Eur Spine J (2007) 16:307–319
DOI 10.1007/s00586-006-0225-6

REVIEW

A systematic review of the diagnostic accuracy of provocative tests of the neck for diagnosing cervical radiculopathy

Sidney M. Rubinstein · Jan J. M. Pool · Maurits W. van Tulder · Ingrid I. Riphagen · Henrica C. W. de Vet

Received: 23 April 2006 / Revised: 3 August 2006 / Accepted: 5 September 2006 / Published online: 30 September 2006

Abstract Clinical provocative tests of the neck, which position the neck and arm inorder to aggravate or relieve arm symptoms, are commonly used in clinical practice in patients with a suspected cervical radiculopathy. Their diagnostic accuracy, however, has never been examined in a systematic review. A comprehensive search was conducted in order to identify all possible studies fulfilling the inclusion criteria. A study was included if: (1) any provocative test of the neck for diagnosing cervical radiculopathy was identified; (2) any reference standard was used; (3) sensitivity and specificity were reported or could be (re-)calculated; and, (4) the publication was a full report. Two reviewers independently selected studies, and assessed methodological quality. Only six studies met the inclusion criteria, which evaluated five provocative tests. In general, Spurling's test demonstrated low to moderate sensitivity and high specificity, as did traction/neck distraction, and Valsalva's maneuver. The upper limb tension test (ULTT) demonstrated high sensitivity and low specificity, while the shoulder abduction test demonstrated low to moderate sensitivity and moderate to high specificity. Common methodological flaws included lack of an optimal reference standard, disease progression bias, spectrum bias, and review bias. Limitations include few primary studies, substantial heterogeneity, and numerous methodological flaws among the studies; therefore, a meta-analysis was not conducted. This review suggests that, when consistent with the history and other physical findings, a positive Spurling's, traction/neck distraction, and Valsalva's might be indicative of a cervical radiculopathy, while a negative ULTT might be used to rule it out. However, the lack of evidence precludes any firm conclusions regarding their diagnostic value, especially when used in primary care. More high quality studies are necessary in order to resolve this issue.

Keywords Sensitivity · Specificity · Diagnostic accuracy · Systematic review · Cervical radiculopathy · Upper limb tension test · Spurling's test · Shoulder abduction test

S. M. Rubinstein (✉) · J. J. M. Pool · M. W. van Tulder · H. C. W. de Vet
Institute for Research in Extramural Medicine (EMGO), VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands
e-mail: SM.Rubinstein@vumc.nl

M. W. van Tulder
Institute for Health Sciences, Faculty of Earth and Life Sciences, Amsterdam, The Netherlands

I. I. Riphagen
Library, VU University Medical Center, Amsterdam, The Netherlands

Introduction

Cervical radiculopathy is a substantial cause of disability and morbidity [7], and is a common condition, affecting both sexes after middle age [31]. Cervical radiculopathy refers to those subjects with signs and symptoms related to dysfunction of the spinal nerve root(s) of the neck [43]. The diagnostic criteria are, however, unclear [7, 43]. Some suggest that cervical radiculopathy is a diagnosis based upon clinical impression [5, 9, 23, 57], which should be confirmed by

advanced testing, such as diagnostic imaging [12, 15, 34, 50], or electrophysiology studies [22, 27, 40]. However, clinical and radiological diagnoses, and electrophysiological testing all have inherent limitations [59].

On the one hand, electrodiagnostic (EDX) testing might appear to be the most suitable reference standard. A recent, comprehensive review determined that needle EMG has a sensitivity of 50–71% for subjects with neurological or radiological signs of a cervical radiculopathy, although the lack of a standardized gold standard may have resulted in an underestimation of this figure [7, 8]. On the other hand, asymptomatic radiological abnormalities are commonly seen with advanced imaging studies [17, 20, 26, 37, 58]. It is well accepted that the diagnostic accuracy of specialized imaging is limited, especially with regards to foraminal nerve root impingement [11, 15, 38]. Unlike EDX testing, imaging cannot distinguish compressive from non-compressive etiologies, such as inflammation. Consequently, nerve root pain can be present in the absence of visible compression [45].

Diagnostic accuracy for neuropathies or radiculopathies might, therefore, be improved when the results of needle EMG are combined with imaging studies and clinical findings [14, 30, 39, 45, 52, 53]. Therefore, for the purpose of this paper, we propose that the optimal reference standard for cervical radiculopathy be based upon both (a) electromyographic evidence of acute denervation in cervical paraspinal muscles and/or in a myotome; and (b) demonstrate abnormalities on advanced imaging studies, such as myelography, CT-myelography, or MRI, which also correspond to the site of signs and symptoms consistent with cervical radiculopathy [43].

Advanced diagnostic testing, however, can be expensive, and in the case of nerve conduction studies, may be intrusive or painful. Clearly, there is a need for a cost-effective, accurate and non-invasive manner for the primary care provider to confirm his diagnostic impression, and to determine whether the patient can be adequately treated in the primary care setting.

To that end, various clinical provocative tests have been proposed which purport to be diagnostic of cervical radiculopathy [4, 42, 46, 59, 62, 63]. Examples of these tests include, but are not limited to the upper limb tension test [ULTT] (formerly called the brachial plexus tension test, or test of Elvey) [3, 24, 32, 42, 44, 49, 63], the shoulder abduction test (also known as the shoulder abduction relief sign) [21, 25], and Spurling's test (also known as the foraminal compression test, neck compression test, or quadrant test) [36, 54, 59, 60]. These tests are not meant to supplant the neurological examination, nor to ignore those subjects with clear neurological deficits. Clinically, these tests are most valuable when the neurological exam is inconclusive, yet the subject has symptoms consistent with a radiculopathy. Therefore, these tests might not only help to confirm the diagnosis when the presenting clinical picture and examination is unclear, but may also help to establish a prognosis, and assist in triage when the clinician is unsure which therapy may be prescribed, or whether the patient should be referred for further diagnostic testing. In order to assess the diagnostic accuracy of these tests, we conducted a systematic review.

Materials and methods

We first searched various orthopaedic textbooks in order to identify all provocative tests suggestive of cervical radiculopathy [1–3, 13, 16]. In addition to the afore-mentioned tests, we identified the following specific-named tests: axial compression test, distraction test, shoulder depression test, and Valsalva's manoeuvre/test.

