

# BMJ Open Measurement properties of the craniocervical flexion test: a systematic review protocol

Francisco Xavier de Araujo,<sup>1,2</sup> Giovanni Esteves Ferreira,<sup>1</sup> Maurício Scholl Schell,<sup>1</sup> Marcelo Peduzzi de Castro,<sup>3</sup> Marcelo Faria Silva,<sup>1</sup> Daniel Cury Ribeiro<sup>4</sup>

**To cite:** Araujo FX, Ferreira GE, Scholl Schell M, *et al*. Measurement properties of the craniocervical flexion test: a systematic review protocol. *BMJ Open* 2018;**8**:e019486. doi:10.1136/bmjopen-2017-019486

► Prepublication history for this paper is available online. To view these files, please visit the journal online <http://dx.doi.org/10.1136/bmjopen-2017-019486>.

Received 5 September 2017

Revised 17 January 2018

Accepted 18 January 2018



<sup>1</sup>Physical Therapy Department, Graduate Program in Rehabilitation Sciences, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSA), Porto Alegre, Brazil

<sup>2</sup>Physical Therapy Department, Centro Universitário Ritter dos Reis (UniRitter), Laureate International Universities, Porto Alegre, Brazil

<sup>3</sup>Neuromusculoskeletal Rehabilitation and Clinical Biomechanics Laboratory (LaBClin), Florianópolis, Brazil

<sup>4</sup>School of Physiotherapy, University of Otago, Dunedin, New Zealand

## Correspondence to

Daniel Cury Ribeiro;  
[daniel.ribeiro@otago.ac.nz](mailto:daniel.ribeiro@otago.ac.nz)

## ABSTRACT

**Introduction** Neck pain is the leading cause of years lived with disability worldwide and it accounts for high economic and societal burden. Altered activation of the neck muscles is a common musculoskeletal impairment presented by patients with neck pain. The craniocervical flexion test with pressure biofeedback unit has been widely used in clinical practice to assess function of deep neck flexor muscles. This systematic review will assess the measurement properties of the craniocervical flexion test for assessing deep cervical flexor muscles.

**Methods and analysis** This is a protocol for a systematic review that will follow the Preferred Reporting Items for Systematic Review and Meta-Analysis statement. MEDLINE (via PubMed), EMBASE, PEDro, Cochrane Central Register of Controlled Trials (CENTRAL), Scopus and Science Direct will be systematically searched from inception. Studies of any design that have investigated and reported at least one measurement property of the craniocervical flexion test for assessing the deep cervical flexor muscles will be included. All measurement properties will be considered as outcomes. Two reviewers will independently rate the risk of bias of individual studies using the updated COnsensus-based Standards for the selection of health Measurement Instruments risk of bias checklist. A structured narrative synthesis will be used for data analysis. Quantitative findings for each measurement property will be summarised. The overall rating for a measurement property will be classified as 'positive', 'indeterminate' or 'negative'. The overall rating will be accompanied with a level of evidence.

**Ethics and dissemination** Ethical approval and patient consent are not required since this is a systematic review based on published studies. Findings will be submitted to a peer-reviewed journal for publication.

**PROSPERO registration number** CRD42017062175.

## INTRODUCTION

Neck pain is the leading cause of years lived with disability worldwide, and it accounts for high economic and societal burden.<sup>1,2</sup> In the general population, 16.7%–75.1% of adults will develop an episode of neck pain in any given year.<sup>3,4</sup> Patients may present recurrent neck pain,<sup>5,6</sup> and the prognosis of recovery is poor.<sup>5</sup> Between 50% and 75% of people who

## Strengths and limitations of this study

- Comprehensive and exhaustive search for relevant studies from several databases.
- A new summary of the evidence on measurement properties of a widely used clinical test: the craniocervical flexion test with the pressure biofeedback unit for the assessment of deep cervical flexor muscles.
- This review used the internationally recognised, validated COnsensus-based Standards for the selection of health Measurement Instruments risk of bias checklist to assess the methodological quality of the included studies when assessing the quality of the craniocervical flexion test.
- The proposed systematic review will adhere to the Preferred Reporting Items for Systematic Review and Meta-Analysis guidelines, ensuring consistency and uniformity in reporting and the full systematic review.
- A limitation of the review is that it will only include papers published in English.

