View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by VU Research Portal

VU UNIVERSITEIT AMSTERDAM

VU Research Portal

Treatment based classification systems for patients with non-specific neck pain

Maissan, Francois; Pool, Jan; de Raaij, Edwin; Wittink, Harriet; Ostelo, Raymond

published in
Musculoskeletal Science and Practice
2020

DOI (link to publisher)
10.1016/j.msksp.2020.102133

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Link to publication in VU Research Portal

citation for published version (APA)
Maissan, F., Pool, J., de Raaij, E., Wittink, H., & Ostelo, R. (2020). Treatment based classification systems for patients with non-specific neck pain: A systematic review. *Musculoskeletal Science and Practice*, *47*, 1-11. [102133]. https://doi.org/10.1016/j.msksp.2020.102133

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 23. May. 2021

Musculoskeletal Science and Practice 47 (2020) 102133

Contents lists available at ScienceDirect

Musculoskeletal Science and Practice

journal homepage: www.elsevier.com/locate/msksp

Systematic Review

Treatment based classification systems for patients with non-specific neck pain. A systematic review.

Francois Maissan [a,b,c,*], Jan Pool [a], Edwin de Raaij [a,b,c], Harriet Wittink [a], Raymond Ostelo [b,c]

[a] Research Group Lifestyle and Health, HU University of Applied Sciences Utrecht, Utrecht, the Netherlands
[b] Department of Health Sciences, VU University, Amsterdam, the Netherlands
[c] Department of Epidemiology and Biostatistics, Amsterdam UMC, Amsterdam, Amsterdam Movement Sciences, the Netherlands

A B S T R A C T

Objective: We aimed to identify published classification systems with a targeted treatment approach (treatment-based classification systems (TBCSs)) for patients with non-specific neck pain, and assess their quality and effectiveness.
Design: Systematic review.
Data sources: MEDLINE, CINAHL, EMBASE, PEDro and the grey literature were systematically searched from inception to December 2019.
Study appraisal and synthesis: The main selection criterium was a TBCS for patients with non-specific neck pain with physiotherapeutic interventions. For data extraction of descriptive data and quality assessment we used the framework developed by Buchbinder et al. We considered as score of ≤3 as low quality, a score between 3 and 5 as moderate quality and a score ≥5 as good quality.
To assess the risk of bias of studies concerning the effectiveness of TBCSs (only randomized clinical trials (RCTs) were included) we used the PEDro scale. We considered a score of ≥ six points on this scale as low risk of bias.
Results: Out of 7664 initial references we included 13 studies. The overall quality of the TBCSs ranged from low to moderate. We found two RCTs, both with low risk of bias, evaluating the effectiveness of two TBCSs compared to alternative treatments. The results showed that both TBCSs were not superior to alternative treatments.
Conclusion: Existing TBCSs are, at best, of moderate quality. In addition, TBCSs were not shown to be more effective than alternatives. Therefore using these TBCSs in daily practice is not recommended.

1. Introduction

Neck pain is the fourth major cause of disability worldwide (Vos et al., 2012). In 2010, the proportion of Years Lived with Disability (YLDs) from all musculoskeletal disorders (MSK) was 21.3% of the total proportion of YLDs. Neck pain was responsible for 20.1% of the total proportion due to MSK (March et al., 2014). In 2015, more than a third of a billion people worldwide had neck pain of more than 3 months duration (Hurwitz et al., 2018).

At least six Cochrane reviews focussing on physiotherapy interventions for patients with neck pain reported inconclusive evidence for their effectiveness (Ezzo et al., 2007; Graham et al., 2008; A. Gross et al., 2012; A. R. Gross et al., 2004; Kay et al., 2012; Monticone et al., 2015). This may be due to heterogeneity of the study population.

One method to deal with this heterogeneity is to match treatment more specifically to subgroups of patients with "non-specific pain". Matching groups of patients with the most appropriate treatment for their risk profile or with treatment that they are most likely to benefit from, i.e. stratified or matched care (Linton et al., 2018), has been a research priority for the last few years (Foster et al., 2011) as it might increase the effectiveness of the interventions (Coupe et al., 2016). However, studies have described the lack of evidence of accurate and reproducible classification systems that aim to subgroup patients into distinct subgroups with a matching intervention (treatment-based classification systems (TBCSs)) (Damgaard et al., 2013; Tsakitzidis et al., 2013) (Fairbank et al., 2011).

The development of a TBCS can be achieved through a (clinical) judgement approach and/or a statistical approach (Riddle, 1998). The judgment approach relies on three types of judgment: (1) traditional custom (to identify the variables in the literature that have been suggested to be the most important); (2) conventional wisdom (common, but unpublished, beliefs of the clinical community); and (3) personal experience (the developers' own clinical experiences). The statistical approach relies on one, or a combination of, statistical procedures (e.g. cluster analysis) designed to identify variables that can be used to distinguish subgroups of patients.

* Corresponding author. University of Applied Sciences Utrecht, Heidelberglaan 7, 3584, CS, Utrecht, the Netherlands.
E-mail address: francois.maissan@hu.nl (F. Maissan).

https://doi.org/10.1016/j.msksp.2020.102133
Received 11 February 2019; Received in revised form 2 February 2020; Accepted 15 February 2020
Available online 20 February 2020
2468-7812/© 2020 Elsevier Ltd. All rights reserved.

Our overall aim is to gain more insight into existing TBCSs and their potential for treatment in people with non-specific neck pain. Therefore, we aim to identify published classification systems with a targeted treatment approach (TBCSs) for patients with non-specific neck pain, and assess their quality and effectiveness.

2. Methods

2.1. Design

This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009) and registered in the international prospective register of systematic reviews PROSPERO (CRD 42018087763).

2.2. Data sources and searches

A sensitive electronic search was completed in collaboration with a medical information specialist, in MEDLINE, CINAHL, EMBASE and PEDro. All databases were searched from inception to December 2019. A MEDLINE search of first authors or the name of the included TBCSs was performed, to include any additional published research. To identify grey literature, we searched the following electronic sources: DART-Europe E-theses Portal, Open Access Theses and Dissertations, Networked Digital Library of Theses and Dissertations (NDLTD), ClinicalTrials.gov and WHO International Clinical Trials Registry Platform (ICTRP) The search strategies for PUBMED, CINAHL, EMBASE, PEDro and the grey literature are described in Appendix 1.

2.3. Study selection

We defined the following selection criteria:

1) *Design.* For the description of TBCSs we included studies on the development of TBCSs. To assess quality of the research into the TBCSs, we included, in addition to studies on the development, studies that investigated the quality of the TBCS such as reliability studies. To assess the effectiveness we included only Randomized Control Trials (RCTs) comparing TBCSs to control conditions or usual care. Case reports and case series were excluded for this review.
2) *Population.* Studies were eligible when including adult patients (>18 years of age) with non-specific neck pain. Non-specific neck pain was defined as pain (with or without radiation) located in the cervical spine and/or occiput region and/or cervicothoracic junction and muscles originating from the cervical region acting on the head and shoulders, without underlying pathology (Hogg-Johnson et al., 2008). A study was excluded if the study was performed in patients with whiplash, headache of non-cervicogenic origin or in patients with temporomandibular joint dysfunctions only.
3) *Intervention.* TBCSs should include physiotherapeutic interventions. Chiropractic care or osteopathy were not considered to be physiotherapeutic interventions.

Two reviewers (FM, JP) independently reviewed the titles, abstracts and the papers retrieved for full text based on the inclusion and exclusion criteria. Differences were discussed until consensus was reached. In case of persistent disagreement, a third independent reviewer (HW) was consulted.

2.4. Data extraction and quality assessment

2.4.1. Description

We used a framework, used in multiple reviews, to describe the characteristics of a classification system (Buchbinder et al., 1996). This framework consists of seven items: *purpose* of the study; *method of development* (i.e. based on a clinical judgment or using statistical methods); *domain of interest* (patient population and setting); *specific exclusions* for patients (i.e. exclusion criteria), one or more *categories* to name the specific subgroup; *criteria* used to assign patients to the subgroup; and, finally, *treatment* matching the categories.

