

SYSTEMATIC REVIEW

Test procedures and positive diagnostic criteria of the upper limb tension tests differ: a systematic review of the DiTA database



Arianne P Verhagen^{a,*}, Hayley Brown^a, Mark Hancock^b, David Anderson^{c,d}

^a *Discipline of Physiotherapy, Graduate School of Health, University of Technology Sydney, Sydney, Australia*

^b *Faculty of Medicine, Health and Human Sciences, Macquarie University, Sydney, Australia*

^c *School of Health Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia*

^d *Sydney Spine Institute Research Group, Sydney Spine Institute, Burwood, Sydney, Australia*

Received 28 July 2022; received in revised form 12 October 2023; accepted 25 October 2023
Available online 10 November 2023

KEYWORDS

Diagnostic accuracy;
Positive diagnostic criteria;
Positivity threshold;
Systematic review;
Test procedures;
Upper limb tension test

Abstract

Background: The validity of the ULTT is unclear, due to heterogeneity of test procedures and variability in the definition of a positive test

Objective: To evaluate test procedures and positive diagnostic criteria for the upper limb tension test (ULTT) in diagnostic test accuracy studies.

Methods: A systematic review of diagnostic accuracy studies was performed. We conducted a search of the DiTA (Diagnostic Test Accuracy) database and selected primary studies evaluating the diagnostic accuracy of the ULTT. We assessed risk of bias, performed data extraction on study characteristics, test procedures, and positive diagnostic criteria, and performed a descriptive analysis.

Results: We included nine studies (681 participants), four diagnosing people with cervical radiculopathy (CR), four diagnosing people with carpal tunnel syndrome (CTS), and one included both CR and CTS. The risk of bias varied between 2 and 6 out of 6 positive items. Eight studies reported on the ULTT1 (median nerve).

Overall, all studies clearly described their test procedures and positive diagnostic criteria although the order of movements and the diagnostic criteria between studies varied. We suggest a more standardised test procedure for the ULTT1 to consist of: 1) stabilising the shoulder in abduction, 2) extending the wrist/fingers, 3) supinating the forearm, 4) externally rotating the shoulder, 5) extending the elbow, and finally 6) performed structural differentiation by side bending (lateral flexion) of the neck. This proposed test procedure should reproduce the symptoms and enables the clinician to evaluate whether symptoms increase/decrease when stressing or relaxing the nerves.

Declarations: This work has not been published previously, or is under consideration for publication elsewhere.

* Corresponding author at: Discipline of Physiotherapy, Graduate School of Health, University of Technology Sydney, Sydney, Australia.
E-mail: arianne.verhagen@uts.edu.au (A.P. Verhagen).

<https://doi.org/10.1016/j.bjpt.2023.100558>

1413-3555/© 2023 Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. Published by Elsevier España, S.L.U. All rights reserved.

Conclusion: Based on our findings we proposed a more standardised test procedure for the ULTT1 with accompanying positive diagnostic criteria to facilitate homogeneity in future diagnostic accuracy studies of the ULTT.

© 2023 Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. Published by Elsevier España, S.L.U. All rights reserved.

Introduction

Upper limb tension tests (ULTT), also called upper limb neurodynamic tests (ULNT), are a commonly used neurodynamic technique by clinicians to evaluate nerve gliding and neural tension in patients.¹ Neurodynamic tests aim to investigate if a peripheral nerve is contributing to the patients' pain, resulting in peripheral neuropathic pain, by moving and stretching the peripheral nerves.^{2,3} When utilizing a neurodynamic test, the assessor manipulates the limb into various positions with the aim of increasing tension on peripheral nerves.^{4,5} The positioning of each joint is added to provoke the pain or reproduce the symptoms.⁶

Variations of the ULTT aim to selectively differentiate between nerves of the upper limb by altering the type of movement and movement sequence. The ULTT1 and 2a aim to test the median nerve, the ULTT2b the radial nerve, and ULTT3 the ulnar nerve. The ability of the ULTT to selectively test specific nerve roots is however questionable.^{3,7,8}

The diagnostic strategy to diagnose a compression of the nerve includes, apart from clinical signs and symptoms and tests like the ULTT, electromyography, nerve conduction studies, and MRI (magnetic resonance imaging).⁹ Clinicians often use the ULTT to help diagnose whether a patient has peripheral neuropathic pain, like cervical radiculopathy (CR), carpal tunnel syndrome, or cubital tunnel syndrome.^{2,3,5}

Before any test can be endorsed for use in clinical practice, its validity needs to be determined. Systematic reviews state that the validity of the ULTT is unclear, due to heterogeneity of test procedures and variability in the definition of a positive test.^{10,11} For clinicians, several websites and videos exist (apart from textbooks) to help them perform these tests accurately and claim to be evidence-based.^{2,12,13} Unfortunately, test procedures of the ULTT vary in the literature as well as on these major websites.

