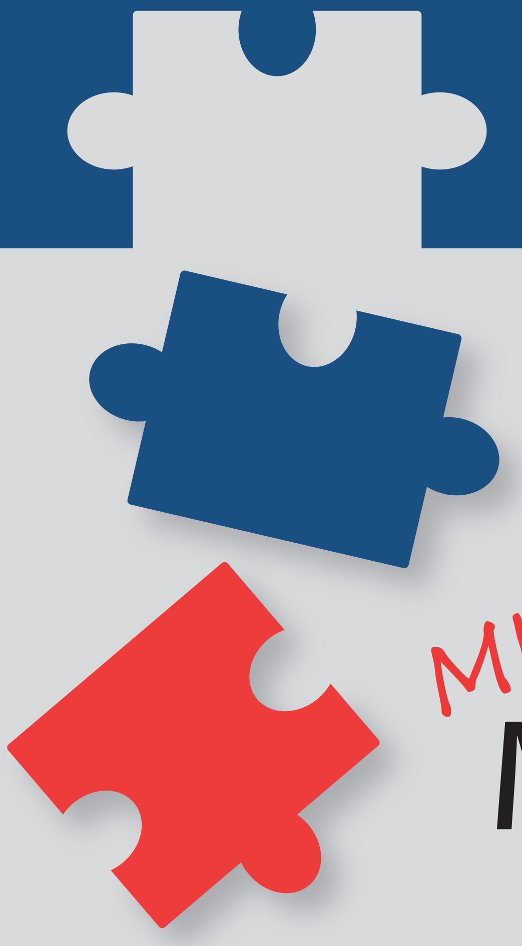
Non-specific neck pain: to match or not to match?

Maissan, Jean François

2022

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

Non-specific neck pain:
to match or not to match?



MIS
MATCH?

Does matching the treatment to diagnose improve outcomes?

François Maissan

Non-specific neck pain: to match or not to match?

Does matching the treatment to diagnose improve outcomes?

François Maissan

Cover: Marieke Schalken

Layout: Marieke Schalken

Printed by: GVO drukkers & vormgevers

ISBN: 978-94-6332-808-1

The research in this thesis was embedded in Amsterdam Movement Sciences Research Institute, at the department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, the Netherlands

The work presented in this thesis was supported by the University of Applied Sciences Utrecht (HU)

The printing of this thesis was financially supported by the Scientific College Physical Therapy (WCF) of the Royal Dutch Society for Physical Therapy (KNGF)

Some terms used in this thesis have been standardized throughout the different chapters.
Therefore, the text might slightly differ from the articles that have been published.

VRIJE UNIVERSITEIT

Non-specific neck pain: to match or not to match?

Does matching the treatment to diagnose improve outcomes?

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Bètawetenschappen
op woensdag 12 oktober 2022 om 15.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door
Jean François Maissan

geboren te Gorinchem

promotor : prof.dr. R.W.J.G. Ostelo

copromotor : dr. J.J.M. Pool
dr. H.M. Wittink

promotiecommissie : dr. S.M. Rubinstein
prof.dr. J.H. van Dieen
prof.dr. B. Cagnie
dr. J.B. Staal
prof.dr. L. Hooft

Table of content

1

General introduction

Introduction

Neck pain is the fourth major cause of disability worldwide [1]. From 1990 to 2010, the disability burden attributable to musculoskeletal disorders (MSK) increased by 46%. It is further expected that this burden will increase in the coming years with the ageing population and an increase in other contributing factors, such as further increase of co-morbidities [2]. In 2010 neck pain was responsible for 20% of the total proportion Years Lived with Disability (YLDs) due to MSK([1]). In 2015, more than one third of a billion people worldwide had neck pain of more than 3 months duration [2], which makes neck pain a serious health threat and also a socio-economic burden on society [3].

In Western Europe, non-specific neck pain has a one-year prevalence of 12 to 71%, among other things depending on the case definitions used, in the general population [4]. The one-year prevalence of neck pain that leads to impaired physical functioning ranges between 1.7 and 11% [5].

A study published in 2003 shows that in the Dutch population aged 25 years and older, the neck, the lower back and the shoulder region are the anatomical regions in which musculoskeletal disorders most frequently occur [6]. Pain, stiffness and/or loss of mobility associated with neck pain often results in health care utilization, such as diagnostic assessments and various treatments [8]. In the Netherlands the costs of care for neck and back complaints amounted to 937 million euros in 2017 [9]. This corresponds to 14% of the total care costs incurred for diseases of the musculoskeletal system and connective tissue.

Since physiotherapists are regularly consulted by people with non-specific neck pain, with substantial associated costs, the question arises whether physiotherapy interventions are effective in this patient population. At least six Cochrane reviews focussing on physiotherapy interventions for patients with neck pain have shown inconclusive evidence for the effect of physiotherapy interventions [10-15]. This observation leads to a number of probing questions: Why is the evidence inconclusive? Are there differential effects of physiotherapy interventions in people with non-specific neck pain? Or are there methodological issues that may (at least partly) explain the inconclusive results in this population?

Evidence-based medicine

Evidence-based medicine (EBM) is "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients" [16]. The aim of EBM is the integration of the best research evidence with clinical expertise and patient values"[16]. The goal of EBM is to improve the quality and consistency of care. The early model of EBM is shown in **Figure 1**.

