BMJ Open

Measurement properties of pressure biofeedback unit for the assessment of deep cervical flexor muscles: a systematic review protocol

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-019486
Article Type:	Protocol
Date Submitted by the Author:	05-Sep-2017
Complete List of Authors:	Araujo, Francisco Ferreira, Giovanni ; Universidade Federal de Ciencias da Saude de Porto Alegre, Physical Therapy Department Scholl Schell, Maurício ; Universidade Federal de Ciencias da Saude de Porto Alegre Castro, Marcelo; Neuromusculoskeletal Assessment and Clinical Biomechanics Laboratory (LaBClin) Silva, Marcelo; Universidade Federal de Ciencias da Saude de Porto Alegre, Physical Therapy Department Ribeiro, Daniel; School of Physiotherapy
Primary Subject Heading:	Rehabilitation medicine
Secondary Subject Heading:	Occupational and environmental medicine
Keywords:	Musculoskeletal disorders < ORTHOPAEDIC & TRAUMA SURGERY, REHABILITATION MEDICINE, SPORTS MEDICINE

MEASUREMENT PROPERTIES OF PRESSURE BIOFEEDBACK UNIT FOR THE ASSESSMENT OF DEEP CERVICAL FLEXOR MUSCLES: A SYSTEMATIC REVIEW PROTOCOL

Francisco Xavier de Araujo, PT, MSc[1,2], franciscoxaraujo@gmail.com, Corresponding author

Giovanni Esteves Ferreira, PT, MSc[1], giovannieferreira@hotmail.com

Maurício Scholl Schell, PT[1], mschell.fisio@gmail.com

Marcelo Peduzzi de Castro, PT, PhD[3], marcelo.peduzzi.castro@gmail.com

Marcelo Faria Silva, PT, PhD[1], marcelofs@ufcspa.edu.br

Daniel Cury Ribeiro, PT, PhD[4], daniel.ribeiro@otago.ac.nz, Corresponding author

[1] Physical Therapy Department, Graduate Program in Rehabilitation Sciences, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Brazil.

[2] Centro Universitário Ritter dos Reis (UniRitter), Laureate International Universities, Porto Alegre, Brazil.

[3] Neuromusculoskeletal Assessment and Clinical Biomechanics Laboratory (LaBClin), Florianópolis, Brazil.

[4] Centre for Health, Activity and Rehabilitation Research –School of Physiotherapy, University of Otago, Dunedin, New Zealand.

ABSTRACT

Introduction: Neck pain is the fourth 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 a pressure biofeedback unit has been widely used in clinical practice to assess the function of deep neck flexor muscles. This systematic review will assess the measurement properties of the pressure biofeedback unit for assessing deep cervical flexor muscles.

Methods and analysis: This review will follow the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement. We will systematically search the following databases: MEDLINE (via PubMed), EMBASE, PEDro, Cochrane Central Register of Controlled Trials (CENTRAL), Scopus and Science Direct. Studies of any design that have investigated and reported at least one measurement property of the pressure biofeedback unit for assessing the deep cervical flexor muscles during the craniocervical flexion test 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 COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist. A structured narrative synthesis will be used for data analysis. Quantitative findings for each measurement property will be summarized. 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.

Trial registration number: This protocol has been registered with the International Prospective Register of Systematic Reviews (PROSPERO), registration number (CRD42017062175).

STRENGTHS AND LIMITATIONS:

- Comprehensive and exhaustive search for relevant studies from several databases;

- An update of the evidence on measurement properties of a widely used clinical test: the pressure biofeedback unit for the assessment of deep cervical flexor muscles;

- A strength of this review is its use of the internationally recognized and validated COSMIN checklist to assess the methodological quality of the included studies;

- The proposed systematic review will adhere to the PRISMA 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.

INTRODUCTION

Neck pain is the fourth cause of years lived with disability worldwide, and it accounts for high economic and societal burden.[1,2] In the general population, 16.7 to 75.1% of adults will experience an episode of neck pain in any given year.[3,4] Patients may present with recurrent neck pain,[5,6] and the prognosis of recovery is poor.[5] Between 50 and 75% of people who experienced neck pain still present with symptoms one to five years after the onset of these symptoms.[6]

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

As the clinical presentation of patients with neck pain is not homogeneous, clinical assessment of neck muscle function is important for identifying musculoskeletal impairments and tailoring treatment to the patients' needs.[17] Several tests have been designed to evaluate different aspects of neck muscle performance.[18–23] Among those tests, the craniocervical flexion test with the 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 (e.g. *sternocleidomastoid*, *anterior scalene*) during an active craniocervical flexion in supine lying.[23] This test has been widely used in clinical practice to assess the function of deep neck flexor muscles.[8,23,24]

To conduct the craniocervical flexion test with pressure biofeedback unit, the patient is positioned in supine crook lying with the head in a neutral starting position, followed by an active head nodding action, during which the patient tries to sequentially target five progressive stages, from 22 to 30 mmHg.[24] 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.[23] Evaluation of the test involves different components: performance of the craniocervical flexion action (contracting adequately the deep cervical flexors without compensatory activity of the superficial flexors), isometric endurance of the deep cervical flexors at test stages that the patient is able to achieve with the correct craniocervical flexion action, and assessment of the quality and range of craniocervical rotation in the sagittal plane, which should proportionally increase as the stages of the test progresses.[8,23]