Today no current uniform set of positive diagnostic criteria exists for the ULTT, creating uncertainty on what constitutes a positive test.^{10,11} Positive test criteria (positivity thresholds) used in literature and clinical practice include, amongst others, reproduction of patient symptoms with differences between sides, relief or exacerbation of symptoms on structural differentiation, and reproduction of neurological pain associated with the nerve distribution.^{14,15}

Current systematic reviews focus on the accuracy data, irrespective of the test procedures and variation in positive diagnostic criteria.^{10,11} Given that there is a lack of test procedure standardization and there are no clear cut-offs for positive tests available, it brings into question the usefulness of these tests in clinical practice and the validity of the accompanying accuracy data. The establishment of a standardized procedure and clear positive diagnostic criteria needs to be done first before we evaluate the diagnostic accuracy of these tests. When the diagnostic accuracy is established and the ULTT can rule in or rule out patients

with a possible CR or CTS, these tests might have greater clinical application and improve patient outcomes. Additionally, this may lead to a reduction in patient waiting time for nerve conduction studies or expensive diagnostic imaging. Therefore, this study aimed to evaluate test procedures and positive diagnostic criteria for the ULTT in diagnostic test accuracy studies and to construct a proposed set of recommended test procedures and positivity thresholds.

Methods

Study design

A systematic review using the DiTA (Diagnostic Test Accuracy) database was performed. DiTA database is a comprehensive index, which is updated monthly using automated searches, of diagnostic test accuracy studies developed specifically for the discipline of physical therapy.^{16,17} It is a sister database of the PEDro (Physiotherapy Evidence Database) which includes intervention studies within the physical therapy domain. We registered our protocol on the UTS (University of Technology Sydney) data repository, and it can be requested from the corresponding author.

Search strategy

The search was conducted on the 22 November 2021 using the search terms: Upper Limb Tension Test and associated synonyms, like ULTT, ULNT, Brachial Plexus Tension, Elvey Test, Upper limb nerve tension test, Peripheral neuropathic pain, Cervical radiculopathy, Carpal tunnel syndrome, Cubital tunnel syndrome.

Study selection

The inclusion criteria consisted of a) adult participants (18 years or over) who present with signs and symptoms that indicate upper limb pathology, b) evaluated sensitivity and/or specificity of the ULTT in diagnosing upper limb pathology by comparing the ULTT to any reference test, c) primary diagnostic test accuracy studies. We excluded studies using non-human subjects or cadavers and where the ULTT was used in combination with other special tests for diagnosis. The eligibility of each study, identified by the search or included in (systematic) reviews, was determined by two out of five independent assessors (GB, CB, CA, HD, SF). Discrepancies were resolved by consensus or by a third independent investigator (AV).

Risk of bias assessment

The QUADAS-2 is the recommended tool to assess risk of bias (RoB) in diagnostic test accuracy studies.¹⁸ As this tool is

Table 1 Criteria for risk of bias in a diagnostic test accuracy study.³¹

1. A consecutive or random sample of patients was enrolled
2. The study avoids inappropriate exclusions
3. The index test and reference test are assessed independently (i.e. without knowledge of the outcome of the other test)
4. The threshold for a positive test is clear
5. All participants received both tests
6. The timing between the performance of the index and reference tests is clearly appropriate (i.e. the result on either test is unlikely to change during that time)

rather difficult to administer for clinicians/researchers, we simplified it to six criteria (Table 1).¹⁹ Each criterion can be scored yes/no/unclear. Two out of five assessors (GB, CB, CA, HD, SF) independently assessed each study and any discrepancies were discussed and resolved by consensus. If a consensus could not be reached, a third independent investigator (AV) made a final decision.

Data extraction

Two of five assessors (GB, CB, CA, HD, SF) extracted the following data from each of the studies: author(s) and year of publication, participant details (number, mean age and range, sex, clinical characteristics), examination details (clinical setting, examiner profession and expertise), reference standard, test procedures, and criteria for positive test result, e.g. pain, range of motion (either actively or passively performed). Discrepancies and a data check was done by two review authors (HB, AV).

Analysis

We considered a specificity or sensitivity value high enough to be clinically useful in ruling in or out a condition, according to the SpPIn (specificity high and positive test rules in a condition) and SnNOut (sensitivity high and negative test rules out a condition) rules. As most diagnostic test accuracy studies are performed in highly specific populations, generalising their accuracy to the public requires sensitivity and specificity to be sufficiently high, with recommended cut-offs between 0.85 and 0.90–0.95 being used.^{20,21} We conducted a frequency analysis of the number of papers outlining the test protocol used for the ULTT to determine the common denominator of those procedures. Similarly, a descriptive analysis was completed to assess the positive diagnostic criteria.

Results

Search results

From the initial search in the DiTA database, 22 original studies were retrieved, in addition to two systematic reviews^{10,11} and three narrative reviews.^{3,22,23} After reviewing the primary research papers included in these reviews and removing duplicates, we retrieved an additional 5

references, resulting in 27 papers being available for full-text review (Fig. 1). After the application of the selection criteria, we excluded three papers as these were narrative reviews,^{3,22,23} three papers did not evaluate the ULTT,²⁴⁻²⁶ and one was in a population with low back pain.²⁷ Finally, we included 10 papers reporting on 9 studies.