Study identification

Subsequently, we conducted a comprehensive search in the following databases: MEDLINE, EMBASE, CINAHL, Medion, OSTMED, and the database of reviews of effectiveness (DARE), in order to identify relevant studies. The electronic search was conducted in June 2005, and developed in collaboration with an experienced librarian (Ingrid I. Riphagen). We chose not to use a filter, such as that outlined by Bachmann et al. [10], because an earlier preliminary search had suggested that there were few studies in this area, and we did not want to miss any publications. Therefore, we elected for a rather broad search (Appendix 1).

No restrictions were applied to year of publication or language. All titles and abstracts were examined that met our search terms and full publications were reviewed, when necessary. Additionally, the reference sections of all primary studies were inspected. Finally, authors of the primary studies were contacted. Four of the six authors responded, although this did not yield any additional studies.

Study selection

Two reviewers (SMR, JJMP), independently, screened the titles and abstracts of the citations identified in the electronic searches using the following criteria. A study was included if: (1) any provocative test of the neck for

diagnosing cervical radiculopathy was identified; (2) the diagnostic test was compared to any reference standard, such as, electromyography (EMG), plain film, or advanced imaging (e.g., MRI, CT, myelography); (3) sensitivity and specificity were reported or a 2×2 contingency table could be (re-)constructed; and (4) the publication was a full report. Case series and case reports, as well as animal, surgical, and cadaveric studies were excluded because these studies cannot evaluate diagnostic accuracy for subjects with neck and radiating pain. Justifications for excluding studies were also noted and discrepancies were resolved by a third reviewer (HCWdeV), where necessary.

Assessment of methodological quality

The methodological quality of all studies was evaluated by QUADAS, which is a tested tool (Table 1; Appendix 2 contains an operationalization of the items) [65]. These items cover the most significant forms of bias in diagnostic research, such as, spectrum bias, disease progression bias, verification bias, review bias, and the potential bias associated with subject withdrawal, as well as aspects of external validity. Two items were dropped concerning partial verification bias and incorporation bias because it was felt that they would not be relevant items in this study.

Pilot study

In order to improve agreement between the two reviewers, the aforementioned assessment tool was piloted using a few studies which have examined the diagnostic accuracy of Lasegue's test for diagnosing herniated discs of the low-back [6, 29, 33].

Scoring

Two reviewers (SMR, JJMP), each blind to the others assessment, scored the criteria items: ''positive'' or ''negative'' when studies satisfied or failed to meet the criteria, respectively, and ''unclear'' when an item was inadequately described. The reviewers then met to discuss differences in coding during a consensus meeting. Agreement between reviewers was quantified by an unweighted Cohen's kappa (k) and we classified the categories positive, unclear, and negative as ordinal categories. The program from the website: http://www.faculty.vassar.edu/lowry/kappa.html was used. Disagreement was also calculated and expressed as a percentage (the number of disagreed upon criteria/the total number of criteria).

Data extraction

The same two reviewers performed data extraction independently, using a pre-determined, self-developed data extraction form. The following data were extracted: author, year of publication, country where the study was performed, characteristics of the study population, the test(s) examined, and results.

Data analysis

Sensitivities and specificities were extracted, where possible. In two studies [21, 62], it was necessary to re-calculate these figures using the raw numbers presented. Diagnostic odds-ratio's (DOR's) were not calculated because of the recent suggestion that a DOR does not meaningfully describe a marker's ability to classify subjects [41]. A meta-analysis was also not

Table 1 Methodological criteria used to assess the quality of studies investigating diagnostic accuracy for subjects with cervical radiculopathy

Item
1. Was the spectrum of subjects used in the study, representative of the patients who will receive the index test in clinical practice?
2. Were the in- and exclusion criteria of the subjects clearly described?
3. Was the reference standard used likely to correctly classify the target condition (i.e. radiculopathy)?
4. Was the time between the application of the reference standard and index test short enough to be reasonably sure that the disease status of the target condition did not change between administration of the two tests?
5. Did all patients receive the same reference standard, regardless of the index test results?
6. Was the execution of the index test described in sufficient detail to permit replication of the test?
7. Was the execution of the reference standard described in sufficient detail to permit its replication?
8. Were the index test results interpreted without knowledge of the results of the reference standard?
9. Were the reference standard results interpreted without knowledge of the results of the index test?
10. Were the same clinical data available when the index test results were interpreted as would be available when the index test is applied in clinical practice?
11. Were uninterpretable/intermediate/unclear index test results reported?
12. Were withdrawals from the study explained or reported?

Adapted from Whiting et al. [65]

conducted because too few studies were identified for individual tests, and there was too much clinical heterogeneity among these studies. Instead, we chose to provide a qualitative descriptive analysis.

Results

Data synthesis

The first database that was consulted (Embase.com), which includes MEDLINE, identified 366 potentially relevant articles. We excluded 354 studies on the basis of their titles and abstracts. We then retrieved and reviewed 12 full reports for possible inclusion [4, 19, 21, 35, 42, 48, 51, 55, 59, 60, 62, 63]. Four studies were excluded because they examined either reliability or repeatability of the test, but not diagnostic accuracy [4, 19, 35, 48]. Two studies were excluded because the tests were used to evaluate clinical conditions other than radiculopathy [55, 60]. This resulted in six studies which fulfilled our inclusion criteria [21, 42, 51, 59, 62, 63], all of which were found in MEDLINE. Subsequently, we searched PubMed and CINAHL for additional relevant articles, which resulted in 218 and 66 additional references, respectively, not found in the Embase.com search. This, however, did not yield any additional studies. A hand search of the references of the primary studies was also conducted, but also did not yield additional studies, as did searches in OSTMED, Medion, and DARE. Finally, a search in PubMed using the specific names of the provocative tests was conducted, which also failed to yield additional studies. Note: a flow chart which depicts the electronic databases used and where the studies were found are available from the primary author Sidney M. Rubinstein upon request.

Study characteristics

Study characteristics are presented in Table 2. Multiple studies evaluated the following provocative tests: ULTT [42, 63], Spurling's test [51, 59, 62, 63], shoulder abduction test [21, 62, 63], traction/neck distraction [63, 62], while only a single study was found for the Valsalva's manoeuvre [63]. No studies were identified which examined the axial compression test or shoulder depression test. (Note: for a description of these tests, we refer the reader to Wainner et al. [63].)