Ideally, the measurement properties (e.g. 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.[25] Selecting instruments with proper measurement properties is fundamental for well-conduced clinical trials.[26] Hence, systematic reviews of measurement properties are useful for identifying instruments and tools with the highest reliability, validity and responsiveness scores.[27]

A previous systematic review[28] evaluated the measurement properties of methods to measure muscle function in patients with non-specific neck pain. In this review, the intra-observer reliability was the only measurement property assessed.[28] This review was conducted over 10 years ago, included only four studies, and used a checklist adapted from two previous studies.[29,30] 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 the methodological quality of individual studies exploring the measurement properties of instruments (i.e. Consensus-based Standards for the selection of health Measurement Instruments – COSMIN).[25] It is likely that a new review evaluating the measurement properties of the craniocervical flexion test with pressure biofeedback unit will provide relevant insights into the state of research in this field. This systematic review will assess the

measurement properties of the pressure biofeedback unit for assessing deep cervical flexor muscles.

METHODS

Protocol and Registration

This protocol was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P).[31] The systematic review has been registered with PROSPERO (CRD42017062175).

Eligibility criteria

We will include studies if they meet the following criteria:

- Studies of any design (e.g. Cross-sectional studies or randomized clinical trials);
- Articles that have investigated and reported at least one measurement property (i.e. validity, reliability or responsiveness) of the pressure biofeedback unit for assessing the deep cervical flexor muscles during the craniocervical flexion test;
- Articles published in English;
- Assessing participants older than 18 years old;
- 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.[32] Reliability is defined as the extent to which scores for patients who have not changed are the same for repeated measurement under several conditions; 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.[32] 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.[32] For the purposes of this review, we will include all outcome measures used in assessing psychometric properties 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 Direct. All databases will be searched from their inception to present time using a published search filter for finding studies on measurement properties.[33] 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 solved by consensus. The following information will be extracted from the included studies: study design, sample characteristics, measurement

properties (e.g. validity, reliability or responsiveness) assessed by included studies, 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 COSMIN checklist.[25] The COSMIN checklist is a validated critical appraisal tool designed for the systematic evaluation of the methodological quality of studies on the measurement properties.[27] The checklist consists of nine domains concerning measurement properties. The number of items for each domain varies from 5 to 18. Each item deals with design characteristics and statistical methods used and reported by authors. Each item will be scored based on a four-point rating scale as "excellent", "good", "fair", or "poor". The lowest rating score of a domain will be used for attributing the quality score for that specific domain. 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 property will be summarized. 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.[34] The criteria used to assign levels of evidence for the quality of each measurement property will follow the framework proposed by the Cochrane Back and Neck Review Group.[35]

We will conduct a narrative synthesis of subgroups, if applicable, based on the sample characteristics (i.e. asymptomatic or symptomatic), and type of disorder (e.g. 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 may be presented at national and international meetings.

AUTHORS' CONTRIBUTIONS

FXA is the leading researcher, responsible for conceiving the study, and designing the protocol. 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. FXA and MSS will screen title, abstracts and full text for eligibility. MSS and GEF will extract data. FXA GEF will rate the methodological quality of individual studies. MPC will will resolve any disagreement between reviewers. GEF and MSS will provide the statistical analysis plan of the study and will conduct the data analysis. FXA, MSS, GEF will write the first version of the paper. MPC, DCR and MFS will provide critical revision of the paper. All authors read and approved this protocol for publication.

FUNDING STATEMENT

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. Araujo FX and Schell MS are supported by a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Brazil). CAPES is not involved in any other aspect of this study protocol.

COMPETING INTEREST

None declared.

REFERENCES

1. Vos, T. *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **386,** 743–800 (2015).
2. Hoy, D. *et al.* The global burden of neck pain: estimates from the Global Burden of Disease 2010 study. *Ann. Rheum. Dis.* **73,** 1309–1315 (2014).
3. Hogg-Johnson, S. *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.* **32,** 46–60 (2009).
4. Fejer, R., Kyvik, K. O. & Hartvigsen, J. The prevalence of neck pain in the world population: A systematic critical review of the literature. *Eur. Spine J.* **15,** 834–848 (2006).
5. Hush, J. M., Lin, C. C., Michaleff, Z. A., Verhagen, A. & Refshauge, K. M. Prognosis of acute idiopathic neck pain is poor: A systematic review and meta-analysis. *Arch. Phys. Med. Rehabil.* **92,** 824–829 (2011).
6. Carroll, L. J. *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.* **32,** 87–96 (2009).
7. Falla, D. & Farina, D. Neuromuscular adaptation in experimental and clinical neck pain. *J. Electromyogr. Kinesiol.* **18,** 255–261 (2008).
8. Jull, G. & Falla, D. Does increased super fi cial neck fl exor activity in the craniocervical fl exion test re fl ect reduced deep fl exor activity in people with neck. **25,** 43–47 (2016).
9. Oliveira, A. C. & Silva, A. G. Neck muscle endurance and head posture: A comparison between adolescents with and without neck pain. *Man. Ther.* **22,** 62–67 (2016).
10. Falla, D., Jull, G., Russell, T., Vicenzino, B. & Hodges, P. Effect of Neck Exercise on Sitting Posture in Patients With Chronic Neck Pain. *Phys. Ther.* **87,** 408–417 (2007).
11. Treleaven, J., Chen, X., Sarig Bahat, H. & Treleaven, Julia; Xiaoqi, Chen; Bahat, H. S. Factors associated with cervical kinematic impairments in patients with neck pain. *Man. Ther.* **22,** 109–115 (2016).
12. Falla, D., Jull, G. & Hodges, P. W. Feedforward activity of the cervical flexor muscles during voluntary arm movements is delayed in chronic neck pain. *Exp. Brain Res.* **157,** 43–48 (2004).
13. Zoete, R. M. J. De, Osmotherly, P. G., Rivett, D. A., Farrell, F. & Snodgrass, S. J. Sensorimotor control in individuals with idiopathic neck pain and healthy individuals: A systematic review and meta-analysis. (2016). doi:10.1016/j.apmr.2016.09.121.This

14. de Vries, J. *et al.* Joint position sense error in people with neck pain: A systematic review. *Man. Ther.* **20,** 736–744 (2015).