Description of included studies

Participants. Four studies aimed at diagnosing CR (386 participants),^{14,28-30} and four studies aimed at diagnosing CTS in 295 participants.³¹⁻³⁴ One study (in two publications) included both patients with CR and CTS (Table 2).^{35,36} Individuals with CR were all referred to a specialized clinic or a neurosurgery department. Three out of the four studies on patients with CR specifically stated they included consecutive patients. The average age of the participants varied between 43.2 and 54.3 years and most of the participants were female (varying between 49 and 83%).

Index test. The most evaluated test was ULTT1; five studies only evaluated the ULTT1, two studies evaluated all four ULTT tests and provided accuracy data for all tests separately,^{14,29} one study did not specify which ULTT was evaluated,²⁸ and one study evaluated the ULTT-A (which was like the ULTT1) and ULTT-B (an alternative to ULTT1).³⁵

Reference test. Nerve conduction study was the most common reference test mainly in the studies investigating CTS,³¹⁻³⁵ followed by clinical presentation and magnetic resonance imaging (MRI), in the studies investigating CR.^{14,29,30} The CR studies reported history taking and physical examination to be part of the reference test, and in two (out of four) studies the reference test was performed by one specialist, while none of the CTS studies did so.

Accuracy data. All studies provided data on the accuracy of the ULTT, and eight studies reported data for the ULTT1. Based on the sparsity of data related to other ULTT tests, we only report outcomes for the ULTT1. Sensitivity of the four test procedures for the ULTT1 for CR varied between 0.35 and 0.83, and the specificity varied between 0.40 and 0.76, meaning none met the lowest cut-off mentioned in literature of 0.85. For CTS studies, which reported on eight test procedures, reported sensitivity was between 0.06 and 0.93 and specificity between 0.10 and 0.93. In three procedures the accuracy data met at least the 0.85 cut-off, but not the 0.95.

Risk of bias. The RoB varied from 2 to 6 criteria scored positive (Table 3).

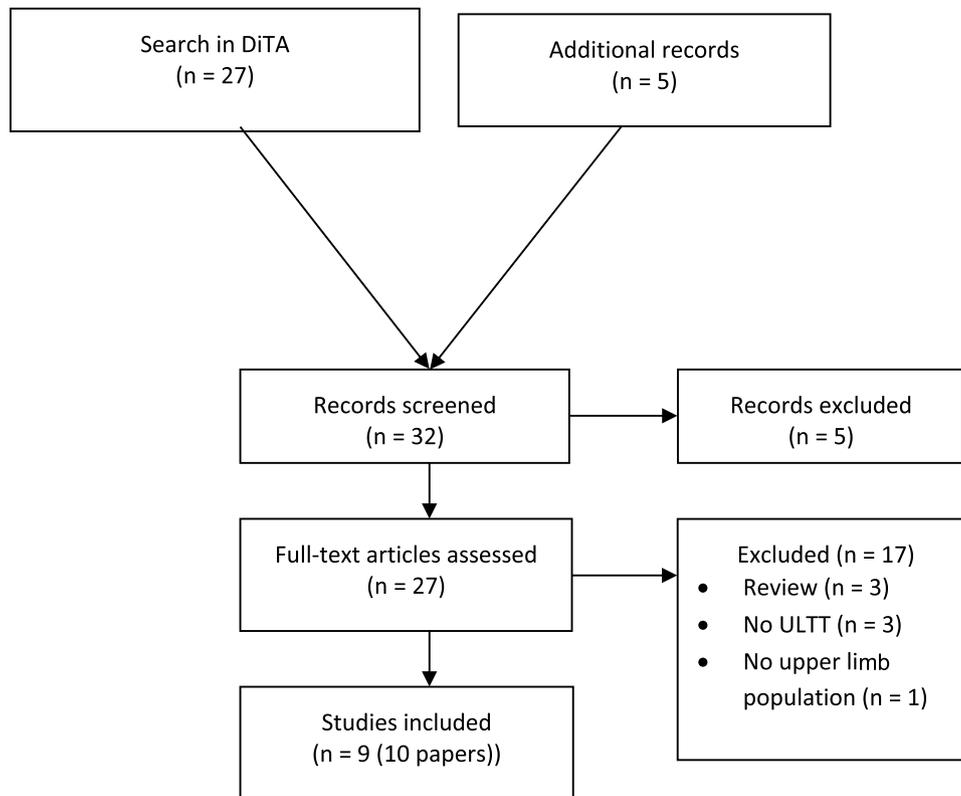


Fig. 1 PRISMA flow diagram.