Three studies were conducted in the 1980s [21, 42, 62], while the remaining three studies were published after 2000 [51, 59, 63]. Some studies examined the accuracy of multiple provocative tests of the neck [62, 63]. Two studies performed the index test to both the symptomatic and asymptomatic side, presumably to serve as an internal control [42, 62], and one of these studies included a control group [42]. The remaining studies used a cross-sectional design. Despite the large cohorts from which subjects were recruited in some studies, the number of diseased subjects was relatively small for all studies (mean = 20, range: 18 [42]–29 [51]).

All, but one study [62], recruited subjects with both neck and radicular-type pain. Patients were recruited from neurosurgical and orthopaedic departments in three studies [21, 51, 62], while in two other studies, subjects were recruited who were referred for electrodiagnostic testing [59, 63]. This might suggest spectrum bias because the severity of the condition is likely to have influenced the decision to see a specialist or to be referred for advanced diagnostic testing. In the remaining study, patients were recruited from a physiatrist practice [42].

Methodological quality assessment

Results of the methodological quality assessment are presented in Table 3. No single study used the optimal reference standard, which consists of both EDX testing and advanced imaging. Two studies used EMG testing as the reference standard [59, 63], three studies used advanced imaging (MRI or myelography) [21, 51, 62], while one study used plain-film radiography [42]. One study used operative findings as a reference standard in a subset of subjects who failed conservative therapy [51]. Two studies presented neurological exam findings for individual patients, and was incorporated into the reference standard in one study [62], while in another study it was unclear how the neurological findings were used [21].

Other forms of bias found in the studies included: disease progression bias (n = 5/6 or 83% of the studies, criterion 4) [21, 42, 51, 59, 62], spectrum bias (n = 5/6 or 83% of the studies, criterion 1) [21, 51, 59, 62, 63], review bias (equivalent to blinded assessment for intervention studies) for the reference standard (n = 3/6 or 50% of the studies, criterion 9) [21, 42, 59], review bias for the index test (n = 2/6 or 33% of the studies, criterion 8) [21, 42], and bias associated with drop-out/withdrawals (n = 3/6 or 50% of the studies, criterion 12) [42, 59, 62]. In one study, clinicians were inexplicably blinded to relevant historical patient information [62]. Despite these shortcomings to internal validity, all studies generally scored well with external validity (criteria 2, 6, 7).

Table 2 Study characteristics for the individual studies which investigated diagnostic accuracy of provocative tests for subjects with cervical radiculopathy

Author, reference, country	Study subjects sample size (N = all subjects) cases[a]: age [mean (SD)], sex (% female (F))	Setting; period of recruitment	Inclusion criteria	Exclusion criteria	Duration of symptoms, (range in months)
Davidson [21], USA	N = 200 cases: n = 15 45 years (range: 31–67) 27% F	Neurosurgical and orthopaedic department; 1978–1980	1. Neck and radicular arm pain 2. Failure with conservative care, consisting of rest, heat, support, analgesia, muscle relaxants, and in most cases, traction	?	?
Quintner [42] Australia	Cases = 37 34 years (range: 12–63) 84%F Controls = 20 42 years (range: 14–75) 55%F	Clinical practice; Jan 1986–Dec 1987	1. Persistent neck and arm pain &/or paresthesias 2. Motor vehicle accident 1. Full and pain-free cervical movement in all directions 2. Full range of passive movement at the shoulder, elbow, wrist	Neck pain prior to motor vehicle accident ?	4–76 months (mean 28 months)
Shah [51], India	Cases = 50 42 years (range: 22–60) 62%F;	Neurosurgical department; May 2001–May 2003	Neck and radicular pain	1. Associated myelopathy 2. Prior surgery on cervical spine 3. History of trauma	0.5–36 months Soft disc prolapse: 6.6 months presence of osteophyte: 8.1 months
Tong [59], USA	N = 255 cases = 192 age= ? sex = ?	Academic general physiatrist practice; 1988–1993	Subjects referred for electrodiagnostic testing for upper extremity nerve disorders	?	?
Viikari-Juntura [62], Finland	N = 69 cases = 43 median age: 52 years (range: 33–80) 37%F	Neurosurgical department; March 1982–Jan 1985	Cervical disc disease (spondylosis &/or cervical disc herniation)	1. Neurological tumors 2. Nonexpansive neurologic diseases 3. "Miscellaneous states" 4. Rheumatoid arthritis 5. Cervical spine cancer 6. Cervical spine malformations	?
Wainner [63], USA	N = 82 cases = 19 mean age 45(12) years 21%F	Four medical facilities, including medical centers, hospitals; Dec 1998–April 2000	Subjects referred from electrophysiologic laboratory with signs/symptoms of CR or CTS	1. Systemic disease causing peripheral neuropathy 2. Bilateral radiating pain 3. Conditions of the upper extremity affecting function 4. > 6 months stopped with work due to the condition 5. History of surgical procedures for pathologies causing neck pain or CTS 6. Previous EMG or NCS for radiculopathy or CTS 7. Workers' compensation or pending litigation for the presenting condition	Radiculopathy only: 1.4–36.5 Radiculopathy w/CTS: 0.7–3.3

CR cervical radiculopathy, *CTS* carpal tunnel syndrome, *EMG* electromyography, *NCS* nerve conduction studies, *?* unclear or not defined

[a]Number of subjects actually undergoing the index test

Table 3 Methodological quality assessment of studies which investigated the diagnostic accuracy of provocative tests for subjects with cervical radiculopathy

Author, reference[a]	Criteria number											
	1	2	3	4	5	6	7	8	9	10	11	12
Davidson [21]	–	–	–	?	+	+	+	?	?	+	?	+
Quintner [42]	+	+	–	?	+	+	+	–	?	+	+	–
Shah [51]	–	+	–	?	–	+	+	+	+	+	?	+
Tong [59]	–	–	–	?	+	+	+	+	?	+	?	–
Viikari-Juntura [62]	–	+	–	?	–	+	+	+	+	–	?	?
Wainner [63]	–	+	–	+	+	+	+	+	+	+	?	+
% of maximum	17	67	0	17	67	100	100	67	50	83	17	50

Items were scored as follows: + means adequate methods,– means inadequate methods, and *?* means an item was inadequately described and therefore, a decision couldnot be made whether it satisfied the criteria or not

[a]Publications are listed alphabetically by author

Agreement between assessors

Agreement between the two reviewers was high [unweighted $k = 0.86$ (95% confidence interval 0.76–0.97)], as is also reflected in the low disagreement for all the criteria scored (6/72 = 8% disagreed upon items). Disagreements appeared quite randomly spread among the criteria and were principally due to either reading errors or differences in interpretation. All disagreements were resolved during the consensus meeting, without intervention from an arbiter.