15. Stanton, T. R., Leake, H. B., Chalmers, K. J. & Moseley, G. L. Evidence of Impaired Proprioception in Chronic, Idiopathic Neck Pain: Systematic Review and Meta-Analysis. *Phys. Ther.* **96,** 876–887 (2016).

16. Lindstrøm, R., Schomacher, J., Farina, D., Rechter, L. & Falla, D. Association between neck muscle coactivation, pain, and strength in women with neck pain. *Man. Ther.* **16,** 80–86 (2011).

17. O'Leary, S., Falla, D., Elliott, J. M. & Jull, G. Muscle Dysfunction in Cervical Spine Pain: Implications for Assessment and Management. *J. Orthop. Sport. Phys. Ther.* **39,** 324–333 (2009).

18. Blizzard, L., Grimmer, K. A. & 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.* **81,** 1204–1210 (2000).

19. Phillips, B. A., Lo, S. K. & Mastaglia, F. L. Muscle force measured using 'break' testing with a hand-held myometer in normal subjects aged 20 to 69 years. *Arch. Phys. Med. Rehabil.* **81,** 653–661 (2000).

20. Silverman, J. L., Rodriquez, A. A. & Agre, J. C. Quantitative cervical flexor strength in healthy subjects and in subjects with mechanical neck pain. *Arch. Phys. Med. Rehabil.* **72,** 679–681 (1991).

21. Grimmer, K. Measuring the endurance capacity of the cervical short flexor muscle group. *Aust. J. Physiother.* **40,** 251–254 (1994).

22. Harris, K. D. *et al.* Reliability of a measurement of neck flexor muscle endurance. *Phys. Ther.* **85,** 1349–1355 (2005).

23. Jull, G. A., O'Leary, S. P. & Falla, D. L. Clinical Assessment of the Deep Cervical Flexor Muscles: The Craniocervical Flexion Test. *J. Manipulative Physiol. Ther.* **31,** 525–533 (2008).

24. Juul, T., Langberg, H., Enoch, F. & Søgaard, K. The intra- and inter-rater reliability of five clinical muscle performance tests in patients with and without neck pain. *BMC Musculoskelet. Disord.* **14,** 339 (2013).

25. Mokkink, L. B. *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.* **19,** 539–549 (2010).

26. Mokkink, L. B. *et al.* Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual. Life Res.* **18,** 313–333 (2009).

27. Terwee, C. B. *et al.* Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual. Life Res.* **21,** 651–657 (2012).

28. de Koning, C. H. P. *et al.* Clinimetric evaluation of methods to measure muscle functioning in patients with non-specific neck pain: a systematic review. *BMC Musculoskelet. Disord.* **9,** 142 (2008).

29. Lohr, K. *et al.* Evaluating Quality of Life and Health status instruments: development of scientific review criteria. *Clin. Ther.* **18,** 979–992 (1996).

30. Bot, S. D. M. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann. Rheum. Dis.* **63,** 335–341 (2004).

31. Moher, D. *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **4,** 1 (2015).

32. Mokkink, L. B. *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.* **63,** 737–745 (2010).

33. Terwee, C. B., Jansma, E. P., Riphagen, I. I. & De Vet, H. C. W. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual. Life Res.* **18,** 1115–1123 (2009).

34. Terwee, C. B. *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* **60,** 34–42 (2007).

35. Furlan, A. D., Pennick, V., Bombardier, C. & van Tulder, M. 2009 Updated Method Guidelines for Systematic Reviews in the Cochrane Back Review Group. *Spine (Phila. Pa. 1976).* **34,** 1929–1941 (2009).

TABLE 1. Search strategy in MEDLINE (via Pubmed).

1#	(((((((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 (health care)"[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 "standard error 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])))))
2#	(((((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 stabilizers[Title/Abstract] OR stabiliser[Title/Abstract] OR stabilisers[Title/Abstract] OR biofeedback[Title/Abstract] OR biofeedbacks[Title/Abstract] OR craniocervical flexion test[Title/Abstract] OR cranio-cervical flexion test[Title/Abstract] OR cranio cervical flexion test[Title/Abstract] OR cranio cervical flexion[Title/Abstract)))))
3#	((((((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]))
4#	1# AND 2# AND 3#

TABLE 2. Quality criteria for measurement properties.