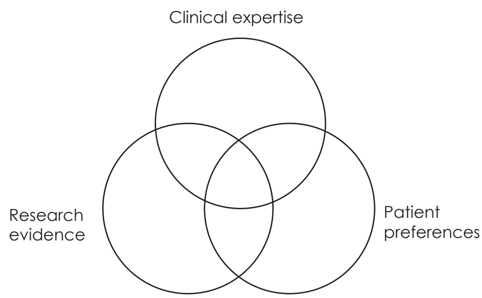


Figure 1: *Early model of the key elements for evidence-based clinical decisions* [17]

The clinical state of the patient and the circumstances under which the patient is located were missing in the first model of Sacket. Haynes updated the model with the element: "clinical state and circumstances" **(Figure 2)**. Patients' clinical state, the clinical setting, and the clinical circumstances they find themselves in when they seek medical attention are key, and often dominant, factors in clinical decisions; "patient preferences" is more broadly defined to include patients' actions and is reversed in position with "research evidence", which signifies its frequent precedence.
Finally, "clinical expertise" is an overarching construct as the means to integrate the other 3 components: thus, constituting a 4th element. Clinical expertise must encompass and balance the patient's clinical state and circumstances, relevant research evidence, and the patient's preferences and actions, to ensure the delivery of best clinical practice [17].

Theoretically, it should be easy to integrate this model into physiotherapy practice. However, sufficient evidence for the application of a specific physiotherapy modality or therapy aiming at a specific subgroup of patients with non-specific neck pain is barely available[18]. The main recommendation in a review of physiotherapy interventions for patients with chronic neck pain was to identify relevant subgroups with matching treatments among patients with non-specific neck pain [18].

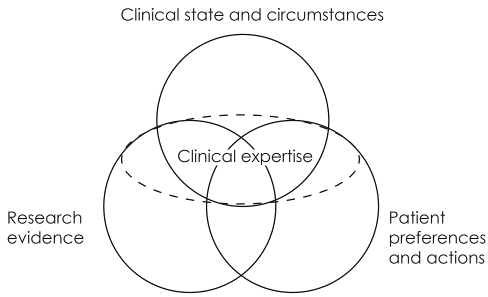


Figure 2: *An updated model for evidence-based clinical decisions* [17]

Physiotherapy

The World Confederation for Physical Therapy has described physiotherapy in a policy statement [19] as: "Physiotherapy provides services by physiotherapists to individuals and populations to develop, maintain and restore maximum movement and functional ability throughout the lifespan. The service is provided in circumstances where movement and function are threatened by ageing, injury, pain, diseases, disorders, conditions or environmental factors and with the understanding that functional movement is central to what it means to be healthy".

Physiotherapists are qualified and professionally required to:

- undertake a comprehensive examination/assessment of the patient/client or needs of a client group
- evaluate the findings from the examination/assessment to make clinical judgments regarding patients/clients
- formulate a diagnosis, prognosis within their expertise and determine when patients/clients need to be referred to another professional
- implement a physiotherapist intervention/treatment programme
- determine the outcomes of any interventions/treatments
- make recommendations for self-management

The first three bullet points describe the diagnostic process to obtain information about the patients' clinical state and circumstances and about patient's preferences and actions (what a patient can or is willing to do) to ultimately arrive at a diagnosis that guides the decision about appropriate intervention(s). The last three bullet points describe the treatment process. It may be clear that the diagnostic process aims to clarify the cause or causes of the health problem with which the patient consults the physiotherapist and that the intervention or interventions are matched to the identified cause(s) [20].

In this dissertation we refer to this match between the diagnostic process and intervention process as "physiotherapeutic validity". In both processes, research evidence should, where possible, be integrated into the clinical decision making of the physiotherapist. Conversely, scientific research must be sufficiently physiotherapeutically valid in order to be able to translate the results of this scientific research into daily practice.

Models applied in physiotherapy

The International Classification of Diseases (ICD) is the classification system used in the medical world for diagnosing and reporting diseases, disorders, injuries and other related health conditions [21].
Physiotherapists, most of the time, do not treat pathology but the consequences of pathology where movement and function are threatened. A theoretical framework is helpful to further operationalize the consequences of a pathology on movement and function. Verbrugge and Jette reported in 1994 that “the disablement process” by Nagi [22] could be used to describe the impact of pathology on physical functioning **(Figure 3)**.