Table 1
Criteria used to appraise the methodological quality of treatment based classification systems (adapted from Buchbinder et al.).

Criteria	Description
Purpose	Is the purpose, population and setting clearly specified?
Content validity	Is the domain and all specific exclusions from the domain clearly specified?
	Are all relevant categories included?
	Is the breakdown of categories appropriate, considering the purpose?
	Are the categories mutually exclusive?
	Was the method of development appropriate?
	If multiaxial, are criteria of content validity satisfied for each additional axis?
Face validity	Is the nomenclature used to label the categories satisfactory?
	Are the terms used based upon empirical (directly observable) evidence?
	Are the criteria for determining inclusion into each category clearly specified?
	If yes do these criteria appear reasonable?
	Have the criteria been demonstrated to have reliability or validity?
	Are the definitions of criteria clearly specified?
	If multiaxial are criteria of face validity satisfied for each additional axis?
Feasibility	Is the classification simple to understand?
	Is classification easy to perform?
	Does it rely on clinical examination alone?
	Are special skills, tools and/or training required?
	How long does it take to perform?
Construct validity	Does it discriminate between entities that are thought to be different in a way appropriate for the purpose?
	Does it perform satisfactorily when compared to other classification systems which classify the same domain?
Reliability	Does the classification system provide consistent results when classifying the same conditions?
	Is the intra-observer and inter-observer reliability satisfactory?
Generalizability	Has it been used in other studies and/or settings?

2.4.2. Quality

A scoring system, using seven criteria, was developed to critically appraise the quality of the TBCSs: purpose, content validity, face validity, feasibility, construct validity, (diagnostic) reliability, and generalizability (Buchbinder et al., 1996), see Table 1. The overall inter-rater reliability of the Buchbinder scale had an Intraclass Correlation Coefficient (ICC) of 0.82 (Buchbinder et al., 1996). In this article, we will refer to these seven criteria as the "Buchbinder appraisal scale". A score of one point (= yes) was awarded for meeting a criterion, a half point for partially meeting a criterion, and zero points (= no) for not meeting a criterion or being unable to score due to lack of evidence or information. Scores were summed up and in total the score could range from 0 to 7.

Two authors (FM,JP) independently extracted the data, using the guidance as described previously (Riddle, 1998). We pilot tested the data extraction on two articles not selected for this review.

Regarding the reliability criterion of the Buchbinder appraisal scale, the inter and/or intra reliability had to be weighted. For this weighting we used the following classification for interpretation of Cohen's kappa values: 0–0.4 slight to fair (= score of "0" on the Buchbinder scale), 0.4–0.8 moderate to substantial (= score of "0.5" on the Buchbinder scale) and >0.8 almost perfect (= score of "1" on the Buchbinder scale) (Landis and Koch, 1977). For the ICC we used 0–0.5 as poor (= score of "0" on the Buchbinder scale), 0.5–0.75 as moderate (score of "0.5" on the Buchbinder scale) and >0.75 as good (= score of "1" on the Buchbinder scale) (Portney and Watkins, 2000).

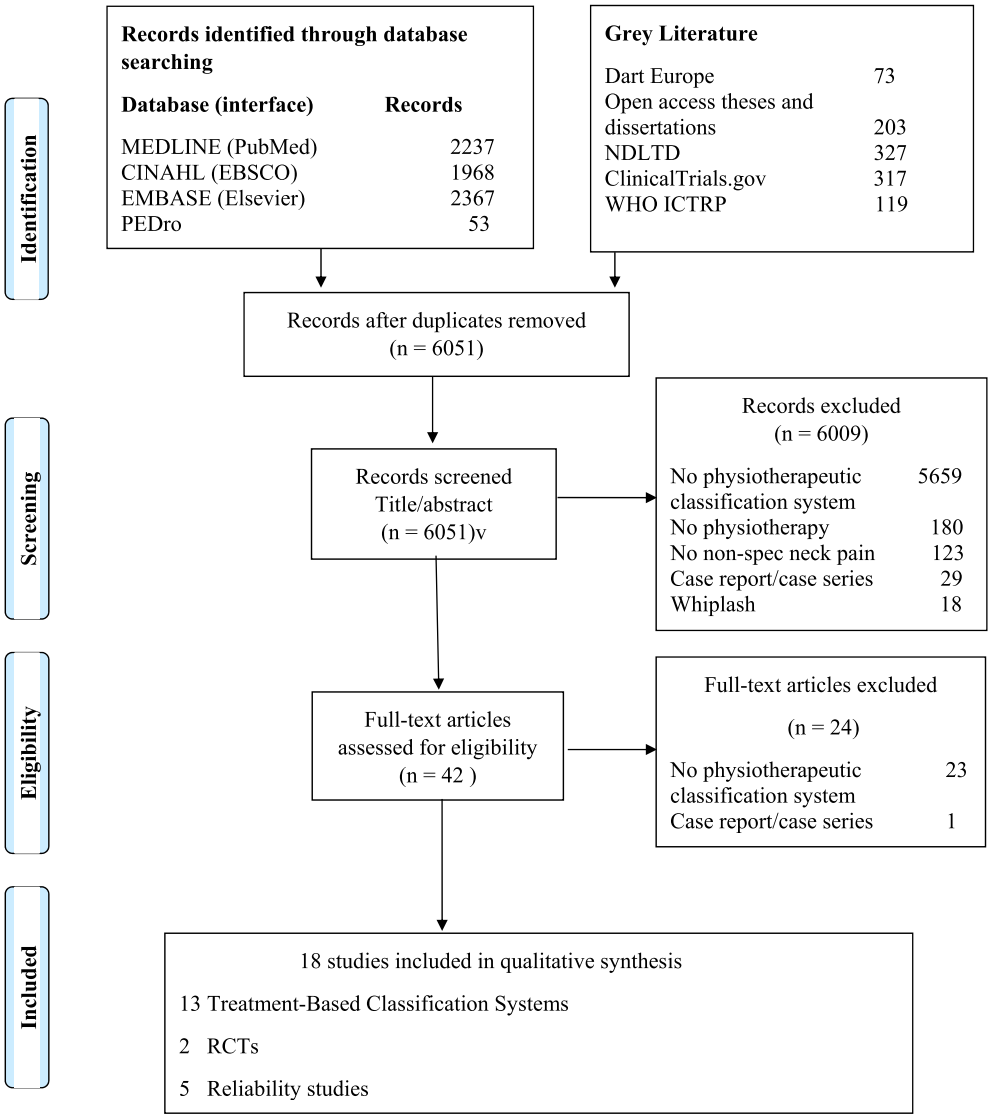

Fig. 1. Flowchart of articles reviewed.

2.4.3. *Effectiveness*

We assessed the risk of bias of the RCTs using the PEDro scale (www.pedro.org.au) (de Morton, 2009). The PEDro scale has moderate-to-good reliability with an ICC of 0.68 (95% confidence interval (CI) 0.57 to 0.76) (Maher et al., 2003). We considered RCTs with a score of ≥ six points on the PEDro scale as studies with a low risk of bias (Veerbeek et al., 2011).

2.5. *Data synthesis and analysis*

We considered the quality of a TBCS on the Buchbinder scale to be low if the score was ≤3, to be moderate if the score was between 3 and 5, and to be good as the score was ≥5. We described the characteristics of the TBCSs included and their quality narratively. Concerning the effectiveness, we assessed the between group differences on the primary outcomes (pain and/or disability), that is, between the TBCS under investigation and the comparator intervention. The clinical relevance was assessed on the basis of the Minimal Important Change (MIC) if it was known for the used outcome measures.