Property	Rating	Quality criteria
Reliability		
Internal consistency	+	Cronbach's alpha(s) ≥ 0.70
	?	Cronbach's alpha not determined or unidimensionality unknown
	-	Cronbach's alpha(s) < 0.70
Reliability	+	ICC/ weighted Kappa ≥ 0.70 OR Pearson's r ≥ 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 ≤ SDC OR MIC equals or inside LoA
Validity		
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 ≥ 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 ≥ 0.70
	?	No convincing arguments that gold standard is "gold"
	-	Correlation with gold standard < 0.70
Responsiveness		
Responsiveness	+	Correlation with changes on instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 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

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

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

Legends: +, positive rating; -, negative rating; ?, indeterminate rating

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Page number
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	NA
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	9
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	NA
Support:			
Sources	5a	Indicate sources of financial or other support for the review	
Sponsor	5b	Provide name for the review funder and/or sponsor	NA
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	4, 5
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	5
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting,	6

		time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	6
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	6
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	7
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	7
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	7
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	6, 7
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	6, 7
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	8
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	8
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	8
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	8
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	8
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication	NA

		bias across studies, selective reporting within studies)	
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	NA

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-019486.R1
Article Type:	Protocol
Date Submitted by the Author:	05-Dec-2017
Complete List of Authors:	Araujo, Francisco Ferreira, Giovanni ; Universidade Federal de Ciencias da Saude de Porto Alegre, Physical Therapy Department Scholl Schell, Maurício ; Universidade Federal de Ciencias da Saude de Porto Alegre Castro, Marcelo; Neuromusculoskeletal Assessment and Clinical Biomechanics Laboratory (LaBClin) Silva, Marcelo; Universidade Federal de Ciencias da Saude de Porto Alegre, Physical Therapy Department Ribeiro, Daniel; School of Physiotherapy
Primary Subject Heading:	Rehabilitation medicine
Secondary Subject Heading:	Occupational and environmental medicine
Keywords:	Musculoskeletal disorders < ORTHOPAEDIC & TRAUMA SURGERY, REHABILITATION MEDICINE, SPORTS MEDICINE

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

Francisco Xavier de Araujo, PT, MSc[1,2]
franciscoxaraujo@gmail.com

Giovanni Esteves Ferreira, PT, MSc[1]
giovannieferreira@hotmail.com

Maurício Scholl Schell, PT[1]
mschell.fisio@gmail.com

Marcelo Peduzzi de Castro, PT, PhD[3]
marcelo.peduzzi.castro@gmail.com

Marcelo Faria Silva, PT, PhD[1]
marcelofs@ufcspa.edu.br

Daniel Cury Ribeiro, PT, PhD[4]
daniel.ribeiro@otago.ac.nz

[1] Physical Therapy Department, Graduate Program in Rehabilitation Sciences, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Brazil.

[2] Centro Universitário Ritter dos Reis (UniRitter), Laureate International Universities, Porto Alegre, Brazil.

[3] Neuromusculoskeletal Assessment and Clinical Biomechanics Laboratory (LaBClin), Florianópolis, Brazil.

[4] School of Physiotherapy, University of Otago, Dunedin, New Zealand.

ABSTRACT:

Introduction: Neck pain is the first 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 (PRISMA) 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 COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist. A structured narrative synthesis will be used for data analysis. Quantitative findings for each measurement property will be summarized. 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.

Trial registration number: This protocol has been registered with the International Prospective Register of Systematic Reviews (PROSPERO), registration number (CRD42017062175).

STRENGTHS AND LIMITATIONS:

- 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 recognized, validated COSMIN 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 PRISMA 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.

1. INTRODUCTION:

Neck pain is the first cause of years lived with disability worldwide, and it accounts for high economic and societal burden.[1,2] In the general population, 16.7 to 75.1% of adults will develop an episode of neck pain in any given year.[3,4] Patients may present recurrent neck pain,[5,6] and the prognosis of recovery is poor.[5] Between 50 and 75% of people who experienced neck pain still present with symptoms one to five years after onset of symptoms.[6]

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

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.[17] Several tests have been designed to evaluate different aspects of neck muscle performance.[18–23] 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 (e.g. *sternocleidomastoid*, *anterior scalene*) during an active craniocervical flexion in supine lying.[23] This test has been widely used in clinical practice to assess function of deep neck flexor muscles.[8,23,24]

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 mmHg.[24] 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.[23] The performance of the test is assessed through the following components: how well the patient performs the active head nodding and achieves that by contractingthe 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).[8,23]

Ideally, the measurement properties (e.g. 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.[25] Selecting instruments with proper measurement properties is fundamental for well-conduced clinical trials.[26] Hence, systematic reviews of measurement properties are useful for identifying instruments and tools with the highest reliability, validity and responsiveness scores.[27]

A previous systematic review[28] evaluated the measurement properties of methods to measure muscle function in patients with non-specific neck pain. In this review, the intra-observer reliability was the only measurement property assessed.[28] This review was conducted over 10 years ago, included only four studies and used a checklist adapted from two previous studies.[29,30] 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 (i.e. Consensus-based Standards for the selection of health Measurement Instruments – COSMIN).[25] 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 summarize the quality of the

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

2. METHODS:

2.1. 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).[31] The systematic review has been registered with PROSPERO (CRD42017062175).

2.2. Eligibility criteria

We will include studies if they meet the following criteria:

- Studies of any design (e.g. Cross-sectional studies or randomized clinical trials);
- Articles that have investigated and reported at least one measurement property (i.e. 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 old;
- 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.

2.3. Outcomes

All measurement properties will be considered as outcomes in this

systematic review. We will adopt the COSMIN terminology and definitions of measurement properties.[32] Reliability is defined as the extent to which scores for patients who have not changed are the same for repeated measurement under several conditions; 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.[32] 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.[32] For the purposes of this review, we will include all outcome measures used assessing psychometric properties that are reported by included studies.