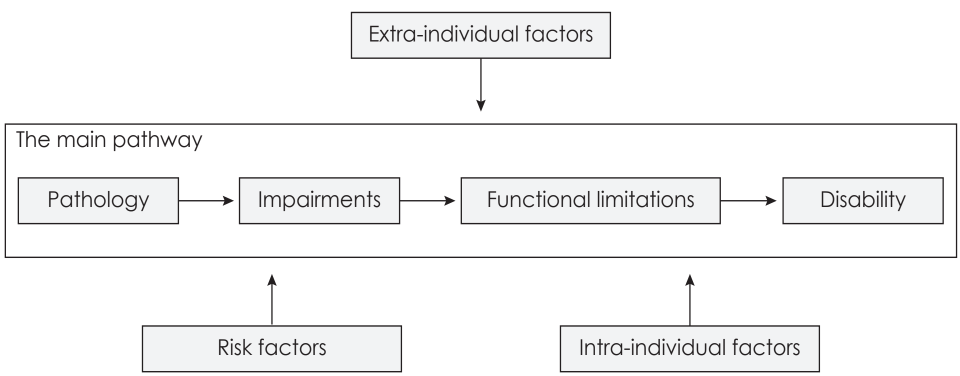


Figure 3: *A model of The Disablement Proces* [22]

In 2001, the World Health Organisation published “the International Classification of Functioning, Disability and Health” (ICF) [23]. The ICF is a framework for describing and organising information on functioning and disability. It provides a standard language and a conceptual basis for the definition and measurement of health and disability **(Figure 4)**. This model has the following domains: body functions and structures, activities and participation. The ICF also lists environmental and personal factors that interact with all these domains. In this manner, it enables the user to record useful profiles of individuals' functioning, disability and health in various domains. In ICF, the term functioning refers to all body functions, activities and participation, while disability is similarly an umbrella term for impairments, activity limitations and participation restrictions.

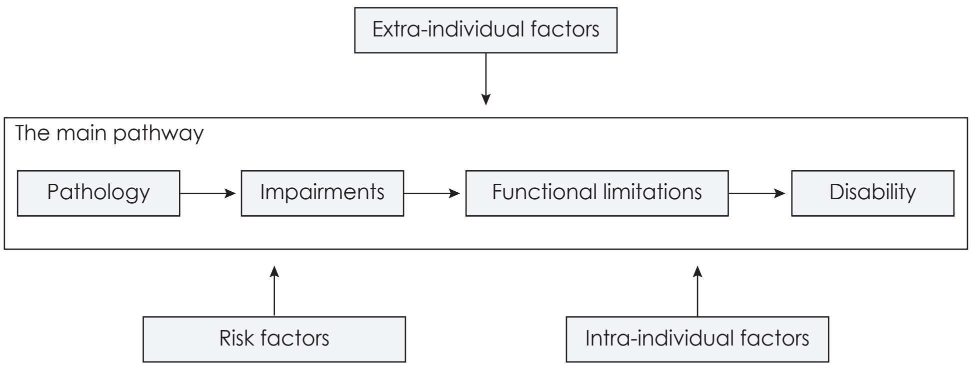


Figure 4*: The ICF model* [23]

The ICF model takes Nagi's patient-centered concept further by expanding on how contextual factors (personal factors and environmental factors) impact physical functioning. This expansion represents an increased focus on how a disease or condition can affect different people in different ways based on the person's unique situation. It also reflects that there is focus on improving health instead of improving (or to cure) the disease.

The ICF framework creates the possibility to describe physical functioning in different domains and how contextual factors influence physical functioning. In order to provide insight into a person's physical functioning, one should strive to measure variables of each domain. Examples of these variables are for:

- Body functions: e.g., mobility of joint functions or muscle functions or movement functions
- Body structure: e.g., structure of the nervous system or structure of extremities
- Activities: e.g., carrying, moving and handling objects or driving
- Participation: e.g., washing oneself or toileting, social activities, work, school, sports

Describing physiotherapy treatment options as variables of the ICF can thus lead to a common language in physiotherapy [24]. The use of outcome measurement tools that measure these variables can in turn help to translate the outcomes into daily practice.

Today, many evidence-based practice guidelines are being developed using the ICF as the basis for describing and classifying care provided by physiotherapists to patients with a variety of musculoskeletal conditions [25]. Although we increasingly speak a common language in physiotherapy, this has not yet resulted in the integration of scientific evidence in our practice [18,26].

In the past twenty-five years 3 approaches (models) for delivering health care have been developed; the stepped care, stratified care and matched care approach [27]. These models reflect daily physiotherapy practice and can be explored in scientific research. Stepped care proposes that more conservative interventions should be tried first, progressing to more complex interventions only when the simpler interventions fail. A limitation is that stepped care may delay proper treatment [27].
The stratified care model attempts to categorize groups of patients into risk categories based on the risk for poor outcome, with matching intervention intensities. Stratified care aims to provide more comprehensive treatment for those with high risk, while allowing those with low risk to recover with little or no treatment. Thus, stratified care individualizes the intervention based on the needs of the group [27]. A limitation of stratified care is that risk identification does not identify the underlying mechanism of why the patient is at low or high risk to develop a chronic condition [27]. Matched care is the new innovation approach in treatment and prevention [28]. Matched care involves identifying those at higher risk; but unlike stratified care, it tailors the intervention to the individual patient's particular risk issues. Therefore, matched care takes also into account individual differences in people's environments, and lifestyles [28]. Of the three models, attention for matched care in the Netherlands is increasing, reports the National Institute for Public Health and the Environment [29]. Possible advantages of matched care are that targeting of treatments more specifically to the individual patient has the potential to revolutionize healthcare delivery through improved effectiveness of treatments and reduced side-effects and associated costs [30]. Matched care is closely aligned with "physiotherapeutic validity" as described earlier. A limitation of matched care is that accurate matching of interventions to profiles/risks is yet under development.