experienced neck pain still present with symptoms 1–5 years after the onset of symptoms.<sup>6</sup>

Altered activation of the neck muscles is a common musculoskeletal impairment presented by patients with neck pain.<sup>7</sup> Compared with asymptomatic individuals, patients with neck pain exhibit increased activity of superficial neck flexors and reduced activity of the deep neck flexors,<sup>8</sup> poor muscle endurance,<sup>9,10</sup> altered kinematics of the cervical spine,<sup>11</sup> delayed feedforward activity<sup>12</sup> and impaired proprioception.<sup>13–15</sup> These impairments are likely to contribute to maintenance of symptoms in patients with chronic neck pain.<sup>16</sup>

As the clinical presentation of patients with neck pain are not homogeneous, clinical assessment of neck muscle function is important for identifying musculoskeletal impairments and tailoring treatment to patients' needs.<sup>17</sup> Several tests have been designed to evaluate different aspects of neck

muscle performance.<sup>18–23</sup> Among those tests, the craniocervical flexion test with pressure biofeedback unit has been developed to evaluate the ability of an individual to selectively recruit the deep neck flexors (*longus capitis* and *longus colli*) while maintaining low activity levels of the superficial neck flexors (eg, *sternocleidomastoid*, *anterior scalene*) during an active craniocervical flexion in supine lying.<sup>23</sup> This test has been widely used in clinical practice to assess function of deep neck flexor muscles.<sup>8 23 24</sup>

To conduct the craniocervical flexion test with pressure biofeedback unit, the patient's head is positioned in neutral, with patient in supine crook lying. The test consists of an active head nodding movement. During this movement, the patient attempts to target five different pressure levels, from 22 to 30 mm Hg.<sup>24</sup> This test is performed with an extrinsic air-filled pressure biofeedback unit placed behind the neck. This device provides feedback and direction for the patient to perform the test and enables an objective and quantitative assessment of the patient performance.<sup>23</sup> The performance of the test is assessed through the following components: how well the patient performs the active head nodding and achieves that by contracting the deep cervical flexors without contraction of superficial flexors), muscle endurance (through isometric contraction) of deep cervical flexors at each test stages with appropriate craniocervical flexion contraction and quality and range of craniocervical movement in the sagittal plane (which is expected to increase as the patient progress through the five different pressure levels).<sup>8 23</sup>

Ideally, the measurement properties (eg, reliability, validity and responsiveness) of an instrument or test, for instance, the craniocervical flexion test with pressure biofeedback unit, should be assessed before its full implementation in clinical practice.<sup>25</sup> Selecting instruments with proper measurement properties is fundamental for well-conducted clinical trials.<sup>26</sup> Hence, systematic reviews of measurement properties are useful for identifying instruments and tools with the highest reliability, validity and responsiveness scores.<sup>27</sup>

A previous systematic review<sup>28</sup> evaluated the measurement properties of methods to measure muscle function in patients with non-specific neck pain. In this review, the intraobserver reliability was the only measurement property assessed.<sup>28</sup> This review was conducted over 10 years ago, included only four studies and used a checklist adapted from two previous studies.<sup>29 30</sup> Since then, the number of published studies evaluating measurement properties of craniocervical flexion test has increased. In addition, new tools have been developed for assessing methodological quality of individual studies exploring measurement properties of instruments (ie, Consensus-based Standards for the selection of health Measurement Instruments (COSMIN)).<sup>25</sup> It is likely that a new review evaluating the measurement properties of the craniocervical flexion test with pressure biofeedback unit will provide relevant insights on the state of research in this field. This systematic review will critically appraise

and summarise the quality of the measurement properties of the craniocervical flexion test for assessing deep cervical flexor muscles.