2.4. 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 Direct. All databases will be searched from their inception to present time using a published search filter for finding studies on measurement properties.[33] MEDLINE full-search strategy is described in Table 1.

2.5. 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 solved by consensus. The following information will be extracted from the included studies: study design, sample characteristics, measurement properties (e.g. validity, reliability or responsiveness) assessed by included studies, craniocarvical flexion test

procedures and results of the measurements properties.

2.6 Risk of bias within included studies

Two reviewers (FXA and GEF) will independently rate the risk of bias of individual studies using the COSMIN checklist.[25] The COSMIN checklist is a validated critical appraisal tool designed for the systematic evaluation of the methodological quality of studies on the measurement properties.[27] The checklist consists of nine domains concerning measurement properties. The number of items for each domain varies from 5 to 18. Each item deals with design characteristics and statistical methods used and reported by authors. Each item will be scored based on a four-point rating scale as "excellent", "good", "fair", or "poor". The lowest rating score of a domain will be used for attributing the quality score for that specific domain. 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).

2.7 Synthesis of results:

A structured narrative synthesis will be used for data analysis. Quantitative findings for each measurement property will be summarized. 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.[34] The criteria used to assign levels of evidence for the quality (Table 3) of each measurement property will follow the framework proposed by the Cochrane Back and Neck Review Group.[35]

Table 3

We will conduct a narrative synthesis of subgroups, if applicable, based on the sample characteristics (i.e. asymptomatic or symptomatic), and type of disorder (e.g. 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).

3. 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 also possibly presented at national and international meetings.

Authors' contributions: FXA is the leading researcher, responsible for conceiving the study, and designing the protocol. 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. FXA and MSS will screen title, abstracts and full text for eligibility. MSS and GEF will extract data. FXA and GEF will rate the methodological quality of individual studies. MPC will will resolve any disagreement between reviewers. GEF and MSS will provide the statistical analysis plan of the study and will conduct the data analysis. FXA, MSS, GEF will write the first version of the paper. MPC, DCR and MFS will provide critical revision of the paper. All authors read and provided final approval of this protocol to be published.

Funding statement: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. Araujo FX and Schell MS are supported by a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Brazil). However, CAPES is not involved in any other aspect of this study protocol.

Competing interest: None declared.

References:

1. Vos, T. *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* **388,** 1545–1602 (2016).

2. Hoy, D. *et al.* The global burden of neck pain: estimates from the Global Burden of Disease 2010 study. *Ann. Rheum. Dis.* **73,** 1309–1315 (2014).

3. Hogg-Johnson, S. *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.* **32,** 46–60 (2009).

4. Fejer, R., Kyvik, K. O. & Hartvigsen, J. The prevalence of neck pain in the world population: A systematic critical review of the literature. *Eur. Spine J.* **15,** 834–848 (2006).

5. Hush, J. M., Lin, C. C., Michaleff, Z. A., Verhagen, A. & Refshauge, K. M. Prognosis of acute idiopathic neck pain is poor: A systematic review and meta-analysis. *Arch. Phys. Med. Rehabil.* **92,** 824–829 (2011).

6. Carroll, L. J. *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.* **32,** 87–96 (2009).

7. Falla, D. & Farina, D. Neuromuscular adaptation in experimental and clinical neck pain. *J. Electromyogr. Kinesiol.* **18,** 255–261 (2008).

8. 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.* **25,** 43–47 (2016).

9. Oliveira, A. C. & Silva, A. G. Neck muscle endurance and head posture: A comparison between adolescents with and without neck pain. *Man. Ther.* **22,**

62–67 (2016).

10. Falla, D., Jull, G., Russell, T., Vicenzino, B. & Hodges, P. Effect of Neck Exercise on Sitting Posture in Patients With Chronic Neck Pain. *Phys. Ther.* **87,** 408–417 (2007).

11. Treleaven, J., Chen, X., Sarig Bahat, H. & Treleaven, Julia; Xiaoqi, Chen; Bahat, H. S. Factors associated with cervical kinematic impairments in patients with neck pain. *Man. Ther.* **22,** 109–115 (2016).

12. Falla, D., Jull, G. & Hodges, P. W. Feedforward activity of the cervical flexor muscles during voluntary arm movements is delayed in chronic neck pain. *Exp. Brain Res.* **157,** 43–48 (2004).

13. de Zoete, R. M. J., Osmotherly, P. G., Rivett, D. A., Farrell, S. F. & Snodgrass, S. J. Sensorimotor Control in Individuals With Idiopathic Neck Pain and Healthy Individuals: A Systematic Review and Meta-Analysis. *Arch. Phys. Med. Rehabil.* **98,** 1257–1271 (2017).

14. de Vries, J. *et al.* Joint position sense error in people with neck pain: A systematic review. *Man. Ther.* **20,** 736–744 (2015).

15. Stanton, T. R., Leake, H. B., Chalmers, K. J. & Moseley, G. L. Evidence of Impaired Proprioception in Chronic, Idiopathic Neck Pain: Systematic Review and Meta-Analysis. *Phys. Ther.* **96,** 876–887 (2016).

16. Lindstrøm, R., Schomacher, J., Farina, D., Rechter, L. & Falla, D. Association between neck muscle coactivation, pain, and strength in women with neck pain. *Man. Ther.* **16,** 80–86 (2011).

17. O'Leary, S., Falla, D., Elliott, J. M. & Jull, G. Muscle Dysfunction in Cervical Spine Pain: Implications for Assessment and Management. *J. Orthop. Sport. Phys. Ther.* **39,** 324–333 (2009).