It can therefore be argued that physiotherapy clinical practice should focus on matched care and as such it should be a form of personalized care. In the ideal physiotherapy world, physiotherapists first determine if physiotherapy is indicated for this specific patient. If so, the physiotherapist, as a part of the clinical reasoning process, subgroups this patient, aiming to match treatment to the patients' signs and symptoms, including the results of diagnostic tests [20,31]. In addition, the use of the correct outcome measurement tools, for example ICF variables, is of great importance to monitor and optimise the treatment process. Both the diagnostic process and the treatment process must therefore be recognizable in research into physiotherapy care by the physiotherapist in order to optimize the clinical relevance of research.

To achieve relevant results for physiotherapy in scientific research, matching individual patients with the most appropriate treatment for their profile (matched care), might be of great importance [32] and has consistently been a research priority for the last few years [33].

Despite the fact that matched care seems to be a priority of the last few years, the question remains to what extent matched care takes place in physiotherapeutic scientific research. It is even questionable whether physiotherapy researchers agree on which 'care' suits which patient.

Research into physiotherapy

We continue with a discussion of forms of scientific research in which matched care should be recognizable. Firstly, a Randomized Controlled Trial (RCT) specifically examines the effect of an intervention. Therefore, as a form of matched care, one might expect a link between diagnostics and intervention. More specifically, one expects a link between in - and exclusion criteria and the aim and/or working mechanism of the intervention under research. Secondly, treatment-based classification systems, as a form of matched care, suggest that treatment is based on a certain diagnostic classification system. Finally, one might ask physiotherapy researchers themselves which diagnostic information they use as a basis for their (matched) interventions, for example with a Delphi method.

Randomised Controlled Trials

In a RCT it is of the utmost importance that the included intervention group has, for example, an impairment, an activity limitation and/or a participation restriction which the intervention aims to remedy. Thus, the impairment or activity limitation and/or participation restriction must be present in the included patient group. As such, the impairment or activity limitation must therefore be diagnosed and defined as an inclusion criterion of a RCT. Therefore, a valid and reliable diagnostic process is essential.

Impairments and activity limitations can be measured using questionnaires or through physical tests. However, questionnaires and physical tests do not necessarily measure the same construct [34]. There are patients who report that an activity cannot be performed, but in the exercise room it turns out to be possible, and vice versa. It may be clear that capacity based measurement tools (i.e. physical test) and patient-self-reported outcome measures (i.e. PROMS or questionnaires) of physical function assess different aspects of physical functioning [34-37]. Considering these different aspects in the light of physiotherapeutic validity, it is important to use PROMs or physical tests in RCTs, depending on the construct to be investigated. This construct must then match with daily practice. The "physiotherapeutic validity" of RCTs in people with non-specific neck pain has not yet been subject to scientific scrutiny.

The CONSORT statement for RCTs recommends precise specification of trial processes including details of the intervention being studied or components of that intervention [47]. Despite this recommendation, health care providers in daily practice are not provided with a complete description of the intervention in most RCTs. Glasziou et al demonstrated that in back pain trials, only 13 % of the interventions could be replicated [48]. Given the importance of adequate reporting of interventions in clinical trials, the Template for Intervention Description and Replication (TIDieR) was developed by Hoffman et al. [49]. The most frequently investigated physiotherapeutic interventions in patients with non-specific neck pain are manipulations or mobilizations [50].

Treatment Based Classification Systems
A second way to match a specific intervention to the individual patient is to develop classification systems that suggests specific treatments for specific subgroups, i.e., treatment-based classification systems (TBCSs). In other words, TBCSs should match the outcome of a classification process (i.e., a diagnostic process) to specific intervention(s) [38]. The "physiotherapeutic validity" of TBCSs in people with non-specific neck pain is unknown.

The Delphi method
The Delphi method is an appropriate method when aiming to reach consensus in a field where a lack of agreement or incomplete knowledge is evident [39]. The Delphi method creates the opportunity to gather information from a group of international

scientifically trained experts in treating patients with non-specific neck pain, without the need of a meeting [40]. Scientifically trained experts can be expected to have an unambiguous opinion about the physiotherapeutic validity of research results that they integrate into their daily practice.

Therefore, a Delphi method is suitable for examining expert opinion by establishing whether there is consensus on their clinical reasoning in people with non-specific neck pain.

Clinical practice

In clinical practice, physiotherapists first determine if physiotherapy is indicated for a patient. If so, they then, as a part of the clinical reasoning process, subgroup their patients aiming to match their treatment to the signs and symptoms and results of the diagnostic tests. Another important part of the clinical reasoning process is the use of measurement instruments. Measurement instruments, such as PROMs and physical tests are used to support and objectify the clinical reasoning process. Understanding which measurement instruments are most suitable for use in the diagnostic and / or therapeutic process is desirable. In addition, it is unclear how they support the clinical reasoning process in patients with nonspecific neck pain.

The clinical reasoning process is complex and often includes a combination of interventions. Campbell et al [45] indicated that we first must understand working mechanisms of unimodal interventions (such as mobilizations, strength training or pain education) before combining them into multimodal interventions. Therefore, it is sensible to first reach consensus on the various aspects of the clinical reasoning process when using unimodal interventions in patients with non-specific neck pain. A linear clinical reasoning process consists of three sequential phases: the diagnostic, the therapeutic and the evaluative phase. In this thesis, we defined sequential linear clinical reasoning as the transition from signs and symptoms to diagnostic tests, from diagnostic tests to an intervention with matching treatment goals and the evaluation based on outcome measurements related to the matched goals.