Test procedures

We found a wide variation with a total of 13 different ULTT1 procedures (Table 4). One study did not report the procedure and just provided a reference.^{31,37} Three other studies provided the procedure and mentioned a reference of the origin of the procedure.^{29,31,34} All papers mentioned between two and seven different steps in the procedure, unfortunately, no two studies performed the exact same procedure except Vanti et al.^{33,34} For one study we presented the procedure for the ULTT-A, as this one is most like the ULTT1.³⁵

In six studies, shoulder depression (which includes shoulder girdle/scapula depression/fixation/stabilisation) is listed as the starting position. One study did not mention shoulder depression and started with the shoulder in abduction,²⁸ one study started with a contralateral flexion of the neck,³¹ and one study combined most procedures that are separate steps in all other studies into one starting position and added just one extra step (elbow extension).³² Six studies ended the procedure with the last step being ‘contralateral/ipsilateral cervical lateral flexion (bending)’, while one study started with this procedure and released it at the last step.³¹ Two studies did not mention cervical lateral flexion at all.^{28,32}

Positive diagnostic criteria

All studies clearly stated their positive diagnostic criteria, although the criteria between studies differed (Table 5). In

most studies, multiple symptoms were required for a test to be regarded positive, either in combination or as an ‘either/or’. Three studies evaluated more than one set of positive diagnostic criteria.^{31,33,34} ‘Reproducible patient symptoms’ and ‘symptoms decrease when relaxing the stress on the nerves’ were the most used criteria ($n = 9$ and $n = 6$, respectively). With ‘symptoms decrease when relaxing the stress on the nerves’ (or increased symptoms with increased stretching) authors mean relaxing the stretch on the arm; ‘structural differentiation’ means that symptoms were evaluated to either increase or decrease when the cervical spine was laterally flexed. Three studies used a combination of criteria (Wainner’s criteria: “reproducible patient symptoms OR limited range of motion of the elbow OR decrease of symptoms when relaxing”) as defined by Wainner et al.^{32,34,35} The criterion ‘symptoms in the first 3 digits’ was only mentioned in three studies including patients with CTS.^{31,33,34}

Discussion

Main findings

We found nine diagnostic test accuracy studies of which four included patients with CR, four studies patients with CTS, and one included both patients. The most evaluated test was the ULTT1, with in total 13 test procedures described in eight studies. We found a wide variety of test procedures and criteria for a positive test. Based on our findings, we propose a more standardised test procedure of the ULTT1 in patients with CR as well as CTS. The associated positive

Table 2 Characteristics of included studies.

Study	Participants	Index test	Reference test
Cervical radiculopathy Apelby-Albrecht 2013 Sweden	Cervical Radiculopathy; <i>n</i> = 51 consecutive, referred to neurosurgeon; age: 51 (25–67); female: 27 (53%)	ULTT all combined and 1, 2A- B, 3 separate (manual thera- pists)	Combination of history, clini- cal examination, and MRI (neurosurgeons)
Ghasemi 2013 Iran	Cervical Radiculopathy; <i>n</i> = 97 referred to specialised diagnostic centre; age: 46.3; female: 72 (74.2%)	ULTT (trained examiners)	Nerve conduction studies (EDX) (neurologist)
Grondin 2021 France	Cervical Radiculopathy; <i>n</i> = 85 consecutive, referred to neurosurgery depart- ment; age: 44; female: not reported?	ULTT 1, 2A-B, 3 (one experi- enced physical therapist)	Combination of history, clini- cal examination, and MRI (one experienced neuro- surgeon)
Sleijser-Koehorst 2020 Netherlands	Cervical Radiculopathy; <i>n</i> = 134, consecutive, referred to specialised clinic; age: 49.9 (10.7); female: 66 (49%)	ULTT1 (physical therapist)	Clinical presentation + MRI (neurosurgeon)
Carpal Tunnel Syndrome Bueno-Garcia 2016 Spain	Carpal Tunnel Syndrome; <i>n</i> = 58; age: 54.3 (14.5); female: 44 (75.8%)	ULTT1 (physical therapist)	NCS (neurophysiologist)
Trillos 2018 Colombia	Carpal Tunnel Syndrome; <i>n</i> = 118 (230 wrists); age: 50.5 (18–86); female: 98 (83.1%)	ULTT1 (physical therapist)	NCS (physiatrist)
Vanti 2011	Carpal Tunnel Syndrome <i>n</i> = 44; age: 46.3 (10.8); female: 33 (75%)	ULTT1 (physical therapist)	NCS (experienced tester)
Vanti 2012	Carpal Tunnel Syndrome; <i>n</i> = 47 (84 limbs); age: 45.9 (10.7); female: 35 (74.5%)	ULTT1 (physical therapist)	NCS (experienced tester)
Mixed population Wainner 2003* / 2005	<i>N</i> = 82: Cervical Radiculop- athy; <i>n</i> = 19; age: 43.2 (11.7) Carpal Tunnel Syndrome; <i>n</i> = 28; age: 48.4 (\pm 11.5) Female: 41 (50%)	ULTT-A, ULTT-B	NCS (physiatrist or neurolo- gist); needle EMG

Data are mean (standard deviation), mean (range), or frequency (proportion).