3. Results

3.1. *Search results for TBCSs*

The literature search retrieved 7664 studies: after removing duplicates, 6051 remained for further screening. Fig. 1 describes the screening process. No additional studies from the grey literature were included. Eighteen studies were included in the qualitative syntheses, i.e. the description of TBCSs and their quality (Bier et al., 2017; Childs et al., 2004; Clare et al., 2004; Clare et al., 2005; Cleland et al., 2006; Cleland et al., 2007; Cleland et al., 2010; Dewitte et al., 2014; Fernandez-de-las-Penas et al., 2011; Fritz and Brennan, 2007; Hanney et al., 2013; Hefford, 2008; Kjellman and Oberg, 2002; Lee et al., 2017; Puentedura et al., 2012; Raney et al., 2009; Saavedra-Hernandez et al., 2011; Wang et al., 2003).

We identified 13 different TBCSs (Bier et al., 2017; Childs et al., 2004; Cleland et al., 2007; Dewitte et al., 2014; Fernandez-de-las-Penas et al., 2011; Fritz and Brennan, 2007; Hanney et al., 2013; Hefford, 2008; Lee et al., 2017; Puentedura et al., 2012; Raney et al., 2009; Saavedra-Hernandez et al., 2011; Wang et al., 2003). Two TBCSs were very similar, but not identical (Childs et al., 2004; Fritz and Brennan, 2007). Fritz et al. used the proposed classification system from Childs et al. to develop an algorithm to prioritize the findings and place each

Table 2
The general part of the framework to describe treatment-based classification systems and the critical appraisal score.

Primary author	Purpose	Method of development	Domain of interest	Specific exclusion	Critical appraisal
Cleland (2007)	CPR to identify patients with neck pain who are likely to benefit from thoracic spine thrust manipulation.	Statistical method. Logistic regression modeling. Patients were dichotomized as success or non-success.	Mechanical neck pain. Subjects had to be between the ages of 18 and 60 years, with a primary complaint of neck pain with or without unilateral upper-extremity symptoms and a baseline Neck Disability Index (NDI) score of 10% or greater.	Evidence of any central nervous system involvement, or signs consistent with nerve root compression (at least 2 of the following had to be diminished to be considered nerve root involvement: myotomal strength, sensation, or reflexes).	4.5
Fernandez-de-las-Penas et al. (2011)	CPR to classify women with tension-type headache.	Statistical method. Logistic regression modeling. Patients were dichotomized as success or non-success.	Tension-type headache, diagnosed according to the criteria established by the International Headache Society.	No patient reported photophobia, phonophobia, vomiting or nausea during headache attacks, medication overuse headache. No apparent evidence of secondary headaches. No pain feature(s) of migraine or other headache. No history of cervical or cranial surgery. No evidence of any central nervous system involvement (e.g, loss of sensation, muscle atrophy, dysarthria).	3.5
Hanney (2013)	A preliminary CPR to determine which patients with neck pain may benefit from a standardized program of stretching and muscle performance exercise.	Statistical method. Logistic regression modeling. Patients were dichotomized as success or non-success.	non-specific neck pain and an NDI score of 10/50 or greater.	Evidence of central nervous system involvement, spasmodic torticollis, previously diagnosed migraines, previously diagnosed fibromyalgia, prior surgery to the neck or thoracic spine.	3.5
Puentedura (2012)	CPR to identify patients with neck pain likely to benefit from thrust joint manipulation to their cervical spine.	Statistical method. Logistic regression modeling. Patients were dichotomized as success or non-success.	Mechanical non-specific neck pain, with or without unilateral upper extremity symptoms, and have a baseline Neck Disability Index (NDI) score of 10 points (out of 50) or greater.	Any medical red flags suggesting that the etiology of symptoms might be nonmusculoskeletal; bilateral upper extremity symptoms; evidence of central nervous system involvement; pending legal action regarding the neck pain; 2 or more positive neurologic signs consistent with nerve root compression (changes in sensation, myotomal weakness, or decreased deep tendon reflexes); or any history of cervical spine surgery.	4
Raney (2009)	CPR to identify patients with neck pain who are likely to benefit from traction.	Statistical method. Logistic regression modeling. Patients were dichotomized as success or non-success.	Non-specific neck pain with or without upper extremity symptoms, and a baseline neck disability index (NDI) score of 20% or greater.	Identification of any medical red flags suggestive of a non-musculoskeletal etiology of symptoms, pregnancy, or any evidence of vascular compromise, central nervous system involvement or multiple-level neurological impairments.	3.5
Saavedra-Hernandez (2011)	CPR to classify patients with mechanical neck pain likely to experience improvements in both pain and disability after the application of an intervention including cervical and thoracic spine thrust manipulations.	Statistical method. Logistic regression modeling. Patients were dichotomized as success or non-success.	Mechanical neck pain with or without upper-extremity symptoms.	Any contraindication to spinal manipulation: positive extension-rotation test or nystagmus; no history of cervical surgery; diagnosis of fibromyalgia; previous treatment with spinal manipulative therapy; or evidence of any central nervous system involvement, or signs consistent with nerve root compression.	3.5
Bier et al. (2017)	A classification system to classify patients with non-specific neck pain into risk groups.	Judgement based on a small group of experts + literature review.	Non-specific neck pain		5
Childs (2004)	A classification system for patients with non-specific neck pain.	Judgement based on a small group of experts + literature review.	Non-specific neck pain.		2.5
Dewitte (2014)	A clinical algorithm to guide specific mobilization and manipulation.	Judgement based on a small group of experts + literature review.	Mechanical neck pain.	No neurological findings in clinical history or manual assessment; no signs of central hyper excitability.	2.5
Fritz (2007)	A classification system based on clinical characteristics for the purpose of specifically directing nonsurgical treatment choices.	Judgement based on a small group of experts + literature review.	Non-specific neck pain.		3.5
Hefford (2008)	A classification system for patients with non-specific neck pain.	Judgement based on a small group of experts + literature review.	Non-specific mechanical neck pain.		4.5
Lee (2017)	A self-classification system for a smartphone-based exercise program feasible for office	Judgement based on a small group of experts + literature review.	Office workers with non-specific neck pain.	No other treatment or surgery within 3 months; or their neck pain was caused by a known trauma.	2.5

(*continued on next page*)

Table 2 (*continued*)

Primary author	Purpose	Method of development	Domain of interest	Specific exclusion	Critical appraisal
	workers as a method of self-managing their neck pain.				
Wang (2003)	A clinical decision-making algorithm to classify patients with cervical pain likely to response to an individualized physical therapy intervention.	Judgement based on one expert.	Neck pain with or without radiating pain.	No long-term use of systemic steroids over 3 months, no spinal surgery within the previous year	2.5

patient into a classification category. This algorithm is slightly different from that of Childs et al. (2004) due to differences in criteria and interventions (Table 3). Therefore, we included both as separate TBCSs and considered both studies as development studies.

3.2. Description of TBCSs

The characteristics of the TBCSs are presented in Tables 2 and 3 Table 2 describes the purpose of the TBCS, the method of development, the domain of interest and the specific exclusions, so when not to use the TBCS. For example, one TBCS aimed to develop a classification system to classify patients with non-specific neck pain into prognostic risk groups (Bier et al., 2017). The method of development was judgement-based in which only a small group of experts was involved. It also included a literature review and the domain of interest was patients with non-specific neck pain. They described no specific exclusion criteria which means that this TBCS can be applied to every patient with non-specific neck pain.

Table 3 presents the TBCSs and the criteria they use to subgroup patients and the treatments that are matched to each subgroup. For example, the above mentioned TBCS had three categories (low, moderate and high risk for persisting disability) with their own criterion (i.e. the score on the StartBackTool) with treatments for each criterion.

Six (out of 13) TBCSs followed a statistical approach (Cleland et al., 2007; Fernandez-de-las-Penas et al., 2011; Hanney et al., 2013; Puentedura et al., 2012; Raney et al., 2009; Saavedra-Hernandez et al., 2011) and are all referred to as Clinical Prediction Rules (CPRs) (Beattie and Nelson, 2006; Randolph et al., 1998) Seven TBCSs used a judgment-based approach (Bier et al., 2017; Childs et al., 2004; Dewitte et al., 2014; Fritz and Brennan, 2007; Hefford, 2008; Lee et al., 2017; Wang et al., 2003).