Aim of the dissertation

Physiotherapists should strive to substantiate their choices within the clinical reasoning process by matching the results of diagnostic tools to the intervention they aim to apply. This leads to a fundamental question that physiotherapists should be able to answer: why do they do what they do? This dissertation focuses on this fundamental question.

The general aim of this dissertation is to gain insight into the physiotherapeutic validity of physiotherapy research as in daily practice in subjects with non-specific neck pain.

Outline of the dissertation

The first aim of this dissertation was to systematically explore the literature in order to assess whether the intervention matches the diagnostic process in RCTs and in TBCSs in patients with nonspecific neck pain (NSNP).

High quality RCTs are generally considered to provide the best evidence for interventions as they tend to be highly internally valid. However, in addition to high internal validity, studies must also be of sufficient external validity in order to be able to generalise the results to the population as seen in clinical practice [41]. Several authors have stressed the importance of assessing the clinical relevance of RCTs or external validity, in addition to the internal validity [42,43]. A prerequisite for external validity is a recognizable clinical reasoning process which can be verified and understood by clinicians. ***Chapter 2*** describes the critical appraisal of the clinical reasoning process in RCTs on patients with NSNP.

The lack of proven effective physiotherapy interventions can potentially be explained by heterogeneous research populations. One method of dealing with this heterogeneity is to match treatment more specifically to subgroups of patients with NSNP. 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 [18,38,44]. ***Chapter 3*** describes the critical appraisal of Treatment Based Classification Systems.

The second aim of this dissertation was to examine expert opinion regarding matching interventions to the results of the diagnostic process in patients with non-specific neck pain.

Chapter 4 describes: 1) expert opinion on the indication for physiotherapy when a patient's only problem is pain, without other signs or symptoms or positive diagnostic tests, 2) which measurement instruments are being used by experts to support their clinical reasoning process, for which purpose (diagnostic or evaluative), 3) consensus regarding the use of unimodal interventions, i.e., sequential linear clinical reasoning.

The third aim of this thesis was to investigate, from the perspective of physiotherapeutic validity, the most commonly used physiotherapy intervention, namely manipulations or mobilizations, and their indication in patients with nonspecific neck pain. More specifically, the aims were to assess how these interventions are described in the literature, to investigate the diagnostic accuracy tests that are used to indicate the need for these interventions and, to investigate the effect of mobilisations/manipulations in a population with a matching indication for these interventions.

Chapter 5 describes the critical appraisal of the description of mobilisations and manipulations. The main aim was to assess if these interventions are described in such a way that they are reproducible in daily practice.

After investigating the description of mobilizations and manipulations the diagnostic process leading to these interventions was investigated in ***chapter 6.*** A limited range of motion (ROM) seems to be the most important criterion for experts as an indication for mobilizations or manipulations. Diagnosis of a limited ROM can be made by asking a diagnostic question or questions or by using one or more diagnostic tests. The diagnostic accuracy of a self-reported ROM test of the cervical spine in combination with the best physical examination test per movement direction was investigated.

Finally, in ***chapter 7*** we conducted an exploratory, practice-oriented pilot study into matched treatments in patients with non-specific neck pain. We investigated the change in ROM in patients with non-specific neck pain with and without a limited ROM of the cervical spine after an intervention aimed at improving the ROM.

Chapter 8 addresses the main findings of our research and discusses its implications. Recommendation for future research is presented.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. 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 2012 Dec 15;380(9859):2163-2196.
2. March L, Smith EU, Hoy DG, Cross MJ, Sanchez-Riera L, Blyth F, et al. Burden of disability due to musculoskeletal (MSK) disorders. Best Pract Res Clin Rheumatol 2014 Jun;28(3):353-366.
3. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006 May;10(4):287-333.
4. Haldeman S, Carroll L, Cassidy JD, Schubert J, Nygren A, Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. The Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders: executive summary. Spine (Phila Pa 1976) 2008 Feb 15;33(4 Suppl):S5-7.
5. Guzman J, Hurwitz EL, Carroll LJ, Haldeman S, Cote P, Carragee EJ, et al. A new conceptual model of neck pain: linking onset, course, and care: the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. J Manipulative Physiol Ther 2009 Feb;32(2 Suppl):S17-28.
6. Picavet H.S.J., Schouten J.S.A.G. Musculoskeletal pain in the Netherlands: Prevalences, consequences and risk groups, the DMC3-study. Pain 2003;102(1-2):167-178.
7. Rijksinstutuut voor Volksgezondheid en Milieu (RIVM). Nek- en rugklachten > cijfers & context. 2019;
8. Korthals-de Bos IB, Hoving JL, van Tulder MW, Rutten-van Molken MP, Ader HJ, de Vet HC, et al. Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial. BMJ 2003 Apr 26;326(7395):911.
9. Rijksinstituut voo Volksgezondheid en Milieu (RIVM). Nek- en rugklachten > kosten. 2019;
10. Ezzo J, Haraldsson BG, Gross AR, Myers CD, Morien A, Goldsmith CH, et al. Massage for mechanical neck disorders: a systematic review. Spine 2007 2007;32(3):353-362.
11. Graham N, Gross A, Goldsmith CH, Klaber Moffett J, Haines T, Burnie SJ, et al. Mechanical traction for neck pain with or without radiculopathy. Cochrane Database Syst Rev 2008 Jul 16;(3):CD006408. doi(3):CD006408.
12. Gross A, Forget M, St George K, Fraser MM, Graham N, Perry L, et al. Patient education for neck pain. Cochrane Database Syst Rev 2012 03(3).
13. Gross A, Miller J, D'Sylva J, Burnie SJ, Goldsmith CH, Graham N, et al. Manipulation or mobilisation for neck pain. Cochrane Database Syst Rev 2010 01:N.PAG-N.PAG.
14. Kay TM, Gross A, Goldsmith CH, Rutherford S, Voth S, Hoving JL, et al. Exercises for mechanical neck disorders. Cochrane Database Syst Rev 2015 01(1):N.PAG-N.PAG.