18. Blizzard, L., Grimmer, K. A. & 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.* **81,** 1204–1210 (2000).

19. Phillips, B. A., Lo, S. K. & Mastaglia, F. L. Muscle force measured using 'break' testing with a hand-held myometer in normal subjects aged 20 to 69 years. *Arch. Phys. Med. Rehabil.* **81,** 653–661 (2000).

20. Silverman, J. L., Rodriquez, A. A. & Agre, J. C. Quantitative cervical flexor strength in healthy subjects and in subjects with mechanical neck pain. *Arch. Phys. Med. Rehabil.* **72,** 679–681 (1991).

21. Grimmer, K. Measuring the endurance capacity of the cervical short flexor muscle group. *Aust. J. Physiother.* **40,** 251–254 (1994).

22. Harris, K. D. *et al.* Reliability of a measurement of neck flexor muscle endurance. *Phys. Ther.* **85,** 1349–1355 (2005).

23. Jull, G. A., O'Leary, S. P. & Falla, D. L. Clinical assessment of the deep cervical flexor muscles: the craniocervical flexion test. *J. Manipulative Physiol. Ther.* **31,** 525–533 (2008).

24. Juul, T. *et al.* The intra- and inter-rater reliability of five clinical muscle performance tests in patients with and without neck pain. *BMC Musculoskelet. Disord.* **14,** 339 (2013).

25. Mokkink, L. B. *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.* **19,** 539–549 (2010).

26. Mokkink, L. B. *et al.* Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual. Life Res.* **18,** 313–332 (2009).

27. Terwee, C. B. *et al.* Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual. Life Res.* **21,** 651–657 (2012).

28. de Koning, C. H. P. *et al.* Clinimetric evaluation of methods to measure muscle functioning in patients with non-specific neck pain: a systematic review. *BMC Musculoskelet. Disord.* **9,** 142 (2008).

29. Lohr, K. *et al.* Evaluating Quality of Life and Health status instruments: development of scientific review criteria. *Clin. Ther.* **18,** 979–992 (1996).

30. Bot, S. D. M. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann. Rheum. Dis.* **63,** 335–341 (2004).

31. Moher, D. *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **4,** 1 (2015).

32. Mokkink, L. B. *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.* **63,** 737–745 (2010).

33. Terwee, C. B., Jansma, E. P., Riphagen, I. I. & De Vet, H. C. W. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual. Life Res.* **18,** 1115–1123 (2009).

34. Terwee, C. B. *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* **60,** 34–42 (2007).

35. Furlan, A. D., Pennick, V., Bombardier, C. & van Tulder, M. 2009 Updated Method Guidelines for Systematic Reviews in the Cochrane Back Review Group. *Spine (Phila. Pa. 1976).* **34,** 1929–1941 (2009).

TABLE 1. Search strategy in MEDLINE (via Pubmed).

1#	(((((((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 (health care)"[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 "standard error 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])))))
2#	(((((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 stabilizers[Title/Abstract] OR stabiliser[Title/Abstract] OR stabilisers[Title/Abstract] OR biofeedback[Title/Abstract] OR biofeedbacks[Title/Abstract] OR craniocervical flexion test[Title/Abstract] OR cranio-cervical flexion test[Title/Abstract] OR cranio cervical flexion test[Title/Abstract] OR cranio cervical flexion[Title/Abstract)))))
3#	((((((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]))
4#	1# AND 2# AND 3#

TABLE 2. Quality criteria for measurement properties.

Property	Rating	Quality criteria
Reliability		
Internal consistency	+	Cronbach's alpha(s) ≥ 0.70
	?	Cronbach's alpha not determined or unidimensionality unknown
	-	Cronbach's alpha(s) < 0.70
Reliability	+	ICC/ weighted Kappa ≥ 0.70 OR Pearson's r ≥ 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 ≤ SDC OR MIC equals or inside LoA
Validity		

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 ≥ 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 ≥ 0.70
	?	No convincing arguments that gold standard is “gold”
	-	Correlation with gold standard < 0.70
Responsiveness		
Responsiveness	+	Correlation with changes on instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 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

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

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

Legends: +, positive rating; -, negative rating; ?, indeterminate rating

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Page number
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	NA
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	9
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	NA
Support:			
Sources	5a	Indicate sources of financial or other support for the review	
Sponsor	5b	Provide name for the review funder and/or sponsor	NA
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	4, 5
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	5
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting,	6

		time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	6
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	6
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	7
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	7
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	7
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	6, 7
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	6, 7
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	8
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	8
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	8
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	8
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	8
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication	NA

		bias across studies, selective reporting within studies)	
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	NA

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-019486.R2
Article Type:	Protocol
Date Submitted by the Author:	17-Jan-2018
Complete List of Authors:	Araujo, Francisco; Universidade Federal de Ciencias da Saude de Porto Alegre; Centro Universitario Ritter dos Reis - Campus Canoas Ferreira, Giovanni ; Universidade Federal de Ciencias da Saude de Porto Alegre, Physical Therapy Department Scholl Schell, Maurício ; Universidade Federal de Ciencias da Saude de Porto Alegre Castro, Marcelo; Neuromusculoskeletal Assessment and Clinical Biomechanics Laboratory (LaBClin) Silva, Marcelo; Universidade Federal de Ciencias da Saude de Porto Alegre, Physical Therapy Department Ribeiro, Daniel; University of Otago, School of Physiotherapy
Primary Subject Heading:	Rehabilitation medicine
Secondary Subject Heading:	Occupational and environmental medicine
Keywords:	Musculoskeletal disorders < ORTHOPAEDIC & TRAUMA SURGERY, REHABILITATION MEDICINE, SPORTS MEDICINE