3.3. Quality of TBCSs

The percentage agreement between the raters was 100% on purpose, face validity, construct validity and reliability, 92% on content validity and generalizability and 83% on feasibility and the total score (see Table 2, Fig. 2 and Appendix 2).

We included five reliability studies (Bier et al., 2017; Clare et al., 2004; Clare et al., 2005; Cleland et al., 2006; Fritz and Brennan, 2007) for four TBCSs: STarT Back tool, McKenzie system, Cleland classification system and Fritz (Bier et al., 2017; Cleland et al., 2007; Fritz and Brennan, 2007; Hefford, 2008). The reliability scores varied between 0.56 and 0.95. Three TBCSs scored half a point on the Buchbinder appraisal scale for the reliability criterion, and only the Fritz system had a score of one point on the Buchbinder scale.

Four TBCSs had the lowest overall quality score of 2.5 point (out of 7) (Childs et al., 2004; Dewitte et al., 2014; Lee et al., 2017; Wang et al., 2003), while one TBCS gained the highest score of 5 (out of 7) points (Bier et al., 2017). We found for all TBCSs that the criterion 'construct validity' scored zero and the criterion 'purpose' scored one. Four TBCSs were also used in other settings than in the studies describing the development of the TBCSs (Bier et al., 2018; Cleland et al., 2010; Farrell and Lampe, 2018; Kjellman and Oberg, 2002) supporting the generalizability of these TBCSs. Fig. 2 shows the summary of the quality of the 13 classification systems. The overall quality of the TBCSs ranged from low to moderate.

3.4. Effectiveness of TBCSs

Two RCTs investigated the effectiveness of two TBCSs: the Cleland classification system and the McKenzie system (Cleland et al., 2010; Kjellman and Oberg, 2002).

The Cleland study investigated the effect of four intervention groups: these were manipulation plus exercise, with one group positive and one group negative on the CPR; and exercise only, with one group positive and one group negative on the CPR (= rule status). The authors found no statistical significant mean differences, nor clinically relevant differences (Hjermstad et al., 2011; Schellingerhout et al., 2012; Williamson and Hoggart, 2005) on function (the Neck disability index (NDI)) for + CPR vs – CPR) of −0.68 (95% CI -3.1 to 1,7) and of 0.9 (95% CI -0.3 to 0.49) or pain (Numeric Pain Rating Scale (NPRS))This finding does not support the use of the CPR (Cleland et al., 2010).

Concerning the McKenzie system; one study compared the effectiveness of three groups: McKenzie treatment, general exercise and a control group (Kjellman and Oberg, 2002). The control group received ultrasound administered at the lowest intensity possible and with the indicator lights on. They found no statistically significant (nor clinically relevant) between-group differences. Results after six months were: Pain (Visual Analog Scale (VAS)): McKenzie 21 (SD 17), general exercise 23 (SD 26) and control group 27 (SD 23); function (NDI): McKenzie 15 (SD 12), general exercise 17 (SD 17) and control group 18 (SD 15). Both studies had a low risk of bias score on the PEDro scale (https://www.pedro.org.au).

4. Discussion

4.1. Main results

This systematic review identified a total of 13 TBCSs. The overall quality of the TBCSs ranged from low to moderate. We found two randomized clinical trials, with low risk of bias, evaluating the effectiveness of two TBCSs, showing that they were not superior to alternative treatments.

4.2. Discussion of findings

No statistically-derived TBCS scored the maximum of one point for the face validity criterion because there is no clear relation (in the clinical sense) between the items of the TBCSs and their presumed matching interventions. For statistically-derived TBCSs to make predictions about an individual response to a treatment it does not matter how the variables relate to the intervention, as long as they are predictive of the outcome. Therefore, face validity seems to plays no direct role in statistically-derived TBCSs. However, in the methodological standards for derivation of a statistically-derived TBCS, it was stated that such a TBCS has to make "clinical sense" (McGinn et al., 2000).

Judgement based TBCSs also had poor face validity. Exemplary for

Table 3
The specific part of the framework to describe classification systems.

Primary author	Categories	Criteria used	Treatment
Cleland (2010)	1. CPR for thoracic manipulation	• Symptoms <30 days • No symptoms distal to the shoulder • Looking up does not aggravate symptoms • FABQPA score <12 • Diminished upper thoracic spine kyphosis • Cervical extension ROM <30°	CROM exercises +3 different thrust manipulation techniques directed at the thoracic spine: a seated "distraction" manipulation, a supine upper thoracic spine manipulation, and a middle thoracic spine manipulation.
Fernandez-de las Peñas (2011)	1. CPR for tension type headache	• Mean age <44.5 years • Presence left sternocleidomastoid muscle TrP • Presence suboccipital muscle TrP • Presence of left superior oblique muscle TrP • Cervical rotation to the left >69° • Total tenderness score <20.5 • NDI <18.5 • Referred pain area of right upper trapezius muscle TrP >42.23	Multimodal physical therapy including joint mobilization and muscle trigger point therapies.
Hanney (2013)	1. CPR for a standardized program of stretching and muscle performance exercise	• FABQ-Physical Activity score <15 • NDI <18/50 • Does not participate in cycling (for regular exercise) • Shoulder protraction • AROM side bending to one side < 32°	1. Stretches were performed: upper trapezius, anterior and middle scalenes, suboccipital, and pectoralis major. Each stretch was held for 30s and repeated two times. 2. Bilateraly muscle performance exercise progressions were instructed: isometric cervical extension, shoulder protraction, craniocervical flexion, seated row, horizontal shoulder abduction with external rotation, and shoulder elevation in the scapular plane. 3. All patients began with thin elastic bands and progressed to medium, heavy and extra heavy for resistance, as appropriate based on the patient's ability.
Puentedura (2012)	1. CPR for thrust joint manipulations in the cervical spine	• symptom duration of less than 38 days • positive expectation that manipulation will help • side-to-side difference in cervical rotation ROM of 10° or greater • pain with postero-anterior spring testing of the middle cervical spine	1. Supine TJM to the cervical spine directed to an appropriate level between C3 and C7. 2. Gentle active ROM exercise (10 repetitions performed 3–4 times daily) and advised to maintain usual activity within the limits of pain.
Raney (2009)	1. CPR for cervical traction	• Age >55 • Positive shoulder abduction test • Positive ULTT A • Symptom peripheralization with central posterior–anterior motion testing at lower cervical (C4–7) spine • Positive neck distraction test	1. Intermittent mechanical traction was performed using one of two traction units: the Chattanooga Triton Traction Table and the Saunders 3D Active Trac Table. 2. An active exercise intervention.
Saavedra-Hernandez (2011)	1. CPR for mechanical neck pain	• Sex: Female • Pain greater than 4.5 • Extension range of motion less than 46° • Hypomobility T1 • ULTT negative	3 thrust manipulation techniques targeted at the mid cervical spine, cervicothoracic junction, and upper thoracic spine region.
Bier (2017)	1. Low risk for persisting disability 2. Moderate risk for persisting disability 3. High risk for persisting disability	LR 1. Total StartBack score ≤3 points MR 1. Total StartBack score ≥4; Score question 5-9 ≤3 points HR 1. Total StartBack score ≥4; Score question 5-9 ≥4 points	LR 1. The GP provided information, advice, and some analgesics or 1 or 2 physiotherapist consultations, and the treatment was hands-off and consisted of offering information, advice, and exercises. MR 1. In addition to the low-risk approach, the GP referred the patient to a physiotherapist, and the physiotherapist performed an evidence-based intervention. HR 1. In addition to the medium-risk approach, the GP referred the patient to either a physiotherapist specialized in treating patients with a psychosomatic approach, a psychologist, or equivalent, and the physiotherapist assessed biopsychosocial risk factors and used cognitive behavioral principles as interventions.
Childs et al. (2004)	1. Mobility 2. Centralization 3. Conditioning and increase exercise tolerance 4. Pain control 5. Reduce headache	MB 1. Recent onset of symptoms 2. No radicular/referred symptoms in the upper quarter 3. Restricted range of motion with side-to-side rotation and/or discrepancy in lateral flexion range of motion 4. No signs of nerve root compression or peripheralization of symptoms in the upper quarter with cervical range of motion CZ 1. Radicular/referred symptoms in the upper quarter 2. Peripheralization and/or centralization of symptoms with range of motion 3. Signs of nerve root compression present 4. May have pathoanatomic diagnosis of cervical radiculopathy CD 1. Lower pain and disability scores 2. Longer duration of symptoms 3. No signs of nerve root compression	MB 1. Cervical and thoracic spine mobilization/ manipulation. 2. Active range of motion exercises. CZ 1. Mechanical/manual cervical traction. 2. Repeated movements to centralize symptoms. CD 1. Strengthening and endurance exercises for the muscles of the neck and upper quarter. 2. Aerobic conditioning exercises. PC 1. Gentle active range of motion within pain tolerance. 2. Range of motion exercises for adjacent regions. 3. Physical modalities as needed. 4. Activity modification to control pain. RH 1. Cervical spine manipulation /mobilization. 2. Strengthening of neck and upper quarter muscles. 3. Postural education.