## METHODS

### Protocol and registration

This is a protocol for a systematic review that was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P).<sup>31</sup> The systematic review has been registered with Prospective Register of Systematic Reviews (PROSPERO) (CRD42017062175).

### Eligibility criteria

We will include studies if they meet the following criteria:

- ▶ Articles that have investigated and reported at least one measurement property (ie, validity, reliability or responsiveness) of the craniocervical flexion test for assessing the deep cervical flexor muscles.
- ▶ Articles published in English.
- ▶ Assessing participants older than 18 years.
- ▶ Articles available in full text.
- ▶ Studies with both asymptomatic and symptomatic individuals (including those with acute, subacute and chronic neck pain with or without nerve root compromise; neck-related shoulder pain, whiplash-associated disorders and neck disorders associated with headache).

Studies assessing only the effectiveness of interventions, but not reporting measurement property outcomes of pressure biofeedback unit for assessing motor control of deep cervical flexor muscles will be excluded.

### Outcomes

All measurement properties will be considered as outcomes in this systematic review. We will adopt the COSMIN terminology and definitions of measurement properties.<sup>32</sup> Reliability is defined as the degree to which a measurement is free from measurement error; validity is the degree to which an outcome instrument measures the construct(s) it purports to measure and responsiveness is the ability of an outcome instrument to detect change over time in the construct to be measured.<sup>32</sup> Among these properties, reliability and validity are further subdivided. For example, reliability is further classified into reliability, internal consistency and measurement error; validity comprises content validity, construct validity and criterion validity.<sup>32</sup> For the purposes of this review, we will include all outcome measures used assessing psychometric properties that are reported by included studies.

### Search strategy

The search strategy was designed through consultation with a health sciences faculty librarian. Our search will include the following databases: MEDLINE (via PubMed), EMBASE, PEDro, Cochrane Central Register of Controlled Trials (CENTRAL), Scopus and Science