Diagnostic accuracy of the tests

Diagnostic accuracy parameters for the six primary studies are presented in Table 4. Most striking was the considerable variability between results for the various studies. This was most remarkable for the shoulder abduction test, which reported sensitivities ranging from 0.17 to 0.78 [21, 62, 63]. This was most likely due to the great variability in quality between these three studies, and choice of the reference standard. In general, the four studies which investigated Spurling's test [51, 59, 62, 63] demonstrated low to moderate sensitivity and high specificity, as did individual studies for traction/neck distraction [62, 63], and the Valsalva's manoeuvre [63]. On the other hand, the two studies which investigated the ULTT [42, 63] demonstrated high sensitivity and low specificity, while the three studies for the shoulder abduction test [21, 62, 63] demonstrated low to moderate sensitivity and moderate to high specificity.

The reference standard

Studies that used imaging as the reference standard [21, 42, 51, 62] most probably overestimated sensitivity and underestimated specificity due to its false positive rate, or in other words, overestimated the number of diseased subjects. On the other hand, studies which used EDX testing [59, 63] would have resulted in an underestimation of sensitivity and overestimation of specificity due to its false negative rate, and therefore, underestimated the number of diseased subjects. Thus, while some tests demonstrated high specificity (Spurling's, traction/neck distraction, Valsalva's), and the ULTT demonstrated high sensitivity, no provocative test demonstrated both components. In general, the methodological quality of all studies, with the exclusion of Wainner et al. [63], was rather meagre.

It should also be noted that there was considerable variability across the studies in performance of the tests. This lack of standardization has been noted elsewhere [36]. For example, of the four studies which examined Spurling's test [51, 59, 62, 63], no two studies performed Spurling's in exactly the same manner.

Reporting of diagnostic accuracy parameters

In two older studies, diagnostic accuracy was either not presented [21] or was unclear [62]. In these studies, neurological exam findings were also presented for individual subjects, so we reconstructed a 2×2 table using the diagnostic criteria of Radhakrishanan et al. [43] as the reference standard. This criterion combines symptoms of pain and/or weakness, neurological signs, and either advanced diagnostic test results (i.e. EMG or imaging) or surgical verification, for the determination of cervical radiculopathy. While this diagnostic criteria has obvious limitations, it was thought to be more accurate than using the results of myelography alone [43]. In the remaining studies, the neurological examination for

individual subjects was not presented, so we were not able to compare the reported accuracy parameters to potentially other calculated values using the aforementioned criteria.

In another study [51], we found it necessary to re-calculate sensitivity and specificity because the categorization used by the authors was thought to be incorrect. While the authors considered a soft disc prolapse on MRI a positive finding and therefore, confirmation of a radiculopathy, they considered an osteophyte a negative finding. On further study, however, the root canal diameters (i.e. lateral recess) were found to be significantly smaller on the symptomatic side (in the presence of either a disc prolapse or osteophyte) than the asymptomatic side, while the lateral recess on the asymptomatic side was equal in size in asymptomatic subjects. This would suggest that an osteophyte in this case series, could have equally caused nerve root impingement, and therefore, should have also been considered a positive finding. A recalculation for the Spurling's test based upon the aforementioned information, yielded a sensitivity of 52.9% and specificity of 93.8%, while the combined results of the MRI findings and operative findings yielded a sensitivity of 73%, and specificity of 92.3%. These recalculated sensitivities are much less than those presented by the authors, and we believe, a more accurate assessment (see Table 4).

Discussion

We conducted this systematic review in order to determine the diagnostic accuracy of provocative tests for subjects with a suspected cervical radiculopathy. Specifically, we wanted to evaluate the diagnostic value of these tests in the primary care setting, and to determine whether these tests could help the family physician, chiropractor, or manual/physical therapist to confirm their diagnosis, without having to resort to advanced diagnostic testing. Unfortunately, our goal was limited by three major obstacles: (1) only six studies were identified, only one of which evaluated subjects in the primary care setting; (2) no study used an optimal reference standard; and (3) the lack of test standardization (or performance), specifically Spurling's test, made it impossible to compare sensitivity and specificity across studies.

It is remarkable that, despite their wide use in clinical practice, the diagnostic accuracy of these tests have received so little attention. It is all the more interesting that various orthopaedic textbooks list these provocative tests as indicative or suggestive of a cervical radiculopathy [1–3, 13, 16, 56, 64], while there is so little evidence to support their accuracy. There are inherently two problems with trying to determine the accuracy of these tests for radiculopathy: (1) there are no universally accepted diagnostic criteria, and (2) it is not clear what a positive test actually measures. It is likely that these tests are not simply a measure of compression or traction of the nerve root (or dural sleeve). The problem may also lie with distinguishing nerve root pain from brachial plexus pathology [47].

Perhaps the most significant methodological consideration we encountered was the choice of reference standard. As noted above, there is no gold standard for the diagnosis of cervical radiculopathy, as clinical radiological, and electrophysiological testing all have inherent limitations [7]. The problem essentially stems from the false positive rate associated with imaging and the false negative rate associated with electrophysiological testing. This is no minor issue, given that the establishment of a gold standard is pivotal to the study of diagnostic accuracy. In the absence of consensus, we determined that the optimal gold standard should combine the findings of MRI with EDX testing. However, no study that we reviewed performed this. Quite clearly, future studies on this topic must somehow resolve this issue. Although some consider the history and physical exam to be sufficient in many cases to make a clinical diagnosis [18, 28], perhaps the best practical solution to this problem might be to accept the criteria established by Radhakrishnan et al. [43], which combines the physical examination findings with advanced testing or surgical verification.