(*continued on next page*)

Table 3 (*continued*)

Primary author	Categories	Criteria used	Treatment
		4. No peripheralization/centralization during range of motion PC 1. High pain and disability scores 2. Very recent onset of symptoms 3. Symptoms precipitated by trauma 4. Referred or radiating symptoms extending into the upper quarter 5. Poor tolerance for examination or most interventions RH 1. Unilateral headache with onset preceded by Neck pain 2. Headache pain triggered by neck movement or positions 3. Headache pain elicited by pressure on posterior neck	
Dewitte (2014)	1. Cervical spine convergence pattern 2. Cervical spine divergence pattern	CCP 1. Subjective examination: Feeling of locking, movement restriction, unilateral compression pain, often in acute cases, antalgic posture 2. Physical examination: Active and passive combined extension, ipsilateral side bending and rotation is limited and evokes comparable signs 3. Articular examination: Provocation tests are positive at impaired segment, downslope restriction ipsilateral, segmental distraction alleviates the pain CDP 1. Subjective examination: Feeling of painful strain at end of ROM, movement restriction at end of ROM, unilateral stretch pain, high intensity or severity of symptoms is rare, antalgic posture is uncommon 2. Physical examination: Active and passive Combined flexion, contralateral side bending and rotation is limited and evokes comparable signs, passive shoulder elevation in this position does not result in increased ROM/decreased pain 3. Articular examination: Provocation tests are positive at the impaired segment, ipsilateral upslope restriction	CCP 1. Distraction technique; gapping Technique. 2.Translatoric technique - indirect upslope technique or direct downslope technique. CDP 1. Distraction technique. 2. Translatoric upslope technique - focus or locking approach.
Fritz (2007)	1. Mobility 2. Centralization 3. Exercise and conditioning 4. Pain control 5. Headache	MOB 1. Mode of onset no whiplash mechanism 2. NPRS <7 or NDI score <52/100 3. No signs of nerve root compression 4. No symptoms below the elbow 5. The chief complaint is not headache with neck pain 6. Duration of symptoms <30 days + patient age <60 years CEN 1. Mode of onset no whiplash mechanism 2. Duration of symptoms >30 days 3. NPRS <7 or NDI score <52/100 4. Any signs of nerve root compression 5. No signs of nerve root compression + symptoms below the elbow EaC 1. Mode of onset no whiplash mechanism 2. NPRS <7 or NDI score <52/100 3. No signs of nerve root compression 4. No symptoms below the elbow 5. The chief complaint is not headache with neck pain 6. Duration of symptoms >30 days + Patient age >60 years PC 1. Mode of onset no whiplash mechanism 2. Duration of symptoms <30 days 3. NPRS >7 or NDI score >52/100 HA 1. Mode of onset no whiplash mechanism 2. Duration of symptoms >30 days 3. NPRS <7 or NDI score <52/100 4. No signs of nerve root compression 5. No symptoms below the elbow 6. The chief complaint is headache with neck pain 7. Headache is affected by neck movement 8. There is a diagnosis or symptoms of migraines	MOB 1. Cervical or thoracic mobilization or manipulation. 2. Strengthening exercises for the deep neck flexor muscles. CEN 1. Mechanical or manual cervical traction (at least 50% of the sessions). 2. Cervical retraction exercises (at least 50% of the sessions). EAC 1. Strengthening exercises for the upper quarter muscles. 2. Strengthening exercises for the neck or deep neck flexor muscles. PC 1. Cervical spine mobilization. 2. Cervical range-of-motion exercises. HA 1. Cervical spine manipulation or mobilization. 2. Strengthening exercises for the deep neck flexor muscles. 3. Strengthening exercises for the upper quarter muscles.
Hefford (2008)	1. Posture syndrome 2. Dysfunction syndrome 3. Derangement syndrome 4. Other	PS 1. Pain arising as a result of mechanical deformation of normal soft tissues from prolonged end range loading of periarticular structures DyS 1. Pain occurring as a result of mechanical deformation of structurally impaired tissues (such as tissue which is scarred, adhered or adaptively shortened). DeS 1. Pain occurring as a result of a disturbance in the normal resting position of the affected joint surfaces. Derangement may be reducible or irreducible. Ot 1. Those who do not fit the mechanical syndromes but who exhibit signs and symptoms of other known pathology	PS 1. Posture correction. DyS 1. Exercise into the direction of the dysfunction with the aim of remodeling the tissue. Des 1. Depends on the clinically induced directional preference, identified by examining the patient's symptomatic and mechanical response to repeated movements or sustained positions. A reducible derangement typically demonstrates one direction of repeated movement (directional preference) which decreases or centralizes (moves towards the midline) referred symptoms, or abolishes midline symptoms, and the opposite repeated movement which produces or increases or peripheralizes (moves more distally) the symptoms.
Lee (2017)	1. Exercise and conditioning 2. Mobility	EaC 1. Pain on the side where the patient's neck was rotated during the Neck Rotation and Extension Test. 2. No restricted ROM 3. No pain, numbness, or weakness in the shoulder or arm of	EaC 1. Strengthening exercise for deep neck muscles and upper-quarter muscles. MoB 1. Stretching exercise. 2. Strengthening exercise for deep neck muscles.

(*continued on next page*)

Table 3 (*continued*)