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

Francisco Xavier de Araujo, PT, MSc[1,2]
franciscoxaraujo@gmail.com

Giovanni Esteves Ferreira, PT, MSc[1]
giovannieferreira@hotmail.com

Maurício Scholl Schell, PT[1]
mschell.fisio@gmail.com

Marcelo Peduzzi de Castro, PT, PhD[3]
marcelo.peduzzi.castro@gmail.com

Marcelo Faria Silva, PT, PhD[1]
marcelofs@ufcspa.edu.br

Daniel Cury Ribeiro, PT, PhD[4]
daniel.ribeiro@otago.ac.nz

1 Physical Therapy Department, Graduate Program in Rehabilitation Sciences, Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Brazil.

2 Centro Universitário Ritter dos Reis (UniRitter), Laureate International Universities, Porto Alegre, Brazil.

3 Neuromusculoskeletal Assessment and Clinical Biomechanics Laboratory (LaBClin), Florianópolis, Brazil.

4 School of Physiotherapy, University of Otago, Dunedin, New Zealand.

Corresponding author: Daniel Cury Ribeiro, Centre for Health, Activity and Rehabilitation Research (CHARR), School of Physiotherapy – University of Otago PO Box 56, Dunedin 9054, New Zealand, E-mail: Daniel.ribeiro@otago.ac.nz, T: +64 03 479 7455

Keywords: feedback, neck pain, rehabilitation

Word count: 1614

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 (PRISMA) 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 (COSMIN) checklist. A structured narrative synthesis will be used for data analysis. Quantitative findings for each measurement property will be summarized. 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.

Trial registration number: This protocol has been registered with the International Prospective Register of Systematic Reviews (PROSPERO), registration number (CRD42017062175).

STRENGTHS AND LIMITATIONS:

- 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 recognized, validated COSMIN 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 PRISMA 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.

1. INTRODUCTION:

Neck pain is the leading cause of years lived with disability worldwide, and it accounts for high economic and societal burden.[1,2] In the general population, 16.7 to 75.1% of adults will develop an episode of neck pain in any given year.[3,4] Patients may present recurrent neck pain,[5,6] and the prognosis of recovery is poor.[5] Between 50 and 75% of people who experienced neck pain still present with symptoms one to five years after onset of symptoms.[6]

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

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.[17] Several tests have been designed to evaluate different aspects of neck muscle performance.[18–23] 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 (e.g. *sternocleidomastoid*, *anterior scalene*) during an active craniocervical flexion in supine lying.[23] This test has been widely used in clinical practice to assess function of deep neck flexor muscles.[8,23,24]

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 mmHg.[24] 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.[23] 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).[8,23]

Ideally, the measurement properties (e.g. 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.[25] Selecting instruments with proper measurement properties is fundamental for well-conduced clinical trials.[26] Hence, systematic reviews of measurement properties are useful for identifying instruments and tools with the highest reliability, validity and responsiveness scores.[27]

A previous systematic review[28] evaluated the measurement properties of methods to measure muscle function in patients with non-specific neck pain. In this review, the intra-observer reliability was the only measurement property assessed.[28] This review was conducted over 10 years ago, included only four studies and used a checklist adapted from two previous studies.[29,30] 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 (i.e. Consensus-based Standards for the selection of health Measurement Instruments – COSMIN).[25] 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 summarize the quality of the measurement properties of the craniocervical flexion test for assessing deep cervical flexor muscles.

2. METHODS:

2.1. 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).[31] The systematic review has been registered with PROSPERO (CRD42017062175).

2.2. Eligibility criteria

We will include studies if they meet the following criteria:

- Articles that have investigated and reported at least one measurement property (i.e. 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 old;
- 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.

2.3. Outcomes

All measurement properties will be considered as outcomes in this systematic

review. We will adopt the COSMIN terminology and definitions of measurement properties.[32] 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.[32] 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.[32] For the purposes of this review, we will include all outcome measures used assessing psychometric properties that are reported by included studies.

2.4. 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 Direct. All databases will be searched from their inception to present time using a published search filter for finding studies on measurement properties.[33] MEDLINE full-search strategy is described in Table 1.

2.5. 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 solved by consensus. The following information will be extracted from the included studies: study design, sample characteristics, measurement properties (e.g. validity, reliability or responsiveness) assessed by included studies, craniocarvical flexion test procedures and results of the measurements properties.

2.6 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 checklist.[34] The COSMIN checklist is a validated critical appraisal tool designed for the systematic evaluation of the methodological quality of studies on the measurement properties.[27] 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).

2.6. Synthesis of results:

A structured narrative synthesis will be used for data analysis. Quantitative findings for each measurement property will be summarized. 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.[35] 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.[36]

We will conduct a narrative synthesis of subgroups, if applicable, based on the sample characteristics (i.e. asymptomatic or symptomatic), and type of disorder (e.g. 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).

3. 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 also possibly presented at national and international meetings.