Other important methodological problems include spectrum bias. Because the suspicion of a radiculopathy was the reason for referral in all but one article [21, 42, 51, 59, 63], and hence, the reason for inclusion in the study, subjects in those studies were more likely to have a worse clinical presentation than those presenting in primary care. Patients would have been most likely referred to the neurologist or for EDX testing because of a strong clinical suspicion of a cervical radiculopathy, or the lack of response to conservative care. Subjects are, therefore, more likely to have been accurately classified according to these tests than had subjects also been recruited with pseudo-radicular pain or pain from the brachial plexus or a peripheral nerve. This would have resulted in an overestimation of sensitivity and an underestimation of the specificity. In the remaining study, while radicular-type pain was not necessarily an inclusion criteria, subjects were recruited from a

Table 4 Diagnostic parameters for studies which investigated diagnostic accuracy of provocative tests for subjects with cervical radiculopathy

Type of index test	Reference standard	Diseased			
Author, reference		+	–	Sensitivity (95% CI)	Specificity (95% CI)
Shoulder abduction test					
Davidson [21]	Myelography[a]	18	4	0.78 (0.52–0.94)	0.75 (0.19–0.99)
Viikari-Juntura [62]	Myelography[a]	13	13	0.46 (0.19–0.75)	0.85 (0.55–0.98)
Wainner [63]	Needle EMG/NCS	19	63	0.17 (0.0–0.34)	0.92 (0.85–0.99)
Spurling's—combined with neck extension					
Shah [51] (no rotation)	Operation (group 1)	20	5	0.90 (0.68–0.99)	1.00 (0.48–1.00)
Ibid	MRI, no operation (group 2)	9	16	1.00 (0.66–1.00)[b]	0.94 (0.70–1.00)[b]
Ibid	Combined MRI + operative findings	29	21	0.93 (0.77–0.99)	0.95 (0.762–1.0)
Tong [59] (with contralateral rotation)	EMG	20	172	0.30 (0.12–0.54)	0.93 (0.88–0.96)
Wainner [63]; (Test part B)[c] (with ipsilateral rotation)	Needle EMG/NCS	19	63	0.50 (0.27–0.73)	0.74 (0.63–0.85)
Spurling's—without neck extension					
Viikari-Juntura [62] (with contralateral rotation)	Myelography[a]	22	56	0.50 (0.28–0.72)	0.93 (0.83–0.98)
Wainner [63]; (Test part A)[c] (no cervical rotation)	Needle EMG/NCS	19	63	0.50 (0.27–0.73)	0.86 (0.77–0.94)
Upper limb tension test (ULTT)					
Quintner [42]	Plain-film radiography of the cervical spine	18	27	0.83 (0.59–0.96)	0.11 (0.02–0.29)
Wainner [63]; (Test part A)[c]	Needle EMG/NCS	19	63	0.97 (0.90-1.0)	0.22 (0.12-0.33)
Wainner [63]; (Test part B)[c]	Needle EMG/NCS	19	63	0.72 (0.52–0.93)	0.33 (0.21–0.45)
Traction/neck distraction test					
Viikari-Juntura [62]	Myelography[a]	9	35	0.44 (0.14–0.79)	0.97 (0.85–1.0)
Wainner [63]	Needle EMG/NCS	19	63	0.44 (0.21–0.67)	0.90 (0.82–0.98)
Valsalva manoeuvre					
Wainner [63]	Needle EMG/NCS	19	63	0.22 (0.03–0.41)	0.94 (0.88–1.0)

EMG electromyography, *NCS* nerve conduction study

[a]Confirmation of the diagnosis was combined with the presence (or absence) of neurological signs according to the criteria of Radhakrishnan et al. further defined in the text

[b]Other scores were calculated by the authors of this review using the data presented in the primary study and is explained in discussion section;

[c]The actual testing procedures for the different provocative tests are summarized and defined in detail by Wainner et al. [63] in his appendix, p. 62. Wainner et al. [63] distinguished distinct aspects of the Spurling's test and the ULTT, and operationalized them as parts A & B. For Spurling's, part A was performed as originally described by Spurling and Scoville [54], while part B represents variations that have been reported, such as the addition of rotation, or rotation and extension [61]. For the ULTT, part A was performed as described by Elvey [24], while part B includes wrist and finger flexion, instead of extension

neurosurgical department referred for cervical myelography, so presumably, these subjects had a worse clinical presentation than the typical patient with spondylosis [62].

Finally, we identified only one study which recruited subjects from the primary care setting [42]. This is quite remarkable because this is the setting in which the test is most likely to be used. This study, however, scored poorly in many of the items of internal validity, and used a completely inadequate reference standard (i.e. plain-film radiography). This means that no study has satisfactorily examined diagnostic accuracy of these tests in primary care.

Conclusions

Provocative tests of the neck for those with a suspected cervical radiculopathy might help to establish the diagnosis, especially in those subjects lacking well-defined neurological deficits. When consistent with the history and other physical findings, a positive Spurling's test, as well as positive findings for traction/neck distraction, and the Valsalva's manoeuvre might be suggestive of a cervical radiculopathy (i.e. given their high specificity), while a negative ULTT might be used to rule it out (i.e. given its high sensitivity). However, the lack of primary studies investigating the

accuracy of these tests, as well as heterogeneity between the various studies, and numerous methodological problems, preclude any strong recommendations for the use of these tests, especially in the primary care setting. Therefore, in absence of other clinical information or corroborating evidence, the value of these tests should be interpreted with caution. Future diagnostic studies should include sufficient diseased subjects, and a composite reference standard, consisting of both advanced imaging and electrodiagnostic testing (or consider the diagnostic criteria of Radhakrishnan et al.), in order to ensure correct classification of cervical radiculopathy.

Key points

- A systematic review was conducted which examined the diagnostic accuracy of five provocative tests of the neck for subjects with cervical radiculopathy.
- Only six studies, with a relatively small number of diseased subjects, met the inclusion criteria. Only one study was considered to have a lower risk of biased results.
- While the Spurling's test, traction/neck distraction, and Valsalva's manoeuvre demonstrated high specificity, and the ULTT demonstrated high sensitivity, no provocative test demonstrated both high sensitivity and high specificity.
- Only one study evaluated these tests in primary care, which is the setting in which these tests are most likely to be used. However, this study was of poor quality.
- More high quality studies are necessary to determine the diagnostic accuracy of these tests for cervical radiculopathy, especially when used in the primary care setting.