**Table 1** Search strategy in MEDLINE (via Pubmed)

1#	<p>((((((patient outcome assessment (Mesh)) OR ((outcome and process assessment (Mesh))) OR treatment outcome (Mesh) OR instrumentation (sh) OR methods (sh) OR 'Validation Studies' (pt) OR 'Comparative Study' (pt) OR 'psychometrics' (MeSH) OR psychometr* (tiab) OR clinimetr* (tw) OR clinometr*(tw) OR 'outcome assessment (healthcare)' (MeSH) OR 'outcome assessment' (tiab) OR 'outcome measure*(tw) OR 'observer variation' (MeSH) OR 'observer variation' (tiab) OR 'Health Status Indicators' (Mesh) OR 'reproducibility of results' (MeSH) OR reproducib* (tiab) OR 'discriminant analysis' (MeSH) OR reliab*(tiab) OR unreliab* (tiab) OR valid* (tiab) OR 'coefficient of variation'(tiab) OR coefficient (tiab) OR homogeneity (tiab) OR homogeneous(tiab) OR 'internal consistency' (tiab) OR (cronbach* (tiab) AND (alpha (tiab) OR alphas (tiab)))) OR (item (tiab) AND (correlation* (tiab) OR selection* (tiab) OR reduction* (tiab))) OR agreement (tw) OR precision (tw) OR imprecision (tw) OR 'precise values' (tw) OR test-retest(tiab) OR (test (tiab) AND retest (tiab)) OR (reliab* (tiab) AND (test(tiab) OR retest(tiab))) OR stability (tiab) OR interrater (tiab) OR inter-rater (tiab) OR intrarater (tiab) OR intra-rater (tiab) OR intertester (tiab) OR inter-tester (tiab) OR intratester (tiab) OR intra-tester (tiab) OR interobserver (tiab) OR inter-observer (tiab) OR intraobserver (tiab) OR intra-observer (tiab) OR intertechnician (tiab) OR inter-technician (tiab) OR intratechnician (tiab) OR intra-technician (tiab) OR interexaminer (tiab) OR inter-examiner (tiab) OR intraexaminer (tiab) OR intra-examiner (tiab) OR interassay (tiab) OR inter-assay (tiab) OR intraassay (tiab) OR intra-assay (tiab) OR interindividual (tiab) OR inter-individual (tiab) OR intraindividual (tiab) OR intra-individual (tiab) OR interparticipant (tiab) OR inter-participant (tiab) OR intraparticipant (tiab) OR intra-participant (tiab) OR kappa (tiab) OR kappa's (tiab) OR kappas (tiab) OR repeatab*(tw) OR ((replicab* (tw) OR repeated (tw)) AND (measure (tw) OR measures (tw) OR findings (tw) OR result (tw) OR results (tw) OR test (tw) OR tests (tw))) OR generaliza* (tiab) OR generalisa* (tiab) OR concordance (tiab) OR (intraclass (tiab) AND correlation* (tiab)) OR discriminative (tiab) OR 'known group' (tiab) OR 'factor analysis' (tiab) OR 'factor analyses' (tiab) OR 'factor structure' (tiab) OR 'factor structures' (tiab) OR dimension* (tiab) OR subscale* (tiab) OR (multitrait (tiab) AND scaling (tiab) AND (analysis(tiab) OR analyses(tiab))) OR 'item discriminant' (tiab) OR 'interscale correlation*' (tiab) OR error(tiab) OR errors(tiab) OR 'individual variability' (tiab) OR 'interval variability' (tiab) OR 'rate variability' (tiab) OR (variability (tiab) AND (analysis (tiab) OR values (tiab))) OR (uncertainty (tiab) AND (measurement (tiab) OR measuring (tiab))) OR 'SE of measurement' (tiab) OR sensitiv* (tiab) OR responsive* (tiab) OR (limit (tiab) AND detection (tiab)) OR 'minimal detectable concentration'(tiab) OR interpretab*(tiab) OR ((minimal (tiab) OR minimally (tiab) OR clinical (tiab) OR clinically (tiab)) AND (important (tiab) OR significant (tiab) OR detectable (tiab)) AND (change (tiab) OR difference (tiab))) OR (small* (tiab) AND (real (tiab) OR detectable (tiab)) AND (change (tiab) OR difference (tiab))) OR 'meaningful change' (tiab) OR 'ceiling effect' (tiab) OR 'floor effect' (tiab) OR 'Item response model' (tiab) OR IRT (tiab) OR Rasch (tiab) OR 'Differential item functioning'(tiab) OR DIF(tiab) OR 'computer adaptive testing' (tiab) OR 'item bank' (tiab) OR 'cross-cultural equivalence' (tiab))))))</p>
2#	<p>((((((pressure biofeedback unit (Title/Abstract) OR pressure biofeedback units (Title/Abstract) OR unit, pressure biofeedback (Title/Abstract) OR units, pressure biofeedback (Title/Abstract) OR stabilizer (Title/Abstract) OR (Title/Abstract) OR stabiliser (Title/Abstract) OR stabilizers (Title/Abstract) OR biofeedback (Title/Abstract) OR biofeedbacks (Title/Abstract) OR craniocervical flexion test (Title/Abstract) OR craniocervical flexion test (Title/Abstract) OR craniocervical flexion test (Title/Abstract) OR craniocervical flexion (Title/Abstract))))))</p>
3#	<p>((((((Muscle, Neck (Title/Abstract) OR Muscles, Neck (Title/Abstract) OR Neck muscle (Title/Abstract)))) OR Neck muscles (MeSH Terms)) OR ((neck (MeSH Terms)) OR Necks (Title/Abstract) OR deep cervical flexor* (Title/Abstract) OR rectus capit* (Title/Abstract) OR longus colli (Title/Abstract) OR longus capiti (Title/Abstract)))</p>
4#	<p>1# AND 2# AND 3#</p>

Direct. All databases will be searched from their inception to present time using a published search filter for finding studies on measurement properties.<sup>33</sup> MEDLINE full-search strategy is described in [table 1](#).