Primary author	Categories	Criteria used	Treatment
	3. Centralization 4. Reduce headache	the side where the neck was rotated during the Neck Rotation and Extension Test 4. No peripheralization/centralization of symptoms with repeated movements during the Repeated Movement Test MoB 1. Pain on the side opposite where the patient's neck was rotated during the Neck Rotation and Extension Test 2. Restricted ROM during the Repeated Movement Test 3. No pain, numbness, or weakness in the shoulder or arm of the side where the neck was rotated during the Neck Rotation and Extension Test 4. No peripheralization/centralization of symptoms with repeated movements during the Repeated Movement Test CeN 1. Pain, numbness, and/or weakness in the shoulder and/or arm of the side where the neck was rotated during the Neck Rotation and Extension Test 2. Centralization or distal symptom reduction with the Repeated Movement Test RH 1. Headache triggered by neck movement or position 2. Headache elicited by pressure on the ipsilateral posterior neck	CeN 1. Cervical retraction exercise and repeated neck extension exercise. 2. Strengthening exercise for deep neck muscles. RH 1. Self-myofascial release technique. 2. Strengthening exercise for deep neck and upper-quarter muscles.
Wang (2003)	1: radicular arm pain or neck pain; 3 patterns 2: referred arm pain or neck pain; 6 patterns 3: cervicogenic headaches: 4 patterns 4: neck pain only; 5 patterns	RAD-1 1. Positive neurologic signs 2. Centralization or distal symptom reduction occurred with repeated movements RAD-2 1. Positive neurologic signs 2. Centralization or distal symptom reduction did not occur with repeated movements 3. Manual traction decreased symptoms RAD-3 1. Positive neurologic signs 2. Centralization or distal symptom reduction did not occur with repeated movements 3. Manual traction did not decrease symptoms REF-1 1. Negative neurologic signs 2. Referred arm pain with or without neck pain 3. Centralization or distal symptom reduction occurred with repeated movements REF-2 1. Negative neurologic signs 2. Referred arm pain with or without neck pain 3. Centralization or distal symptom reduction did not occur with repeated movements 4. Positive upper limb tension tests (ULTTs) REF-3 1. Negative neurologic signs 2. Referred arm pain with or without neck pain 3. Centralization or distal symptom reduction did not occur with repeated movements 4. Positive ULTTs 5. Manual traction did not decrease symptoms 6. Negative thoracic outlet syndrome tests REF-4 1. Negative neurologic signs 2. Referred arm pain with or without neck pain 3. Centralization or distal symptom reduction did not occur with repeated movements 4. Positive ULTTs 5. Manual traction did not decrease symptoms 6. Positive thoracic outlet syndrome tests 7. Positive shoulder depression provocation /release tests REF-5 1. Negative neurologic signs 2. Referred arm pain with or without neck pain 3. Centralization or distal symptom reduction did not occur with repeated movements 4. Positive ULTTs 5. Manual traction did not decrease symptoms 6. Positive thoracic outlet syndrome tests 7. Negative shoulder depression provocation /release tests REF-6 1. Negative neurologic signs 2. Referred arm pain with or without neck pain 3. Centralization or distal symptom reduction did not occur with repeated movements 4. Negative ULTTs 5. Manual traction decreased symptoms HA-1 1. Negative neurologic signs 2. Referred pain into the temporal/facial area HA-2 1. Negative neurologic signs 2. Referred pain in non-facial areas 3. Occipito-atlantal (OA) joint distraction provoked or reduced symptoms	RAD-1 1. Mechanical traction. 2. Repeated movement exercises that can centralize the symptoms. 3. May need manual therapy techniques to enhance the repeated movement exercises. 4. Postural exercise. 5. Education. RAD-2 1. Mechanical positional cervical Traction. 2. Specific level manual traction (with foraminal opening as needed). 3. Postural exercise. 4. Education. 5. Continue monitoring treatment response, may evolve to pattern 1 if beginning centralization with repeated movement. RAD-3 1. Trial of strong prolonged mechanical traction. 2. Neural mobilization to distract and release tension on neural tissues. 3. Continue monitoring treatment response, may evolve to pattern 2 or 1; if no progress, refer back to physician. REF-1 1. Mechanical traction. 2. Repeated movement exercises that can centralize the symptoms. 3. May need manual therapy techniques To enhance the repeated movement exercises. 4. Postural exercise. 5. Education. REF-2 1. Mechanical positional cervical traction 2. Specific level manual traction 3. Postural exercise 4. Education REF-3 1. Neural mobilization to desensitize. 2. Postural exercise. 3. Activity tolerance training. 4. If no progress with treatment, refer back to physician. REF-4 1. Thoracic outlet release techniques (tissue- specific). 2. Specific mobilization and stretching. 3. Postural exercise. 4. Activity tolerance training. 5. Education. REF-5 1. Trial treatment using neural mobilization, strong mechanical traction, postural exercise, and activity tolerance. 2. If no progress with treatment, refer back to physician. REF-6 1. Specific joint mobilization. 2. Mechanical traction. HA-1 1. Temporomandibular joint treatment protocol. HA-2 1. Suboccipital muscle stretching. 2. Specific joint mobilization or muscle energy techniques to OA, AA, and C2–3. 3. Postural exercise. HA-3 1. Specific joint mobilization or muscle energy techniques to AA joint. 2. Postural exercise. HA-4 1. Specific joint mobilization or muscle energy

(*continued on next page*)

Table 3 (*continued*)

Primary author	Categories	Criteria used	Treatment
		HA-3 1. Negative neurologic signs 2. Referred pain in non-facial areas 3. OA joint distraction did not provoke or reduce symptoms 4. Atlantoaxial (AA) joint distraction provoked or educed symptoms HA-4 1. Negative neurologic signs 2. Referred pain in non-facial areas 3. Either OA or AA joint distraction did not provoke or reduce symptoms 4. Joint distraction on other cervical spinal Level provoked or reduced symptoms NP-1 1. Negative neurologic signs 2. Neck pain only 3. Gross movement tests showed capsular pattern restriction NP-2 1. Negative neurologic signs 2. Neck pain only 3. Gross movement tests showed non-capsular pattern restriction 4. Pain on the same side of side bending /rotation 5. Segmental mobility test showed hypomobile segment(s) of the involved level NP-3 1. Negative neurologic signs 2. Neck pain only 3. Gross movement tests showed non-capsular pattern restriction 4. Pain on the same side of side bending /rotation 5. Segmental mobility test showed hypermobile segment(s) of the involved level NP-4 1. Negative neurologic signs 2. Neck pain only 3. Gross movement tests showed non-capsular pattern restriction 4. Pain on the opposite side of side bending /rotation 5. Segmental mobility test showed hypomobile segment(s) of the involved level NP-5 1. Negative neurologic signs 2. Neck pain only 3. Gross movement tests showed non-capsular pattern restriction 4. Pain on the opposite side of side bending /rotation 5. Segmental mobility test showed hypermobile segment(s) of the involved level	techniques to the involved level(s). 2. Mechanical traction in the absence of specific joint mobilization technique. NP-1 1. Mechanical traction. 2. Manual traction. 3. Neck range of motion exercise. 4. Postural exercise. NP-2 1. Specific joint mobilization (muscle energy or dorsal gliding techniques) to the involved level(s). 2. Postural exercise. NP-3 1. Specific joint mobilization (muscle energy or gliding techniques) to the adjacent level(s). 2. Stabilization exercise. NP-4 1. Specific joint mobilization (muscle energy or ventral gliding techniques) to the involved level(s). 2. Postural exercise. NP-5 1. Specific joint mobilization (muscle energy or gliding techniques) to the adjacent level(s). 2. Stabilization exercise.

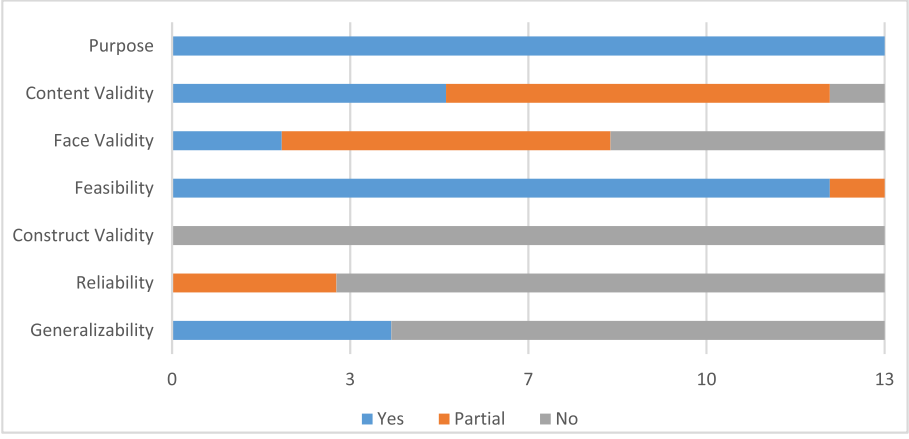

Fig. 2. Quality summary of the 13 classification systems, based on the appraisal tool.

this was that many criteria of the framework did not match the interventions. For example, the category 'Exercise and conditioning' with the intervention 'Strengthening exercises' for deep neck muscles and upper-quarter muscles. In this category, is at least one diagnostic criterion that relates to reduced muscle strength seems to be missing. If muscle strength is not reduced, why apply 'Strengthening exercises'. Or in other words, how could muscle strength be effective if muscle strength is not reduced in the first place (Jull et al., 2009). Apparently, it is difficult to link diagnostic criteria to clinically-relevant matching interventions. A further explanation for moderate-to-low face validity may be the lack of convincing evidence for which subgroups should be matched to which treatments. In a recently-published systematic review,

RCTs typically lacked a clear and recognizable clinical reasoning process (Maissan et al., 2018).