Acknowledgments The principal author Sidney M. Rubinstein would like to thank Dr. Scott Haldeman and Dr. Tammy de Koekkoek for their valuable contributions during the preparation of this manuscript. This study has been funded, in part, by The Netherlands Chiropractic Association and The Netherlands Association for Manual Therapists.

Appendix 1

Search strategy used to identify studies on diagnostic accuracy of provocative tests for cervical radiculopathy.

Medion Database

Search conducted on: June 7, 2005

radiculopath* OR radiculit* OR monoradiculopath* OR polyradiculopath* OR cervical OR neck > No reviews were identified.

Dare

Search conducted on: June 7, 2005

1. radiculopath* OR radiculit* OR monoradiculopath* OR polyradiculopath*
2. (cervical OR cervico* OR neck) AND (root* OR radical) AND (nerve* OR spine OR spinal OR vertebr*)

>No reviews were identified.

Ostmed

Search conducted on: June 7, 2005

1. radiculopath* OR radiculit* OR monoradiculopath* OR polyradiculopath*
2. (cervical OR cervico* OR neck) AND (root* OR radical) AND (nerve* OR spine OR spinal OR vertebr*) AND (pain* OR complain* OR compression*) AND (diagnos* OR test* OR screen* OR examinat*)

Five studies were originally identified for possible inclusion, but were found not to meet the inclusion criteria:

1. Gifford L (2001) Acute low cervical nerve root conditions: symptom presentations and pathobiological reasoning. J Osteopath Med 4(2):69
2. Gifford L (2001) Acute low cervical nerve root conditions: symptom presentations and pathobiological reasoning. Man Ther 6(2):106–115
3. Biondi DM Cervicogenic headache: mechanisms, evaluation, and treatment strategies. J Am Osteopath Assoc 100(Suppl 9):S7–S14; quiz S27
4. Stoll ST, Caffrey JL, and Wright TJ (1999) Dermatomal somatosensory evoked potentials: evaluation of manipulative medicine in the treatment of cervical and lumbar radicular symptoms. J Am Osteopath Assoc 99(8):428
5. Gunn CC and Milbrandt WE (1977) Tenderness at motor points: an aid in the diagnosis of pain in the

shoulder referred from the cervical spine. J Am Osteopath Assoc 77(3):196–212; quiz 179–180

Cinahl

Search conducted on: June 7, 2005

1982-May, week 4 2005

109 hits > after eliminating duplicates also found in Embase.com, there remained 66 references over.

Embase.com (Medline and Embase combined) search

Search conducted on: April 6, 2005

#1

(Radiculopath* OR Radiculopathy/exp OR radiculit* OR ((spinal-root OR Spinal-root/exp OR nerve-root OR nerve-roots OR radicular OR brachial-plexus) AND (pain* OR complain* OR nerve-root-compression/exp)) OR monoradiculopath* OR polyradiculopath*) AND (cervical OR cervico* OR Cervical-spine/de OR neck OR Neck/exp OR Neck-pain/de OR Neck-injury/exp)

#2

(Tension OR abduction OR reflex OR compression OR traction OR retraction OR depression OR distraction OR elvey* OR spurling* OR orthopedic OR orthopaedic) AND test*

#3

Clinical-examination OR clinical-test OR clinical-tests OR neurologic-examination OR neurologic-examination/de OR Physical-examination/exp OR (physical AND examin*) OR Exercise-test/de

#1 AND (#2 OR #3) > Result 366 hits (311 references were found in Embase only)

PubMed search limited to MEDLINE:

Search conducted on: June 8, 2005

#1

(Radiculopath* OR Polyradiculopathy[mesh] OR radiculit* OR ((spinal root OR Spinal nerve roots[mesh] OR nerve root OR nerve roots OR radicular OR brachial plexus) AND (pain* OR complain* OR nerve compression syndromes[mesh])) OR monoradiculopath* OR polyradiculopath*) AND (cervical OR cervico* OR Cervical vertebrae[mesh] OR neck OR Neck[mesh] OR Neck pain[mesh] OR Neck injuries[mesh] OR Cervical plexus[mesh])

#2

(Tension OR abduction OR reflex OR compression OR traction OR retraction OR depression OR distraction OR elvey* OR spurling* OR orthopedic OR orthopaedic) AND (test OR tests OR tested OR testing)

#3

Clinical examination* OR clinical test OR clinical tests OR neurologic examination* OR Physical examination[mesh] OR (physical AND examin*) OR Diagnostic tests, routine[mesh] OR Exercise-test [mesh]

#1 AND (#2 OR #3) NOT (animals[mesh] NOT humans[mesh]) result 413 hits > after eliminating duplicates from Embase.com en Cinahl: 218 references remaining.

Note: In some cases it would appear that capital and small letters are randomly listed in the search profile, above. Capital letters are official keywords generated in Embase.com; however, this does not have any consequences for the search because the search engines are not sensitive to letter size. In some cases, words are connected to one another by lines (i.e. ''-'') in order to indicate that the entire phrase should be located, and not just separate terms. Also it is important to note the difference between ''/exp'' and ''/de'': /exp searches for other terms 'coupled' on the given term (i.e. not just the key word listed, but other terms using that key word), while /de searches only the key word used (i.e. the specific keyword). ''Exp'' is an abbreviation for 'explode', while ''de'' means 'descriptor'.

Appendix 2

Explanation of items and operationalization of terms used in this review.

General

Diagnostic test or index test = the test under examination.

Reference standard = gold standard (i.e. the test performed for which the diagnostic/index test is to be compared).

Criteria number. Definition of terms

1. Positive if the spectrum of patients included in the study was representative of those for whom the test will be used in primary care/clinical practice. Subjects must have the following characteristics:

a. Neck- and radiating pain in an upper extremity.
b. Diagnosis was not yet known.

c. Possible confounders, such as gender and age were reported. This item was scored negative if subjects were recruited from the neurologist, orthopaedist, or a electrophysiology laboratory.

2. Positive if the selection criteria were clear, including for example, the time period of recruitment, whether subjects were consecutively recruited, and age of the subjects.
3. Positive if both electrodiagnostic testing and advanced imaging (e.g. MRI) were used as the reference standard.
4. Positive if the delay between the application of the index test and reference standard was not more than 7 days.
5. Positive if it was clear that all subjects or a random selection of subjects received verification of their disease status, regardless of the index test results. Also known as ''work-up bias''.