CTS, carpal tunnel syndrome; EDX, electro diagnostic studies; EMG, electromyography; MRI, magnetic resonance imaging; NCS, nerve conduction study; ULTT, upper limb tension test (ULTT1 and 2A aim to test the median nerve, the ULTT2B the radial nerve and ULTT3 the ulnar nerve).

* data presented from the 2003 paper.

Table 3 Risk of bias assessment.

Study	Consecutive or random sample	No inappropriate exclusions	Index and reference test assessed independently	Clear threshold for positive test	All participants received both test	Timing between tests is appropriate	Total score
Apelby-Albrecht 2013	Yes	Yes	Unclear	Yes	Yes	Unclear	4
Bueno-Garcia 2016	Unclear	Yes	Yes	Yes	Yes	Unclear	4
Ghasemi 2013	Unclear	Unclear	Unclear	Yes	Yes	Unclear	2
Grondin 2021	Yes	Yes	Yes	Yes	Yes	Yes, 1 hr	6
Sleijser-Koehorst 2020	Yes	Yes	Yes	Yes	No	Unclear	4
Trillos 2018	Yes	Unclear	Yes	Yes	Yes	Yes, 20 min	5
Vanti 2011	Yes	Yes	Yes	Yes	Yes	Yes, 20–30 min	6
Vanti 2012	Yes	Unclear	Yes	Yes	Yes	Yes, 20–30 min	5
Wainner 2003/2005	Yes	Yes	Yes	Yes	Yes	Unclear	5

Table 4 Test procedure (order of movements) of ULTT1.							
Study	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7
<i>Cervical radiculopathy</i>							
Apelby-Albrecht 2013	Shoulder depression	Shoulder abduction, 110°	Wrist and finger extension	Forearm supination	Shoulder external rotation	Elbow extension	Contra/ipsi-lateral cervical lateral flexion
Ghasemi 2013		Shoulder abduction	Elbow				
extension + forearm supination	Wrist extension						
Grondin 2021 (Nee et al. 2012)	Shoulder stabilisation	Shoulder abduction	Wrist/finger extension	Forearm supination	Shoulder external rotation	Elbow extension	Cervical side bending OR release wrist extension
Sleijser-Koehorst 2020	Shoulder fixation	Shoulder abduction, 90°	Forearm supination	Wrist and finger extension	Shoulder external rotation	Elbow extension	Contra/ipsi-lateral cervical lateral flexion
<i>Carpal tunnel syndrome</i>							
Bueno-Garcia 2016; (Shacklock 2005)	Contralateral flexion of the neck	Shoulder abduction				Elbow extension	Release of cervical lateral flexion
Trillos 2018	Starting position: 90° shoulder abduction + 90° shoulder external rotation + 90° elbow flexion + forearm supination + maximum extension of wrist and fingers + abduction of the thumb					Elbow extension	
Vanti 2011 (Butler 2000)	Shoulder stabilisation	Shoulder abduction, 110°	Wrist and finger extension	Forearm supination	Shoulder external rotation	Elbow extension	Contra/ipsi-lateral cervical lateral flexion
Vanti 2012	Shoulder stabilisation	Shoulder abduction	Wrist and finger extension	Forearm supination	Shoulder external rotation	Elbow extension	Contra/ipsi-lateral cervical lateral flexion
<i>Mixed population</i>							
Wainner 2003 / 2005 (ULTT-A)	Shoulder depression	Shoulder abduction	Forearm supination, wrist and finger extension	Shoulder external rotation		Elbow extension	Contra/ipsi-lateral cervical lateral flexion

Table 5 Number of studies that identified each criterion for a positive test for ULTT1.