### Data extraction

Two reviewers (FXA and MSS) will independently screen titles and abstracts for eligibility. A third reviewer (MPC) will resolve any disagreement. The full text of potentially eligible articles will be screened independently by two reviewers (FXA and MSS). Data from included studies will be extracted independently by the two reviewers, using a piloted data collection form. Data will then be compared for accuracy, and disagreements will be resolved by consensus. The following information will be extracted from the included studies: study design, sample characteristics, measurement properties (eg, validity, reliability or responsiveness) assessed by included studies,

craniocervical flexion test procedures and results of the measurements properties.

### Risk of bias within included studies

Two reviewers (FXA and GEF) will independently rate the risk of bias of individual studies using the updated COSMIN risk of bias checklist.<sup>34</sup> The COSMIN risk of bias checklist is a validated critical appraisal tool designed for the systematic evaluation of the methodological quality of studies on the measurement properties.<sup>27</sup> For each study, only applicable domains to the study being assessed will be used for assessing the quality of the study. Disagreements between reviewers will be resolved by a third reviewer (MPC).

### Synthesis of results

A structured narrative synthesis will be used for data analysis. Quantitative findings for each measurement

**Table 2** Quality criteria for measurement properties

Property	Rating	Quality criteria
<b>Reliability</b>		
Internal consistency	+	Cronbach's alpha(s) $\geq 0.70$
	?	Cronbach's alpha not determined or unidimensionality unknown
	-	Cronbach's alpha(s) $< 0.70$
Reliability	+	ICC/weighted Kappa $\geq 0.70$ OR Pearson's r $\geq 0.80$
	?	Neither ICC/weighted Kappa, nor Pearson's r determined
	-	ICC/weighted Kappa $< 0.70$ OR Pearson's r $< 0.80$
Measurement error	+	MIC $>$ SDC OR MIC outside the LoA
	?	MIC not defined
	-	MIC $\leq$ SDC OR MIC equals or inside LoA
<b>Validity</b>		
Content validity	+	All items are considered to be relevant for the construct to be measured, for the target population and for the purpose of the measurement AND the questionnaire is considered to be comprehensive
	?	Not enough information available
	-	Not all items are considered to be relevant for the construct to be measured, for the target population and for the purpose of the measurement OR the questionnaire is considered not to be comprehensive
Construct validity Structural validity	+	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	-	Factors explain $< 50\%$ of the variance
Hypothesis testing	+	Correlations with instruments measuring the same construct $\geq 0.50$ OR at least 75% of the results are in accordance with the hypotheses AND correlations with related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with instruments measuring the same construct $< 0.50$ OR $< 75\%$ of the results are in accordance with the hypotheses OR correlations with related constructs are lower than with unrelated constructs
Cross-cultural validity	+	No differences in factor structure OR no important DIF between language versions
	?	Multiple group factor analysis not applied AND DIF not assessed
	-	Differences in factor structure OR important DIF between language versions
Criterion validity	+	Convincing arguments that gold standard is 'gold' AND correlation with gold standard $\geq 0.70$
	?	No convincing arguments that gold standard is 'gold'
	-	Correlation with gold standard $< 0.70$
<b>Responsiveness</b>		
Responsiveness	+	Correlation with changes on instruments measuring the same construct $\geq 0.50$ OR at least 75% of the results are in accordance with the hypotheses OR AUC $\geq 0.70$ AND correlations with changes in related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with changes on instruments measuring the same construct $< 0.50$ OR $< 75\%$ of the results are in accordance with the hypotheses OR AUC $< 0.70$ OR correlations with changes in related constructs are lower than with unrelated constructs

+, positive rating; -, negative rating; ?, indeterminate rating; AUC, area under the curve; DIF, differential item functioning; ICC, intraclass correlation coefficient; LoA, limits of agreement; MIC, minimal important change; SDC, smallest detectable change.

property will be summarised. The overall rating for a measurement property will be classified as 'positive', 'indeterminate' or 'negative'. The overall rating will be accompanied with a level of evidence (strong, moderate, limited, conflicting, unknown—[table 2](#)) as

proposed by Terwee *et al.*<sup>35</sup> The criteria used to assign levels of evidence for the quality of each measurement property ([table 3](#)) will follow the framework proposed by the Cochrane Back and Neck Review Group.<sup>36</sup>



**Table 3** Levels of evidence for the quality of the measurement property

Level	Rating	Criteria
Strong	+++ or ---	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++ or --	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	+/-	Conflicting findings
Unknown	?	Only studies of poor methodological quality

–, negative rating; ?, indeterminate rating; +, positive rating.