6–7. Positive if the study includes sufficient details which permit replication of the index test and reference standard.

8–9. Positive if the results of the index test were interpreted without knowledge of the results of the reference standard, and vice versa. Also known as review bias when clinicians are not ''blinded'' to the results of either the index test or reference standard.

10. Positive if clinical data (e.g. clinical presentation, symptoms, severity, etc) were also available to the clinician in order to interpret the results of the index test. Since clinical data will also be available to the clinician in the practice, the clinical picture should also be available to clinicians in the study.
11. Positive if all test results, including uninterpretable/intermediate/equivocal test results were included. This item was scored as a ''?'' in those studies where no equivocal results were reported.
12. Positive if it is clear what happened to all subjects who entered the study. Results are biased if subjects dropped-out systematically. This item was scored negative when some of the subjects did not receive both the index test and reference standard, and these subjects were not described anywhere in the text.

References

1. DJ Magee (ed) (1997) Cervical spine. In: Orthopedic physical assessment, 3rd edn. W.B. Saunders, Philadelphia, pp 101–151
2. R Evans (ed) (2001) Cervical spine. In: Illustrated orthopedic physical assessment, 2nd edn. C.V. Mosby, St. Louis
3. AN de Wolf, JMA Mens (eds) (2001) De cervicale wervelkolom. In: Onderzoek van het bewegingsapparaat. Fysische diagnostiek in de algemene praktijk. Derde, Derde edn. Bohn, Stafleu, van Loghn, Houten/Diegem, The Netherlands, pp 137–156
4. Abdulwahab SS, Sabbahi M (2000) Neck retractions, cervical root decompression, and radicular pain. J Orthop Sports Phys Ther 30(1):4–9
5. Ahlgren BD, Garfin SR (1996) Cervical radiculopathy. Orthop Clin North Am 27(2):253–263
6. Albeck MJ (1996) A critical assessment of clinical diagnosis of disc herniation in patients with monoradicular sciatica. Acta Neurochir (Wien) 138(1):40–44
7. American Association of Electrodiagnostic Medicine (1999) Guidelines in electrodiagnostic medicine, Chap 9. The electrodiagnostic evaluation of patients with suspected cervical radiculopathy: literature review on the usefulness of needle electromyography. Muscle Nerve 22(Suppl 8):S213–S221
8. American Association of Electrodiagnostic Medicine (1999) Practice parameter for needle electromyogrphic evaluation of patients with suspected cervical radiculopathy: summary statement. Muscle Nerve 22(Suppl 8):S209–S211
9. An HS (1996) Cervical root entrapment. Hand Clin 12(4):719–730
10. Bachmann LM, Coray R, Estermann P, Ter Riet G (2002) Identifying diagnostic studies in MEDLINE: reducing the number needed to read. J Am Med Inform Assoc 9(6):653–658
11. Bartlett RJ, Hill CA, Devlin R, Gardiner ED (1996) Two-dimensional MRI at 1.5 and 0.5 T versus CT myelography in the diagnosis of cervical radiculopathy. Neuroradiology 38(2):142–147
12. Bartlett RJ, Hill CR, Gardiner E (1998) A comparison of T2 and gadolinium enhanced MRI with CT myelography in cervical radiculopathy. Br J Radiol 71(841):11–19
13. Bednar MS, Light TR (2003) Hand surgery. In: HB Skinner (ed) Current diagnosis and treatment in orthopedics, 3rd edn. Lange Medical Books/McGraw-Hill, New York, pp 525–588
14. Bendszus M, Wessig C, Reiners K et al (2003) MR imaging in the differential diagnosis of neurogenic foot drop. AJNR Am J Neuroradiol 24(7):1283–1289
15. Birchall D, Connelly D, Walker L, Hall K (2003) Evaluation of magnetic resonance myelography in the investigation of cervical spondylotic radiculopathy. Br J Radiol 76(908):525–531
16. Bland JH (1987) Clinical methods. In: Bland J.H. (ed) Disorders of the cervical spine. Diagnosis and medical management. W.B. Saunders, Philadelphia, pp 79–112
17. Boden SD, McCowin PR, Davis DO et al (1990) Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am 72(8):1178–1184
18. Carette S, Fehlings MG (2005) Clinical practice. Cervical radiculopathy. N Engl J Med 353(4):392–399
19. Coppieters M, Stappaerts K, Janssens K, Jull G (2002) Reliability of detecting 'onset of pain' and 'submaximal pain' during neural provocation testing of the upper quadrant. Physiother Res Int 7(3):146–156
20. Daniels DL, Grogan JP, Johansen JG et al (1984) Cervical radiculopathy: computed tomography and myelography compared. Radiology 151(1):109–113
21. Davidson RI, Dunn EJ, Metzmaker JN (1981) The shoulder abduction test in the diagnosis of radicular pain in cervical