We were not able to appraise the construct validity as none of the included studies compared their TBCS (or parts of this TBCS) to other relevant classification systems. Although this may be challenging, we still think it is important to establish the construct validity of a TBCS.

4.3. Comparison with other literature

Our study is broadly in line with a recent published review. This review only included statistically-derived TBCSs but came to the same conclusions and also recommended not to use statistically-derived TBCSs in daily practice (Kelly et al., 2017). Another review that critically appraised statistically-derived TBCSs focused on musculoskeletal conditions (Stanton et al., 2010) concluded that "at present, there is little evidence that statistically-derived TBCSs can be used to predict effects of treatment for musculoskeletal conditions".

4.4. Strengths and weaknesses

As far as we know, this review is the first review focussing specifically on TBCSs in patients with non-specific neck pain, but its results should be interpreted in the light of some limitations. First, the validity of the Buchbinder scale has not been established. In addition, as the quality criteria of the Buchbinder scale could not always be clearly operationalized, this may have affected scores. To overcome this limitation, we defined, a-priori, agreements how to score (based on the pilot test). A strength of this study was the use of sensitive search strategies in multiple databases, developed in collaboration with a medical information specialist, and also the searching of grey literature to avoid missing relevant studies (Rethlefsen et al., 2015).

4.5. Implications

One important feature of a TBCS is the clinical relevance (McGinn et al., 2000). For most of the included TBCSs, the clinical relevance was not always clear. Therefore, if we continue to develop TBCSs, attention should be paid to the clinical relevance within the design.

Only two of the 13 TBCSs were evaluated on the impact on clinical outcomes. As only TBCSs that have an impact in daily practice should be recommended, we recommend to evaluate the impact of existing TBCS instead of developing new ones (Kappen et al., 2018; van Giessen et al., 2017).

Due to the low to moderate quality and the lack of effectiveness of the existing TBCSs we do not recommend their use in daily clinical practice.

In conclusion, we identified 13 TBCSs with overall a low to moderate quality. In addition, the effectiveness of the majority of these TBCSs was not evaluated. Two TBCSs were evaluated on effectiveness and found to be equally effective compared to other approaches. Furthermore, the clinical relevance of the included TBCSs was not always clear. Therefore, we conclude that these TBCSs should not be used in clinical practice.

Ethical approval

Not applicable.

Funding

None.

This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (16) and registered in the international prospective register of systematic reviews PROSPERO (CRD 42018087763).

Declaration of competing interest

None.

Acknowledgements

We want to thank Arianne Verhagen for her assistance and comments that have greatly improved the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.msksp.2020.102133.

References

Beattie, P., Nelson, R., 2006. Clinical prediction rules: what are they and what do they tell us? Aust. J. Physiother. 52 (3), 157–163.
Bier, J.D., Ostelo, R.W.J.G., Koes, B.W., Verhagen, A.P., 2017. Validity and reproducibility of the modified STarT back tool (Dutch version) for patients with neck pain in primary care. Muscoskel. Sci. Pract. 31, 22–29.
Bier, J.D., Sandee-Geurts, J.J.W., Ostelo, R.W.J.G., Koes, B.W., Verhagen, A.P., 2018. Can primary care for back and/or neck pain in The Netherlands benefit from stratification for risk groups according to the STarT back tool classification? Arch. Phys. Med. Rehabil. 99 (1), 65–71.
Buchbinder, R., Goel, V., Bombardier, C., Hogg-Johnson, S., 1996. Classification systems of soft tissue disorders of the neck and upper limb: do they satisfy methodological guidelines? J. Clin. Epidemiol. 49 (2), 141–149.
Childs, J.D., Fritz, J.M., Piva, S.R., Whitman, J.M., 2004. Proposal of a classification system for patients with neck pain. J. Orthop. Sports Phys. Ther. 34 (11), 686–696.
Clare, H.A., Adams, R., Maher, C.G., 2004. Reliability of the McKenzie spinal pain classification using patient assessment forms. Physiotherapy 90, 114–119.
Clare, H.A., Adams, R., Maher, C.G., 2005. Reliability of McKenzie classification of patients with cervical or lumbar pain. J. Manip. Physiol. Ther. 28 (2), 122–127.
Cleland, J.A., Childs, J.D., Fritz, J.M., Whitman, J.M., 2006. Interrater reliability of the history and physical examination in patients with mechanical neck pain. Arch. Phys. Med. Rehabil. 87 (10), 1388–1395.
Cleland, J.A., Childs, J.D., Fritz, J.M., Whitman, J.M., Eberhart, S.L., 2007. Development of a clinical prediction rule for guiding treatment of a subgroup of patients with neck pain: use of thoracic spine manipulation, exercise, and patient education. Phys. Ther. 87 (1), 9–23.
Cleland, J.A., Mintken, P.E., Carpenter, K., Fritz, J.M., Glynn, P., Whitman, J., Childs, J. D., 2010. Examination of a clinical prediction rule to identify patients with neck pain likely to benefit from thoracic spine thrust manipulation and a general cervical range of motion exercise: multi-center randomized clinical trial. Phys. Ther. 90 (9), 1239–1250.
Coupe, V.M.H., van Hooff, M.L., de Kleuver, M., Steyerberg, E.W., Ostelo, R.W.J.G., 2016. Decision support tools in low back pain. Best Pract. Res. Clin. Rheumatol. 30 (6), 1084–1097.
Damgaard, P., Bartels, E.M., Ris, I., Christensen, R., Juul-Kristensen, B., 2013. Evidence of physiotherapy interventions for patients with chronic neck pain: a systematic review of randomised controlled trials. ISRN Pain 15, 567175.
de Morton, N.A., 2009. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. Aust. J. Physiother. 55 (2), 129–133.
Dewitte, V., Beernaert, A., Vanthillo, B., Barbe, T., Danneels, L., Cagnie, B., 2014. Articular dysfunction patterns in patients with mechanical neck pain: a clinical algorithm to guide specific mobilization and manipulation techniques. Man. Ther. 19 (1), 2–9.
Ezzo, J., Haraldsson, B.G., Gross, A.R., Myers, C.D., Morien, A., Goldsmith, C.H., Peloso, P.M., 2007. Massage for mechanical neck disorders: a systematic review. Spine 32 (3), 353–362.
Fairbank, J., Gwilym, S.E., France, J.C., Daffner, S.D., Dettori, J., Hermsmeyer, J., Andersson, G., 2011. The role of classification of chronic low back pain. Spine 36 (21 Suppl. l), S19–S42.
Farrell, K.P., Lampe, K.E., 2018. Patient outcomes with and without implementation of a neck pain classification system: a preliminary analysis. Orthop. Phys. Ther. Prac. 30 (2), 82–90.
Fernandez-de-las-Penas, C., Cleland, J.A., Palomeque-del-Cerro, L., Caminero, A.B., Guillem-Mesado, A., Jimenez-Garcia, R., 2011. Development of a clinical prediction rule for identifying women with tension-type headache who are likely to achieve short-term success with joint mobilization and muscle trigger point therapy. Headache 51 (2), 246–261.
Foster, N.E., Hill, J.C., Hay, E.M., 2011. Subgrouping patients with low back pain in primary care: are we getting any better at it? Man. Ther. 16 (1), 3–8.
Fritz, J.M., Brennan, G.P., 2007. Preliminary examination of a proposed treatment-based classification system for patients receiving physical therapy interventions for neck pain. Phys. Ther. 87 (5), 513–524.
Graham, N., Gross, A., Goldsmith, C.H., Klaber Moffett, J., Haines, T., Burnie, S.J., Peloso, P.M., 2008. Mechanical traction for neck pain with or without radiculopathy. Cochrane Database Syst. Rev. 3, CD006408.