Study	Crit 1	Crit 2	Crit 3	Crit 4	Sensitivity/Specificity (95%CI)
Cervical radiculopathy					
Apelby-Albrecht 2013	Reproducible patient symptoms		Side differences	Symptoms increase during stretching	Sens: 0.83 (0.66, 0.93)
Ghasemi 2013					Spec: 0.75 (0.48, 0.93)
Grondin 2021 (Nee et al. 2012)	Reproducible patient symptoms			Structural differentiation*	Sens: 0.35 – 0.60 [§]
Sleijser-Koehorst 2020	Reproducible patient symptoms			Structural differentiation*	Spec: 0.40
Carpal tunnel syndrome					
Bueno-García 2016 (Shacklock 2005)	Reproducible patient symptoms	Symptoms in first three digits	Limited Range of Motion	Structural differentiation*	Sens: 0.58 (0.45, 0.71)
Trillos 2018					Spec: 0.84 (0.72, 0.96)
Vanti 2011 (Butler 2000)	Reproducible patient symptoms		Limited Range of Motion	Symptoms increase during stretching	Sens: 0.74 (0.61, 0.83)
Vanti 2012		Symptoms in first three digits		Structural differentiation*	Spec: 0.50 (0.35, 0.65)
		Symptoms in first three digits		Symptoms increase during stretching	Sens: 0.93 (0.88, 0.98)
		Symptoms in first three digits		Symptoms increase during stretching	Spec: 0.10 (0.00, 0.34)
		Symptoms in first three digits		Structural differentiation*	Sens: 0.92 (0.74, 0.97)
Mixed population					
Wainner 2003 / 2005 (ULTT-A)	Reproducible patient symptoms		Limited Range of Motion	Symptoms increase during stretching	Spec: 0.15 (0.05, 0.36)
				Structural differentiation*	Sens: 0.54 (0.35, 0.71)
				Structural differentiation*	Spec: 0.70 (0.48, 0.85)
				Structural differentiation*	Sens: 0.40 (0.26, 0.89)
				Structural differentiation*	Spec: 0.78 (0.66, 0.89)
				Structural differentiation*	Sens: 0.29 (0.16, 0.45)
				Structural differentiation*	Spec: 0.82 (0.69, 0.91)
				Structural differentiation*	Sens: 0.06 (0.02, 0.19)
				Structural differentiation*	Spec: 0.93 (0.82, 0.98)

CI, confidence interval; crit, criterion, Sens, sensitivity; Spec, specificity.

* = symptoms were evaluated to either increase or decrease when the cervical spine was side bended/laterally flexed.

§ = no 95% confidence intervals presented.

diagnostic criteria are reproduction of the symptoms and an increase or decrease of symptoms when stressing or relaxing the nerves.

Comparison with existing literature

Previously, several reviews evaluating the ULTT concluded that there is a lack of a clear definition of terms ('investigators definition of a positive test') and procedures for the ULTT1.^{3,10,23} The authors all concluded that the validity of ULTT1 is probably hampered by the diversity in the procedure and interpretation of the index test. The major online support tools also use a variety of procedures in their videos, which reflects the variability in test procedures found in the literature and hampers clear implementation.^{2,12,13} By evaluating the test procedures of the ULTT1 we found that the order of movements varied between studies. Based on the included studies, we suggest a more standardised test procedure and positive diagnostic criteria. Hopefully in future research the diagnostic accuracy of the ULTT1 can be more accurately assessed through a standardised procedure and positive diagnostic criteria.

Proposed set of test procedures and positive diagnostic criteria

Overall, all studies stabilised the shoulder in abduction, extended the elbow, supinated the forearm, extended the wrist, and performed structural differentiation by side bending of the neck, although the order of movements varied. We propose a more standardised test procedure to consist of

these movements. The suggested order is like the one performed in the studies with the lowest risk of bias (score of 6): 1) stabilising the shoulder in abduction, 2) extending the wrist/fingers, 3) supinating the forearm, 4) externally rotate the shoulder, 5) extending the elbow, and 6) structural differentiation by laterally flexing (i.e., side bending) the neck. This proposed test procedure aims to assess whether the patient's symptoms are reproduced and whether symptoms increase or decrease when increasing or decreasing the tension on the nerves. When symptoms are present in the first three digits during the ULTT, this acts as an extra diagnostic criteria for patients with CTS, and can be regarded as 'reproduction of symptoms', same for 'limited range of motion'.

Strengths and limitations

To our knowledge, this is the first systematic review evaluating the test procedures and positive diagnostic criteria of the ULTT1 in patients with neck or arm pain. A possible limitation of our review may exist in the search strategy, which was conducted solely in the DiTA database. DiTA is updated monthly, drawing on automated optimised searches of MEDLINE, EMBASE, CINAHL, and the Cochrane Database of Systematic Reviews.³⁸ Therefore, we consider the DiTA database up to date and when we performed a search in PubMed (January 2022) we did not find any additional or missed studies. Another limitation may be present in the study selection, as patients selected in the included studies are specific patients, referred to specialised clinics. Although this finding probably does not influence test

procedures and positive diagnostic criteria, it might influence the accuracy data as these are dependent on prevalence of the condition.³⁹ Concerning the reference test, in the CR studies it included physical examination in three of the four studies, potentially leading to risk of confirmation bias that might result in higher accuracy data. In two CR studies it was also unclear whether the index and reference test were performed independently. Our findings, however, did not show higher accuracy data in the CR studies, therefore we believe that confirmation bias might not have played an important role. Another limitation could be our risk of bias tool. We used a simpler 6 item 'checklist' for risk of bias assessment, to allow clinicians and students to conduct risk of bias assessments more easily. We are confident that the outcomes of the risk of bias assessment will not differ much from the QUADAS-2. We registered our protocol on the UTS data repository only, as we were unable to register the protocol in Prospero which focusses on systematic reviews with a patient related/relevant outcome. Lastly, due to the limited data, we were only able to make recommendations about the ULTT1, and not the ULTT2 (a and b) and ULTT3.