Authors' contributions: FXA is the leading researcher, responsible for conceiving the study, and designing the protocol. 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. FXA and MSS will screen title, abstracts and full text for eligibility. MSS and GEF will extract data. FXA and GEF will rate the methodological quality of individual studies. MPC will will resolve any disagreement between reviewers. GEF and MSS will provide the statistical analysis plan of the study and will conduct the data analysis. FXA, MSS, GEF will write the first version of the paper. MPC, DCR and MFS will provide critical revision of the paper. All authors read and provided final approval of this protocol to be published.

Funding statement: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors. Araujo FX and Schell MS are supported by a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Brazil). However, CAPES is not involved in any other aspect of this study protocol.

Competing interest: None declared.

References:

1 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. doi:10.1016/S0140-6736(16)31678-6

2 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. doi:10.1136/annrheumdis-2013-204431

3 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. doi:10.1016/j.jmpt.2008.11.013

4 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. doi:10.1007/s00586-004-0864-4

5 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. doi:10.1016/j.apmr.2010.12.025

6 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. doi:10.1016/j.jmpt.2008.11.013

7 Falla D, Farina D. Neuromuscular adaptation in experimental and clinical neck pain. *J Electromyogr Kinesiol* 2008;**18**:255–61. doi:10.1016/j.jelekin.2006.11.001

8 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. doi:10.1016/j.math.2016.05.336

9 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. doi:10.1016/j.math.2015.10.002

10 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. doi:10.2522/ptj.20060009

11 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. doi:10.1016/j.math.2015.10.015

12 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. doi:10.1007/s00221-003-1814-9

13 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. doi:10.1016/j.apmr.2016.09.121

14 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. doi:10.1016/j.math.2015.04.015

15 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. doi:10.2522/ptj.20150241

16 Lindstrøm 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. doi:10.1016/j.math.2010.07.006

17 O'Leary S, Falla D, Elliott JM, *et al.* Muscle Dysfunction in Cervical Spine Pain: Implications for Assessment and Management. *J Orthop Sport Phys Ther*

2009;**39**:324–33. doi:10.2519/jospt.2009.2872

18 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. doi:10.1053/apmr.2000.7168

19 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. doi:10.1053/mr.2000.4413

20 Silverman JL, Rodriquez 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. doi:0003-9993(91)90214-4 [pii]

21 Grimmer K. Measuring the endurance capacity of the cervical short flexor muscle group. *Aust J Physiother* 1994;**40**:251–4. doi:10.1016/S0004-9514(14)60461-X

22 Harris KD, Heer DM, Roy TC, *et al.* Reliability of a measurement of neck flexor muscle endurance. *Phys Ther* 2005;**85**:1349–55. doi:10.1054/math.1997.0298

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. doi:10.1016/j.jmpt.2008.08.003

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. doi:10.1186/1471-2474-14-339

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. doi:10.1007/s11136-010-9606-8

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**:313–32. doi:10.1007/s11136-009-9451-9

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. doi:10.1007/s11136-011-9960-1

28 de Koning CHP, Heuvel SP van den, 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. doi:10.1186/1471-2474-9-142

29 Lohr K, Aaronson N, 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 SDM. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004;**63**:335–41. doi:10.1136/ard.2003.007724

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. doi:10.1186/2046-4053-4-1

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. doi:10.1016/j.jclinepi.2010.02.006

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. doi:10.1007/s11136-009-9528-5

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;**0**. doi:10.1007/s11136-017-1765-4

35 Terwee CB, Bot SDM, de Boer MR, *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;**60**:34–42. doi:10.1016/j.jclinepi.2006.03.012

36 Furlan AD, Pennick V, Bombardier C, *et al.* 2009 Updated Method Guidelines for Systematic Reviews in the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 2009;**34**:1929–41. doi:10.1097/BRS.0b013e3181b1c99f

TABLE 1. Search strategy in MEDLINE (via Pubmed).

1#	(((((((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 (health care)"[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 "standard error 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])))))
2#	(((((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 stabilizers[Title/Abstract] OR stabiliser[Title/Abstract] OR stabilisers[Title/Abstract] OR biofeedback[Title/Abstract] OR biofeedbacks[Title/Abstract] OR craniocervical flexion test[Title/Abstract] OR cranio-cervical flexion test[Title/Abstract] OR cranio cervical flexion test[Title/Abstract] OR cranio cervical flexion[Title/Abstract)))))
3#	((((((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]))
4#	1# AND 2# AND 3#

TABLE 2. Quality criteria for measurement properties.

Property	Rating	Quality criteria
Reliability		
Internal consistency	+	Cronbach's alpha(s) ≥ 0.70
	?	Cronbach's alpha not determined or unidimensionality unknown

	-	Cronbach's alpha(s) < 0.70
Reliability	+	ICC/ weighted Kappa ≥ 0.70 OR Pearson's r ≥ 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 ≤ SDC OR MIC equals or inside LoA
Validity		
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 ≥ 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 ≥ 0.70
	?	No convincing arguments that gold standard is "gold"
	-	Correlation with gold standard < 0.70
Responsiveness		
Responsiveness	+	Correlation with changes on instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 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

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

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

Legends: +, positive rating; -, negative rating; ?, indeterminate rating

PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Page number
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	1
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	NA
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	2
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	9
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	NA
Support:			
Sources	5a	Indicate sources of financial or other support for the review	
Sponsor	5b	Provide name for the review funder and/or sponsor	NA
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	4, 5
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	5
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting,	6

		time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	
Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	6
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	6
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	7
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	7
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	7
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	6, 7
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	6, 7
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	8
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	8
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	8
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	8
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	8
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication	NA

		bias across studies, selective reporting within studies)	
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	NA

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.