15. Monticone M., Ambrosini E., Cedraschi C., Rocca B., Fiorentini R., Restelli M., et al. Cognitive-behavioral Treatment for Subacute and Chronic Neck Pain: A Cochrane Review. Spine 2015;40(19):1495-1504.
16. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. Clin Orthop Relat Res 2007 Feb;455:3-5.
17. Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient choice. ACP J Club 2002 Mar-Apr;136(2):A11-4.
18. Damgaard P, Bartels EM, Ris I, Christensen R, Juul-Kristensen B. Evidence of Physiotherapy Interventions for Patients with Chronic Neck Pain: A Systematic Review of Randomised Controlled Trials. ISRN Pain 2013 Apr 15;2013:567175.
19. World Confederation for Physical Therapy. Position statement: Description of Physical Therapy. 2011; .
20. Hendriks HJM, Oostendorp R.A.B., Bernards A.T.M., van Ravensberg C.D., Heerkens Y.F., Nelson R.M. The diagnostic proces and indication for physiotherapy: a prerequisite for treatment and outcome evaluation. PHYS THER REV 2000;5:29-47.
21. World Health Organisation. ICD 11 International Classification of diseases for mortality and morbidity statistics.Eleventh revision 2018.
22. Verbrugge LM, Jette AM. The disablement process. Soc Sci Med 1994 Jan;38(1):1-14.
23. World Health Organization. International Classification of Functioning, Disability, and Health (ICF). ICF full version. Geneva, Switzerland. 2001.
24. Jette AM. Toward a common language for function, disability, and health. Phys Ther 2006 May;86(5):726-734.
25. Godges JJ, Irrgang JJ. ICF-based practice guidelines for common musculoskeletal conditions. J Orthop Sports Phys Ther 2008 Apr;38(4):167-168.
26. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials 2017 Mar 14;18(1):122-017-1870-2.
27. Linton SJ, Nicholas M, Shaw W. Why wait to address high-risk cases of acute low back pain? A comparison of stepped, stratified, and matched care. Pain 2018 Dec;159(12):2437-2441.
28. Meiliana A, Dewi NM, Wijaya A. Personalized Medicine: The futuure of health care. Indones Biomed J. 2016;8(3):127-46.
29. Schooneveldt BC, Veldwijk J, Weda M. Application of personalized medicine. 2015;2015-0177.
30. Godman B, Finlayson AE, Cheema PK, Zebedin-Brandl E, Gutierrez-Ibarluzea I, Jones J, et al. Personalizing health care: feasibility and future implications. BMC Med 2013 Aug 13;11:179-7015-11-179.
31. Rothstein JM. Sick and tired of reliability? Phys Ther 2001 Feb;81(2):774-775.
32. Coupe VMH, van Hooff ML, de Kleuver M, Steyerberg EW, Ostelo RWJG. Decision support tools in low back pain. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1084-1097.

33. Foster NE, Hill JC, Hay EM. Subgrouping patients with low back pain in primary care: are we getting any better at it? Man Ther 2011 Feb;16(1):3-8.
34. Wittink H. Functional capacity testing in patients with chronic pain. Clin J Pain 2005 May-Jun;21(3):197-199.
35. Bean JF, Olveczky DD, Kiely DK, LaRose SI, Jette AM. Performance-based versus patient-reported physical function: what are the underlying predictors? Phys Ther 2011 Dec;91(12):1804-1811.
36. Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. J Clin Epidemiol 2006 Feb;59(2):160-167.
37. Lee CE, Simmonds MJ, Novy DM, Jones S. Self-reports and clinician-measured physical function among patients with low back pain: a comparison. Arch Phys Med Rehabil 2001 Feb;82(2):227-231.
38. Fairbank J, Gwilym SE, France JC, Daffner SD, Dettori J, Hermsmeyer J, et al. The role of classification of chronic low back pain. Spine (Phila Pa 1976) 2011 Oct 1;36(21 Suppl):S19-42.
39. Giannarou L, Zervas E. Using Delphi technique to build consensus in practice. Int. journal ofbusiness science and applied management 2014;9(2):66-82.
40. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, et al. Consensus development methods, and their use in clinical guideline development. Health Technol Assess 1998;2(3):i-iv, 1-88.
41. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. A systematic review of the content of critical appraisal tools. BMC Med Res Methodol 2004 Sep 16;4:22.
42. van Tulder M, Furlan A, Bombardier C, Bouter L, Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine (Phila Pa 1976) 2003 Jun 15;28(12):1290-1299.
43. Herbert RD, Bo K. Analysis of quality of interventions in systematic reviews. BMJ 2005 Sep 3;331(7515):507-509.
44. Tsakitzidis G, Remmen R, Dankaerts W, van Royen P. Non-specific neck pain and evidence-based practice. ESJ 2013;9(3):1-19.
45. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696.
46. Hoogeboom TJ, Oosting E, Vriezekolk JE, Veenhof C, Siemonsma PC, de Bie RA, et al. Therapeutic validity and effectiveness of preoperative exercise on functional recovery after joint replacement: a systematic review and meta-analysis. PLoS One 2012;7(5):e38031.
47. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Int J Surg 2011;9(8):672-677.
48. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. BMJ 2010 Aug 13;341:c3852.

49. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348:g1687.
50. Maissan F, Pool JJM, Raaij de E, Mollema J, Ostelo RWJG, Wittink H. The clinical reasoning process in randomised clinical trials with patients with non-specific neck pain is incomplete: A systematic review. Musculoskeletal Science and Practice 2018;35:8-17.

2

The clinical reasoning process in randomised clinical trials with patients with non-specific neck pain is incomplete: A systematic review

François Maissan
Jan Pool
Edwin de Raaij
Jürgen Mollema
Raymond Ostelo
Harriët Wittink

Musculoskeletal Science and Practice 35 (2018) 8–17

Abstract

Objective: Primarily to evaluate the completeness of the description of the clinical reasoning process in RCTs with patients with non-specific neck pain with an argued or diagnosed cause i.e. an impairment or activity limitation. Secondly, to determine the association between the completeness of the clinical reasoning process and the degree of risk of bias.

Data Sources: Pubmed, Cinahl and PEDro were systematically searched from inception to July 2016.

Study Selection: RCTs (n=122) with patients with non-specific neck pain receiving physiotherapy treatment published in English were included.

Data Extraction: Data extraction included study characteristics and important features of the clinical reasoning process based on the Hypothesis-Oriented Algorithm for Clinicians II (HOAC II)].

Data Synthesis: Thirty-seven studies (30%) had a complete clinical reasoning process of which 8 (6%) had a 'diagnosed cause' and 29 (24%) had an 'argued cause'. The Spearmans rho association between the extent of the clinical reasoning process and the risk of bias was -0.2.

Conclusions: In the majority of studies (70%) the described clinical reasoning process was incomplete. A very small proportion (6%) had a 'diagnosed cause'. Therefore, a better methodological quality does not necessarily imply a better described clinical reasoning process.

Key Words: systematic review, neck pain, evidence based medicine, physiotherapy modalities

Introduction

Non-specific neck pain is a major concern in the adult Western world population. A recent review reports a 12-month prevalence ranging from 30%-50%, with activity limitations ranging from 2-11%. About 10% of these patients will develop a chronic pain disorder [1]. Additionally, neck pain poses an important socio-economic burden on society because pain, stiffness or loss of mobility associated with neck pain often results in utilization of diagnostic assessments and treatments [2]. For effective treatment of non-specific neck pain, physiotherapists should be able to rely, within their clinical reasoning process, on the evidence from scientific research. However, scientific research evidence is poorly integrated in physiotherapy [3,4]. One possibility is that RCTs do not reflect "real world" of physiotherapy clinical practice [5,6].

High quality randomized controlled trials (RCTs) are generally considered to provide the best evidence for interventions as they tend to be highly internally valid. Internal validity refers to how "well" the research was performed [7]. High internal validity of the included studies is of paramount importance as this determines the level of confidence for making recommendations for treatment methods. However, in addition to high internal validity, studies must also be of sufficient external validity in order to be able to generalise the results to the population as seen in clinical practice [8]. External validity refers to the "real world" applicability of the research findings or generally the clinical relevance [7]. Several authors have stressed the importance of assessing the clinical relevance of RCTs, in addition to the internal validity [9,10]. A prerequisite for external validity is a recognisable clinical reasoning process which can be verified and understood by clinicians.

An instrument that supports the description of the clinical reasoning process is the Hypothesis-Oriented Algorithm for Clinicians II (HOAC II) [11]. The HOAC II provides a systematic algorithm, consisting of key components, for the clinical reasoning process of physiotherapists. Within this clinical reasoning process, hypothetico-deductive strategies [12] and/or pattern recognition are used [13,14]. In the clinical practice of a physiotherapist a diagnostic strategy is used, which includes history taking and clarification of the patients complaints, i.e. the patient-experienced problems. Next, the physiotherapist needs to generate one or more (alternative) hypotheses as to the cause or causes of the complaint. The HOAC defines the term "cause(s)" as the possible reason(s) for the neck pain or disability; i.e. impairments, limitations in activities or restrictions in participation. These hypotheses guide the physical examination, which

serves to refute or to confirm these hypotheses. The final clinical hypothesis guides the choice for an intervention to eliminate or reduce the cause of the problem. Finally outcome measures should be used to test the clinical hypothesis. Unlike the HOAC II, we consider these outcome measures as twofold:
1) at the level of the patient, i.e. they measure the patients complaint (problem related outcome)
2) at the level of the physiotherapist. i.e. they measure the effect of the intervention (intervention related outcome). In this way, there is a distinction between the immediate effect of the intervention, reflecting the working mechanism of the intervention and, the experienced effect of the patient [15].