We will conduct a narrative synthesis of subgroups, if applicable, based on the sample characteristics (ie, asymptomatic or symptomatic) and type of disorder (eg, acute, subacute and chronic non-specific neck pain; acute, subacute and chronic neck pain with nerve root compromise; neck-related shoulder pain; whiplash-associated disorders and neck disorder associated with headache).

### Ethics and dissemination

Ethical approval and patient consent are not required since this is a systematic review based on published studies. This protocol has been registered on the international PROSPERO and the systematic review will be conducted according to the PRISMA statement. The results of this systematic review will be submitted to a peer-reviewed journal for publication and will be presented at national and international conferences.

**Contributors** FXdeA is a PhD student and the leading researcher, responsible for conceiving the study and designing the protocol. FXdeA and MSS were involved in the screening of title, abstracts and full text for eligibility. MSS and GEF extracted data, provided the statistical analysis plan of the study and conducted the data analysis. FXdeA and GEF rated the methodological quality of individual studies. MPdeC resolved any disagreement between reviewers. FXdeA, MSS and GEF wrote the first version of the paper. MPdeC, DCR and MFS provided the critical revision of the paper. All authors read and provided final approval of this protocol to be published. All authors have contributed to the conception and design of the study protocol, development of the search strategy, the establishment of the inclusion and exclusion criteria, data extraction criteria, analyses and interpretation.

**Funding** This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. FXdeA, GEF and MSS are supported by a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Brazil).

**Disclaimer** CAPES was not involved in any aspect of this study protocol.

**Competing interests** None declared.

**Patient consent** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** This is a protocol and all data is available on the protocol reporting.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

