15. Ghasemi M, Golabchi K, Mousavi SA, et al. The value of provocative tests in  
401 diagnosis of cervical radiculopathy. *J Res Med Sci* . 2013 Mar;18(Suppl 1):S35-8
- 402 16. Schiffman, E., Ohrbach, R., Truelove, E., Look, J., Anderson, G., Goulet, JP, ...  
403 Dworkin, SF. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for  
404 Clinical and Research Applications: Recommendations of the International

- 405 RDC/TMD Consortium Network and Orofacial Pain Special Interest Group. *J Oral*  
406 *Facial Pain Headache*. 2014;28(1):6-27. Doi: 10.11607/jop.1151
- 407 17. Nee RJ, Jull GA, Vicenzino B, Coppieiers MW. The validity of upper-limb  
408 neurodynamic tests for detecting peripheral neuropathic pain. *The Journal of*  
409 *Orthopaedic and Sports Physical Therapy*. 2012; 42(5):413-424. Doi:  
410 10.2519/jospt.2012.3988
- 411 18. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. .  
412 STARD 2015: an updated list of essential items for reporting diagnostic accuracy  
413 studies. *BMJ*. (2015) 351:h5527. 10.1148/radiol.2015151516
- 414 19. Horn ME, Brennan GP, George SZ, Harman JS, Bishop MD. A value proposition for  
415 early physical therapist management of neck pain: a retrospective cohort analysis.  
416 *BMC Health Serv Res*. 2016; 16: 253.
- 417 20. Masaracchio M, Cleland JA, Hellman M, Hagins M. Short-term combined effects of  
418 thoracic spine thrust manipulation and cervical spine nonthrust manipulation in  
419 individuals with mechanical neck pain: a randomized clinical trial. *J Orthop Sports*  
420 *Phys Ther* 2013;43(3):118e27.
- 421 21. Schmid AB, Brunner F, Luomajoki H, et al. Reliability of clinical tests to evaluate  
422 nerve function and mechanosensitivity of the upper limb peripheral nervous system.  
423 *BMC Musculoskelet Disord*. 2009;10:11. Published 2009 Jan 21. doi:10.1186/1471-  
424 2474-10-11
- 425 22. Walsh MT. (2005). Upper limb neural tension testing and mobilization. Fact, fiction,  
426 and a practical approach. *J Hand Ther*. 2005;18(2):241-258. DOI:  
427 10.1197/j.jht.2005.02.010
- 428 23. Cook CJ, Cook CE, Reiman MP, Joshi AB, Richardson W, Garcia AN. Systematic  
429 review of diagnostic accuracy of patient history, clinical findings, and physical tests in

430 the diagnosis of lumbar spinal stenosis. *Eur Spine J.* 2020;29(1):93-112.  
431 doi:10.1007/s00586-019-06048-4

432 24. Koulididis K, Veremis Y, Anderson C, Heneghan NR. Diagnostic accuracy of upper  
433 limb neurodynamic tests for the assessment of peripheral neuropathic pain: A  
434 systematic review. *Musculoskelet Sci Pract.* 2019 Apr;40:21-33. doi:  
435 10.1016/j.msksp.2019.01.001. Epub 2019 Jan 12. PMID: 30665045

436 25. Hegedus EJ, Stern B. Beyond SpPIN and SnNOUT: Considerations with Dichotomous  
437 Tests during Assessment of Diagnostic Accuracy. *J Man Manip Ther.* 2009;17(1):E1-  
438 E5. doi:10.1179/jmt.2009.17.1.1E

439 26. Baeyens JP, Serrien B, Goossens M, Clijsen R. Questioning the “SPIN and SNOUT”  
440 rule in clinical testing. *Arch Physiother* 9, 4 (2019). [https://doi.org/10.1186/s40945-](https://doi.org/10.1186/s40945-019-0056-5)  
441 [019-0056-5](https://doi.org/10.1186/s40945-019-0056-5)

442 27. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical  
443 informatics. *J Biomed Informatics.* 2014;48:193-204.

444 28. Bujang MA, Adnan TH. Requirements for Minimum Sample Size for Sensitivity and  
445 Specificity Analysis. *J Clin Diagn Res.* 2016;10(10):YE01-YE06.  
446 doi:10.7860/JCDR/2016/18129.8744

447 29. Schmid, A.B., Nee, R.J., Coppieters, M.W., 2013. Reappraising entrapment  
448 neuropathies– mechanisms, diagnosis and management. *Man. Ther.* 18 (6), 449–457.  
449  
450  
451  
452  
453  
454  
455  
456  
457

458

459 **Acknowledgments** : The authors thank the assistance of François-Xavier Bénard, Adeline  
460 Vergé, Cécile Roesch and Audrey Tseng Qun.

461

462

463

464

465

466