extradural compressive monoradiculopathies. Spine 6(5):441–446
22. Eisen A (1991) The utility of proximal nerve conduction in radiculopathies: the cons. Electroencephalogr Clin Neurophysiol 78(3):171–172
23. Ellenberg MR, Honet JC, Treanor WJ (1994) Cervical radiculopathy. Arch Phys Med Rehabil 75(3):342–352
24. Elvey RL (1994) The investigation of arm pain: signs of adverse responses to the physical examination of the brachial plexus, related tissues. In: Boyling JD, Palastanga N (eds) Grieve's modern manual therapy, 2nd edn. Churchill Livingstone, New York, pp 577–585
25. Fast A, Parikh S, Marin EL (1989) The shoulder abduction relief sign in cervical radiculopathy. Arch Phys Med Rehabil 70(5):402–403
26. Friedenberg Z, Miller W (1963) Degenerative disc disease of the cervical spine. J Bone Joint Surg Am 45:1171–1178
27. Han JJ, Kraft GH (2003) Electrodiagnosis of neck pain. Phys Med Rehabil Clin N Am 14(3):549–567
28. Honet JC, Ellenberg MR (2003) What you always wanted to know about the history and physical examination of neck pain but were afraid to ask. Phys Med Rehabil Clin N Am 14(3):473–491
29. Jonsson B, Stromqvist B (1995) The straight leg raising test and the severity of symptoms in lumbar disc herniation. A preoperative evaluation. Spine 20(1):27–30
30. Katirji MB, Agrawal R, Kantra TA (1988) The human cervical myotomes: an anatomical correlation between electromyography and CT/myelography. Muscle Nerve 11(10):1070–1073
31. Kelsey JL (1978) Epidemiology of radiculopathies. Adv Neurol 19:385–398
32. Kleinrensink GJ, Stoeckart R, Mulder PG et al (2000) Upper limb tension tests as tools in the diagnosis of nerve and plexus lesions. Anatomical and biomechanical aspects. Clin Biomech (Bristol, Avon) 15(1):9–14
33. Kosteljanetz M, Bang F, Schmidt-Olsen S (1988) The clinical significance of straight-leg raising (Lasegue's sign) in the diagnosis of prolapsed lumbar disc. Interobserver variation and correlation with surgical finding. Spine 13(4):393–395
34. Larsson EM, Holtas S, Cronqvist S, Brandt L (1989) Comparison of myelography, CT myelography and magnetic resonance imaging in cervical spondylosis and disk herniation. Pre- and postoperative findings. Acta Radiol 30(3):233–239
35. Lindgren KA, Leino E, Manninen H (1992) Cervical rotation lateral flexion test in brachialgia. Arch Phys Med Rehabil 73(8):735–737
36. Malanga GA, Landes P, Nadler SF (2003) Provocative tests in cervical spine examination: historical basis and scientific analyses. Pain Physician 6:199–205
37. Malanga G, Landes P, Nadler S (2003) Provocative tests in cervical spine examination: historical basis and scientific analyses. Pain Physician 6:199–205
38. Modic MT, Masaryk TJ, Mulopulos GP et al (1986) Cervical radiculopathy: prospective evaluation with surface coil MR imaging, CT with metrizamide, and metrizamide myelography. Radiology 161(3):753–759
39. Negrin P, Lelli S, Fardin P (1991) Contribution of electromyography to the diagnosis, treatment and prognosis of cervical disc disease: a study of 114 patients. Electromyogr Clin Neurophysiol 31(3):173–179
40. Partanen J, Partanen K, Oikarinen H, Niemitukia L, Hernesniemi J (1991) Preoperative electroneuromyography and myelography in cervical root compression. Electromyogr Clin Neurophysiol 31(1):21–26
41. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P (2004) Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol 159(9):882–890
42. Quintner JL (1989) A study of upper limb pain and paraesthesiae following neck injury in motor vehicle accidents: assessment of the brachial plexus tension test of Elvey. Br J Rheumatol 28(6):528–533
43. Radhakrishnan K, Litchy WJ, O'Fallon WM, Kurland LT (1994) Epidemiology of cervical radiculopathy. A population-based study from Rochester, Minnesota, 1976 through 1990. Brain 117(Pt 2):325–335
44. Reilly PA, Littlejohn GO (1990) The brachial plexus tension test of Elvey. Br J Rheumatol 29(3):234–235
45. Saal JS (1995) The role of inflammation in lumbar pain. Spine 20(16):1821–1827
46. Sandmark H, Nisell R (1995) Validity of five common manual neck pain provoking tests. Scand J Rehabil Med 27(3):131–136
47. Schwartzman RJ (1991) Brachial plexus traction injuries. Hand Clin 7(3):547–556
48. Selvaratnam PJ, Matyas TA, Glasgow EF (1994) Noninvasive discrimination of brachial plexus involvement in upper limb pain. Spine 19(1):26–33
49. Shacklock MO (1996) Positive upper limb tension test in a case of surgically proven neuropathy: analysis and validity. Man Ther 1(3):154–161
50. Shafaie FF, Wippold FJ, Gado M, Pilgram TK, Riew KD (1999) Comparison of computed tomography myelography and magnetic resonance imaging in the evaluation of cervical spondylotic myelopathy and radiculopathy. Spine 24(17):1781–1785
51. Shah KC, Rajshekhar V (2004) Reliability of diagnosis of soft cervical disc prolapse using Spurling's test. Br J Neurosurg 18(5):480–483
52. Spengler DM, Freeman CW (1979) Patient selection for lumbar discectomy. An objective approach. Spine 4(2):129–134
53. Spengler DM, Ouellette EA, Battie M, Zeh J (1990) Elective discectomy for herniation of a lumbar disc. Additional experience with an objective method. J Bone Joint Surg Am 72(2):230–237
54. Spurling RG, Scoville WB (1944) Lateral rupture of the cervical intervertebral discs: a common cause of shoulder and arm pain. Surg Gynecol Obstet 78:350–358
55. Sterling M, Treleaven J, Jull G (2002) Responses to a clinical test of mechanical provocation of nerve tissue in whiplash associated disorder. Man Ther 7(2):89–94
56. Stratton SA, Bryan JM (1989) Dysfunction, evaluation, and treatment of the cervical spine and thoracic inlet. In: Donatelli R, Wooden MJ (eds) Orthopaedic physical therapy. Churchill Livingstone, New York, pp 71–108
57. Sweeney PJ (1995) Clinical evaluation of cervical radiculopathy and myelopathy. Neuroimaging Clin N Am 5(3):321–327
58. Teresi LM, Lufkin RB, Reicher MA et al (1987) Asymptomatic degenerative disk disease and spondylosis of the cervical spine: MR imaging. Radiology 164(1):83–88
59. Tong HC, Haig AJ, Yamakawa K (2002) The Spurling test and cervical radiculopathy. Spine 27(2):156–159
60. Uchihara T, Furukawa T, Tsukagoshi H (1994) Compression of brachial plexus as a diagnostic test of cervical cord lesion. Spine 19(19):2170–2173
61. Viikari-Juntura E (1987) Interexaminer reliability of observations in physical examinations of the neck. Phys Ther 67(10):1526–1532

62. Viikari-Juntura E, Porras M, Laasonen EM (1989) Validity of clinical tests in the diagnosis of root compression in cervical disc disease. Spine 14(3):253–257
63. Wainner RS, Fritz JM, Irrgang JJ et al (2003) Reliability and diagnostic accuracy of the clinical examination and patient self-report measures for cervical radiculopathy. Spine 28(1):52–62
64. Wetzel FT, Reider B (2005) Cervical and thoracic spine. In: B Reider (eds) The orthopaedic physical exam. 2nd edn. Elsevier Saunders, Philadelphia, pp 297–334
65. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 3(1):25