Gross, A., Forget, M., St George, K., Fraser, M.M., Graham, N., Perry, L., Brunarski, D., 2012. Patient education for neck pain. Cochrane Database Syst. Rev. 3, CD005106.
Gross, A.R., Hoving, J.L., Haines, T.A., Goldsmith, C.H., Kay, T., Aker, P., Cervical Overview Group, 2004. A cochrane review of manipulation and mobilization for mechanical neck disorders. Spine 29 (14), 1541–1548.
Hanney, W.J., Kolber, M.J., George, S.Z., Young, I., Patel, C.K., Cleland, J.A., 2013. Development of a preliminary clinical prediction rule to identify patients with neck pain that may benefit from a standardized program of stretching and muscle performance exercise: a prospective cohort study. Int. J. Sports Phys. Ther. 8 (6), 756–776.
Hefford, C., 2008. McKenzie classification of mechanical spinal pain: profile of syndromes and directions of preference. Man. Ther. 13 (1), 75–81.
Hjermstad, M.J., Fayers, P.M., Haugen, D.F., Caraceni, A., Hanks, G.W., Loge, J.H., European Palliative Care Research Collaborative (Epcrc), 2011. Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. J. Pain Symptom Manag. 41 (6), 1073–1093.
Hogg-Johnson, S., van der Velde, G., Carroll, L.J., Holm, L.W., Cassidy, J.D., Guzman, J., Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders, 2008. 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. Spine 33 (4 Suppl. l), S39–S51.
Hurwitz, E.L., Randhawa, K., Yu, H., Cote, P., Haldeman, S., 2018. The global spine care initiative: a summary of the global burden of low back and neck pain studies. Eur. Spine J. 27 (Suppl. 6), 796–801.
Jull, G.A., Falla, D., Vicenzino, B., Hodges, P.W., 2009. The effect of therapeutic exercise on activation of the deep cervical flexor muscles in people with chronic neck pain. Man. Ther. 14 (6), 696–701.
Kappen, T.H., van Klei, W.A., van Wolfswinkel, L., Kalkman, C.J., Vergouwe, Y., Moons, K.G.M., 2018. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn. Prog. Res. 2, 11-018-0033-6.
Kay, T.M., Gross, A., Goldsmith, C.H., Rutherford, S., Voth, S., Hoving, J.L., Santaguida, P.L., 2012. Exercises for mechanical neck disorders. Cochrane Database Syst. Rev. 8, CD004250.
Kelly, J., Ritchie, C., Sterling, M., 2017. Clinical prediction rules for prognosis and treatment prescription in neck pain: a systematic review. Muscoskel. Sci. Pract. 27, 155–164.
Kjellman, G., Oberg, B., 2002. A randomized clinical trial comparing general exercise, McKenzie treatment and a control group in patients with neck pain. J. Rehabil. Med. 34 (4), 183–190.
Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33 (1), 159–174.
Lee, M., Lee, S.H., Kim, T., Yoo, H.J., Kim, S.H., Suh, D.W., Yoon, B., 2017. Feasibility of a smartphone-based exercise program for office workers with neck pain: an individualized approach using a self-classification algorithm. Arch. Phys. Med. Rehabil. 98 (1), 80–87.
Linton, S.J., Nicholas, M., Shaw, W., 2018. Why wait to address high-risk cases of acute low back pain? A comparison of stepped, stratified, and matched care. Pain 159 (12), 2437–2441.
Maher, C.G., Sherrington, C., Herbert, R.D., Moseley, A.M., Elkins, M., 2003. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys. Ther. 83 (8), 713–721.
Maissan, F., Pool, J.J.M., Raaij de, E., Mollema, J., Ostelo, R.W.J.G., Wittink, H., 2018. The clinical reasoning process in randomised clinical trials with patients with non-specific neck pain is incomplete: a systematic review. Muscoskel. Sci. Pract. 35, 8–17.
March, L., Smith, E.U., Hoy, D.G., Cross, M.J., Sanchez-Riera, L., Blyth, F., Woolf, A.D., 2014. Burden of disability due to musculoskeletal (MSK) disorders. Best Pract. Res. Clin. Rheumatol. 28 (3), 353–366.
McGinn, T.G., Guyatt, G.H., Wyer, P.C., Naylor, C.D., Stiell, I.G., Richardson, W.S., 2000. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. evidence-based medicine working group. Jama 284 (1), 79–84.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Prisma Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Open Med. 3 (3), e123–e130.
Monticone, M., Ambrosini, E., Cedraschi, C., Rocca, B., Fiorentini, R., Restelli, M., Moja, L., 2015. Cognitive-behavioral treatment for subacute and chronic neck pain: a cochrane review. Spine 40 (19), 1495–1504.
Portney, L.G., Watkins, M.P., 2000. Foundations of Clinical Research, second ed. ed. Prentice Hall Health, New Jersey.
Puentedura, E.J., Cleland, J.A., Landers, M.R., Mintken, P.E., Louw, A., Fernandez-de-Las-Penas, C., 2012. Development of a clinical prediction rule to identify patients with neck pain likely to benefit from thrust joint manipulation to the cervical spine. J. Orthop. Sports Phys. Ther. 42 (7), 577–592.
Randolph, A.G., Guyatt, G.H., Calvin, J.E., Doig, G., Richardson, W.S., 1998. Understanding articles describing clinical prediction tools. evidence based medicine in critical care group. Crit. Care Med. 26 (9), 1603–1612.
Raney, N.H., Petersen, E.J., Smith, T.A., Cowan, J.E., Rendeiro, D.G., Deyle, G.D., Childs, J.D., 2009. Development of a clinical prediction rule to identify patients with neck pain likely to benefit from cervical traction and exercise. Eur. Spine J. 18 (3), 382–391.
Rethlefsen, M.L., Farrell, A.M., Osterhaus Trzasko, L.C., Brigham, T.J., 2015. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. J. Clin. Epidemiol. 68 (6), 617–626.
Riddle, D.L., 1998. Classification and low back pain: a review of the literature and critical analysis of selected systems. Phys. Ther. 78 (7), 708–737.
Saavedra-Hernandez, M., Castro-Sanchez, A.M., Fernandez-de-Las-Penas, C., Cleland, J. A., Ortega-Santiago, R., Arroyo-Morales, M., 2011. Predictors for identifying patients with mechanical neck pain who are likely to achieve short-term success with manipulative interventions directed at the cervical and thoracic spine. J. Manip. Physiol. Ther. 34 (3), 144–152.
Schellingerhout, J.M., Verhagen, A.P., Heymans, M.W., Koes, B.W., de Vet, H.C., Terwee, C.B., 2012. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. Qual. Life Res. 21 (4), 659–670.
Stanton, T.R., Hancock, M.J., Maher, C.G., Koes, B.W., 2010. Critical appraisal of clinical prediction rules that aim to optimize treatment selection for musculoskeletal conditions. Phys. Ther. 90 (6), 843–854.
Tsakitzidis, G., Remmen, R., Dankaerts, W., van Royen, P., 2013. Non-specific neck pain and evidence-based practice. Eur. Sci. J. 9 (3), 1–19.
van Giessen, A., Peters, J., Wilcher, B., Hyde, C., Moons, C., de Wit, A., Koffijberg, E., 2017. Systematic review of health economic impact evaluations of risk prediction models: stop developing, start evaluating. Value Health J. Int. Soc. Pharmacoecon. Outcomes Res. 20 (4), 718–726.
Veerbeek, J.M., Koolstra, M., Ket, J.C., van Wegen, E.E., Kwakkel, G., 2011. Effects of augmented exercise therapy on outcome of gait and gait-related activities in the first 6 months after stroke: a meta-analysis. Stroke 42 (11), 3311–3315.
Vos, T., Flaxman, A.D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., Memish, Z.A., 2012. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the global burden of disease study 2010. Lancet 380 (9859), 2163–2196.
Wang, W.T., Olson, S.L., Campbell, A.H., Hanten, W.P., Gleeson, P.B., 2003. Effectiveness of physical therapy for patients with neck pain: an individualized approach using a clinical decision-making algorithm. Am. J. Phys. Med. Rehabil. 82 (3), 203–218 quiz 219-21.
Williamson, A., Hoggart, B., 2005. Pain: a review of three commonly used pain rating scales. J. Clin. Nurs. 14 (7), 798–804.