Implications

Clinical implications. We proposed a standardised test procedure and positive diagnostic criteria for use in clinical practice. A more standardised set of test procedures will help clinicians as it is a consistent message to patients. Reproduction of symptoms and a decrease of symptoms when reducing the tension are the recommended criteria for a positive test.

Research implications. We suggest that future research should validate our proposed set of test procedures as well as our proposed set of positive diagnostic criteria. This may lead to greater consistency and standardisation between studies, reducing heterogeneity regarding treatment effectiveness studies, systematic reviews, and meta-analysis of test accuracy studies.

Conclusion

Although the ULTT continues to be used by clinicians, its diagnostic accuracy remains uncertain. Given the ULTT is likely to remain in use by clinicians, we have provided a recommendation on the test protocol, based on the nine studies we found, which includes: 1) stabilising the shoulder in abduction, 2) extending the wrist/fingers, 3) supinating the forearm, 4) externally rotating the shoulder, 5) extending the elbow, and 6) performing structural differentiation by lateral flexion (side bending) of the neck aiming to reproduce the symptoms and enabling the clinician to evaluate whether symptoms increase or decrease when stressing or relaxing the nerves.

Author contributions

APV Conceptualization; APV Data curation; All Formal analysis; NA Funding acquisition; all Investigation; APV, DA Methodology; APV, DA, HB Project administration; APV Resources; APV, DA Software; APV, DA Supervision; APV, DA, MH Validation; all Writing; APV, DA, MH Review & editing

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgement

Georgia Bisset (GB), Christian Blanda (CB), Cassandra Armenio (CA), Helen Dickson (HD), Sarah Fensom (SF) for their assistance in the data extraction.

References

- Walsh MT. Upper limb neural tension testing and mobilization. Fact, fiction, and a practical approach. *J Hand Ther.* 2005;18(2):41–58.
- Physiopeedia: [https://www.physio-peedia.com/Upper_Limb_Tension_Tests_\(ULTTs\)](https://www.physio-peedia.com/Upper_Limb_Tension_Tests_(ULTTs)).
- Nee RJ, Jull GA, Vicenzino B, Coppieters MW. The validity of upper-limb neurodynamic tests for detecting peripheral neuropathic pain. *J Orthop Sports Phys Ther.* 2012;42(5):413–424.
- Lewis J, Ramot R, Green A. Changes in mechanical tension in the median nerve: possible implications for the upper limb tension test. *Physiotherapy.* 1998;84(6):254–261.
- Mahmud MA, Merlo AR, Gomes I, Becker J, Nora DB. [Relationship between adverse neural tension and nerve conduction studies in patients with symptoms of the carpal tunnel syndrome]. (*Portigues*) *Arq Neuropsiquiatr.* 2006;64(2A):277–282.
- Nee RJ, Yang C-H, Liang C-C, Tseng G-F, Coppieters MW. Impact of order of movement on nerve strain and longitudinal excursion: a biomechanical study with implications for neurodynamic test sequencing. *Man Ther.* 2010;15(4):376–381.
- Kleinrensink GJ, Stoeckart R, Mulder PG, et al. Upper limb tension tests as tools in the diagnosis of nerve and plexus lesions. Anatomical and biomechanical aspects. *Clin Biomech.* 2000;15:9–14.
- Manvell JJ, Manvell N, Snodgrass SJ, Reid SA. Improving the radial nerve neurodynamic test: an observation of tension of the radial, median and ulnar nerves during upper limb positioning. *Man Ther.* 2015;20(6):790–796.
- Kang K, Lee H, Lee J. Cervical radiculopathy focus on characteristics and differential diagnosis. *Asian Spine J.* 2020;14:921–930.
- Koulidis K, Veremis Y, Anderson C, Heneghan NR. Diagnostic accuracy of upper limb neurodynamic tests for the assessment of peripheral neuropathic pain: a systematic review. *Musculoskelet Sci Pract.* 2019;40:21–33.
- Thoomes EJ, van Geest S, van der Windt DA, et al. Value of physical tests in diagnosing cervical radiculopathy: a systematic review. *Spine J.* 2018;18(1):179–189.
- Physiotutors: <https://www.youtube.com/watch?v=rir6x6liqc4>.
- Clinical physio: https://www.youtube.com/watch?v=PQIc_etPXNI.
- Apelby-Albrecht M, Apelby-Albrecht M, Andersson L, et al. Concordance of upper limb neurodynamic tests with medical examination and magnetic resonance imaging in patients with cervical radiculopathy: a diagnostic cohort study. *J Manip Physiol Ther.* 2013;36(9):626–632.
- Davis DS, Anderson IB, Carson MG, Elkins CL, Stuckey LB. Upper limb neural tension and seated slump tests: the false positive rate among healthy young adults without cervical or lumbar symptoms. *J Man Manip Ther.* 2008;16(3):136–141.
- Kaizik MA, Hancock MJ, Herbert RD. A description of the primary studies of diagnostic test accuracy indexed on the DiTA database. *Physiother Res Int.* 2020;25(4):e1871.