### REFERENCES

- Vos T, Allen C, Arora M, *et al*. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388:1545–602.
- Hoy D, March L, Woolf A, *et al*. The global burden of neck pain: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:1309–15.
- Hogg-Johnson S, van der Velde G, Carroll LJ, *et al*. The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000–2010 task force on neck pain and its associated disorders. *J Manipulative Physiol Ther* 2009;32:46–60.
- Fejer R, Kyvik KO, Hartvigsen J. The prevalence of neck pain in the world population: a systematic critical review of the literature. *Eur Spine J* 2006;15:834–48.
- Hush JM, Lin CC, Michaleff ZA, *et al*. Prognosis of acute idiopathic neck pain is poor: a systematic review and meta-analysis. *Arch Phys Med Rehabil* 2011;92:824–9.
- Carroll LJ, Hogg-Johnson S, van der Velde G, *et al*. Course and prognostic factors for neck pain in the general population: results of the Bone and Joint Decade 2000–2010 Task Force on Neck Pain and Its Associated Disorders. *J Manipulative Physiol Ther* 2009;32:87–96.
- Falla D, Farina D. Neuromuscular adaptation in experimental and clinical neck pain. *J Electromyogr Kinesiol* 2008;18:255–61.
- Jull G, Falla D. Does increased superficial neck flexor activity in the craniocervical flexion test reflect reduced deep flexor activity in people with neck pain? *Man Ther* 2016;25:43–7.
- Oliveira AC, Silva AG. Neck muscle endurance and head posture: a comparison between adolescents with and without neck pain. *Man Ther* 2016;22:62–7.
- Falla D, Jull G, Russell T, *et al*. Effect of neck exercise on sitting posture in patients with chronic neck pain. *Phys Ther* 2007;87:408–17.
- Treleaven J, Chen X, Sarig Bahat H, *et al*. Factors associated with cervical kinematic impairments in patients with neck pain. *Man Ther* 2016;22:109–15.
- Falla D, Jull G, Hodges PW. Feedforward activity of the cervical flexor muscles during voluntary arm movements is delayed in chronic neck pain. *Exp Brain Res* 2004;157:43–8.
- de Zoete RMJ, Osmotherly PG, Rivett DA, *et al*. Sensorimotor control in individuals with idiopathic neck pain and healthy individuals: a systematic review and meta-analysis. *Arch Phys Med Rehabil* 2017;98:1257–71.
- de Vries J, Ischebeck BK, Voogt LP, *et al*. Joint position sense error in people with neck pain: a systematic review. *Man Ther* 2015;20:736–44.
- Stanton TR, Leake HB, Chalmers KJ, *et al*. Evidence of impaired proprioception in chronic, idiopathic neck pain: systematic review and meta-analysis. *Phys Ther* 2016;96:876–87.
- Lindström R, Schomacher J, Farina D, *et al*. Association between neck muscle coactivation, pain, and strength in women with neck pain. *Man Ther* 2011;16:80–6.
- O'Leary S, Falla D, Elliott JM, *et al*. Muscle dysfunction in cervical spine pain: implications for assessment and management. *J Orthop Sports Phys Ther* 2009;39:324–33.
- Blizzard L, Grimmer KA, Dwyer T. Validity of a measure of the frequency of headaches with overt neck involvement, and reliability of measurement of cervical spine anthropometric and muscle performance factors. *Arch Phys Med Rehabil* 2000;81:1204–10.
- Phillips BA, Lo SK, Mastaglia FL. Muscle force measured using 'break' testing with a hand-held myometer in normal subjects aged 20 to 69 years. *Arch Phys Med Rehabil* 2000;81:653–61.
- Silverman JL, Rodriguez AA, Agre JC. Quantitative cervical flexor strength in healthy subjects and in subjects with mechanical neck pain. *Arch Phys Med Rehabil* 1991;72:679–81.
- Grimmer K. Measuring the endurance capacity of the cervical short flexor muscle group. *Aust J Physiother* 1994;40:251–4.
- Harris KD, Heer DM, Roy TC, *et al*. Reliability of a measurement of neck flexor muscle endurance. *Phys Ther* 2005;85:1349–55.

23. Jull GA, O'Leary SP, Falla DL. Clinical assessment of the deep cervical flexor muscles: the craniocervical flexion test. *J Manipulative Physiol Ther* 2008;31:525–33.
24. Juul T, Langberg H, Enoch F, *et al.* The intra- and inter-rater reliability of five clinical muscle performance tests in patients with and without neck pain. *BMC Musculoskelet Disord* 2013;14:339.
25. Mokkink LB, Terwee CB, Patrick DL, *et al.* The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49.
26. Mokkink LB, Terwee CB, Stratford PW, *et al.* Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual Life Res* 2009;18:313–33.
27. Terwee CB, Mokkink LB, Knol DL, *et al.* Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;21:651–7.
28. de Koning CH, van den Heuvel SP, Staal JB, *et al.* Clinimetric evaluation of methods to measure muscle functioning in patients with non-specific neck pain: a systematic review. *BMC Musculoskelet Disord* 2008;9:142.
29. Lohr KN, Aaronson NK, Alonso J, *et al.* Evaluating quality-of-life and health status instruments: development of scientific review criteria. *Clin Ther* 1996;18:979–92.
30. Bot SD, Terwee CB, van der Windt DA, *et al.* Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004;63:335–41.
31. Moher D, Shamseer L, Clarke M, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
32. Mokkink LB, Terwee CB, Patrick DL, *et al.* The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737–45.
33. Terwee CB, Jansma EP, Riphagen II, *et al.* Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;18:1115–23.
34. Mokkink LB, de Vet HCW, Prinsen CAC, *et al.* COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2017.
35. Terwee CB, Bot SD, de Boer MR, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
36. Furlan AD, Pennick V, Bombardier C, *et al.* 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine* 2009;34:1929–41.