17. Kaizik MA, Hancock MJ, Herbert RD. DiTA: a database of diagnostic test accuracy studies for physiotherapists. *J Physiother.* 2019;65(3):119–120.
18. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529–536.
19. Verhagen A, Hancock M. Research Note: diagnostic test accuracy studies. *J Physiother.* 2021;67(1):69–71.
20. Sackett DL, Straus S. On some clinical useful measures of the accuracy of diagnostic tests. *APC J Club.* 1998;129:A17–A19.
21. Hegedus EJ &, Stern B. Beyond SpIn and SnNOut: considerations with dichotomous tests during assessment of diagnostic accuracy. *JMPT.* 2009;17(1):E1–E5.
22. Pereson A;, della Torre E. Crainich U Efficacia diagnostica dei test provocativi nella radicolopatia cervicale: una meta-review (Efficacy of provocative tests in cervical radiculopathy: a meta-review) [Italian]. *Scienza Riabilitativa.* 2020;22(3):5–12.
23. Wainner RS, Gill H. Diagnosis and nonoperative management of cervical radiculopathy. *J Orthop Sports Phys Ther.* 2000;30(12):728–744.
24. Inal EE, Eser F, Aktekin LA, Oksüz E, Bodur H. Comparison of clinical and electrophysiological findings in patients with suspected radiculopathies. *J Back Musculoskelet Rehabil.* 2013;26(2):169–173. <https://doi.org/10.3233/BMR-2012-00364>.
25. Lauder TD, Dillingham TR, Andary M, et al. Predicting electrodiagnostic outcome in patients with upper limb symptoms: are the history and physical examination helpful? *Arch Phys Med Rehabil.* 2000;81(4):436–441. <https://doi.org/10.1053/mr.2000.4426>.
26. Shabat S, Leitner Y, David R, Folman Y. The correlation between Spurling test and imaging studies in detecting cervical radiculopathy. *J Neuroimaging.* 2012;22(4):375–378. <https://doi.org/10.1111/j.1552-6569.2011.00644.x>.
27. Trainor K, Pinnington MA. Reliability and diagnostic validity of the slump knee bend neurodynamic test for upper/mid lumbar nerve root compression: a pilot study. *Physiotherapy.* 2011;97(1):59–64. <https://doi.org/10.1016/j.physio.2010.05.004>.
28. Ghasemi M, Golabchi K, Mousavi SA, et al. The value of provocative tests in diagnosis of cervical radiculopathy. *J Res Med Sci.* 2013;18(Suppl 1):S35–S38.
29. Grondin F, Cook C, Hall T, Maillard O, Perdrix Y, Freppel S. Diagnostic accuracy of upper limb neurodynamic tests in the diagnosis of cervical radiculopathy. *Musculoskelet Sci Pract.* 2021;55:102427.
30. Sleijser-Koehorst MLS, Coppieters MW, Epping R, Rooker S, Verhagen AP, Scholten-Peeters GGM. Diagnostic accuracy of patient interview items and clinical tests for cervical radiculopathy. *Physiotherapy.* 2021;111:74–82.
31. Bueno-Gracia E, Tricás-Moreno JM, Fanlo-Mazas P, et al. Validity of the upper limb neurodynamic test 1 for the diagnosis of Carpal Tunnel Syndrome. The role of structural differentiation. *Man Ther.* 2016;22:190–195.
32. Trillos MC, Soto F, Briceno-Ayala L. Upper limb neurodynamic test 1 in patients with clinical diagnosis of carpal tunnel syndrome: a diagnostic accuracy study. *J Hand Ther.* 2018;31(3):333–338.
33. Vanti C, Bonfiglioli R, Calabrese M, et al. Upper Limb Neurodynamic Test 1 and symptoms reproduction in carpal tunnel syndrome. A validity study. *Man Ther.* 2011;16(3):258–263.
34. Vanti C, Bonfiglioli R, Calabrese M, Marinelli F, Violante FS, Pillastrini P. Relationship between interpretation and accuracy of the upper limb neurodynamic test 1 in carpal tunnel syndrome. *J Manip Physiol Ther.* 2012;35(1):54–63.
35. Wainner RS, Fritz JM, Irrgang JJ, Boninger ML, Delitto A, Allison S. Reliability and diagnostic accuracy of the clinical examination and patient self-report measures for cervical radiculopathy. *Spine.* 2003;28(1):52–62.
36. Wainner RS, Fritz JM, Irrgang JJ, Delitto A, Allison S, Boninger ML. Development of a clinical prediction rule for the diagnosis of carpal tunnel syndrome. *Arch Phys Med Rehabil.* 2005;86(4):609–618.
37. Shacklock M. *Clinical Neurodynamics: a New System of Musculoskeletal Treatment.* Oxford: Elsevier; 2005.
38. DiTA: <https://dita.org.au>.
39. Leeflang MMG, Bossuyt PMM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol.* 2009;62(1):5–12.