A complete clinical reasoning process starts therefore with the physiotherapeutic diagnostic process. Diagnosis in physiotherapy is the result of a clinical reasoning process which results in the identification of existing or potential impairments, limitations in activities and restrictions in participation and of factors affecting functioning positively or negatively [16,17]

The physiotherapist has to determine which impairments, limitations in activities and restrictions in participation are a potential cause or causes of the experienced problem of the patient. The dictionary definition of diagnosis is "the identification of the nature of an illness or other problem by examination of the symptoms" [18]. Therefore, part of the diagnostic process is performing one or more applicable test(s) for identifying a possible cause of the patient experienced problem. In our paper we consider this to be a 'diagnosed cause'. In RCTs these tests should be used to make sure that every participant actually has the assumed cause and can be included in the study. When the diagnostic process only consists of propositions, of what could be a cause, without testing, we consider this an 'argued cause'. In RCTs this argumentation is often found in the introduction section. Hence, the main difference between a study with a 'diagnosed cause' and an 'argued cause' is that in the "argued cause" studies it is possible that the study sample did not have an impairment or activity limitation at all, despite a complaint of pain. In research it is of great importance to know if the population under research actually did have the impairment or activity limitation the intervention intents to influence. Without the presence of an impairment or activity limitation, there is no need to intervene. This is why, unlike the HOAC, we distinguish between a physiotherapeutic 'diagnosed cause' and 'argued cause'.

Therefore the key components of the physiotherapeutic clinical reasoning process based on the HOACII and extended with our distinction between problem versus goal-related outcome and diagnosed versus argued cause are:

- a patient experienced problem (the complaint)
- a cause (either diagnosed or argued)
- a goal aimed at the diagnosed impairment, activity limitation or restriction in participation.
- a matched intervention to the goal
- an outcome measure related to the diagnosed cause (intervention related outcome)
- an outcome measure related to the patient's experienced problem (problem related outcome)

The assessment of the clinical relevance is increasingly important as evidenced by the updated method guideline for systematic reviews in the Cochrane back and neck group [19]. Now they recommend to specifically describe the type, intensity, dosage, frequency and duration of treatment. However, there is still little attention to the clinical reasoning process. Consequently, it remains unclear if risk of bias of a study is associated with the extent to which this study used (and described) a clinical reasoning process.

Therefore, the research questions are:

- *Are the key components of the clinical reasoning process described within the methodology of RCTs on patients with non-specific neck pain?*
- *How many studies with a complete clinical reasoning process have a diagnosed cause?*
- *What is the association between the extent of a complete clinical reasoning process and the risk of bias?*

Methods

This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [20].

Data sources and Searches

A comprehensive literature search was performed in MEDLINE, CINAHL and PE-

Dro from inception to July 2016. The search was completed in collaboration with a medical information specialist (JM)[21]. A sensitive search strategy was developed for MEDLINE with the acceptance of false positive findings **(Appendix 1)**. To collect as many potentially eligible RCTs as possible, the search strategy combined two primary pathways. The first combined neck pain with physiotherapy and the second concerned the combination neck pain with the subheadings "rehabilitation", "therapy" and "prevention and control" because these subheadings included most likely also physiotherapy. The first and second pathways were combined with the Boolean term "OR". Subsequently, the outcome was limited for RCTs with the "Cochrane Highly Sensitive Search Strategy" for identifying randomized trials". In CINAHL the same strategy was used as in MEDLINE with an adapted Cochrane search strategy. In PEDro the Abstract and Title box was filled with "neck", the problem box with "pain" and the method box with "clinical trial".

The selection process and data extraction were performed by two independent researchers. The titles and abstracts were judged by these researchers based on the in- and exclusion criteria. Full text was reviewed for hits that could not be excluded based on title/abstract. After independently selecting the studies, they discussed differences until consensus was reached. If no consensus was reached, a third researcher (HW) was consulted and consensus was reached based on discussion between them.

Study selection

A study was included if it met the following criteria: full-text original article, published in English, adult patients (>18 years old) with non-specific neck pain, mono disciplinary physiotherapy intervention and randomized controlled trial (RCT). RCTs with mixed population were included if the clinical reasoning process was described specifically for patients with non-specific neck pain instead of a mixed population. Non-specific neck pain was defined as pain (with or without radiation) located in the cervical spine and/or occiput region and/or cervico thoracic junction and muscles originating from the cervical region acting on the head and shoulders, without underlying pathology, such as: trauma (fractures), infection, inflammatory disorders, neurologic pathology or systemic disease [1].

A study was excluded if: if the study was performed in patients with headache with or without non-specific neck pain, temporomandibular joint dysfunctions or trigger points in the trapezius region or trapezius myalgia. Also studies in patients with whiplash related neck pain were excluded.

Data extraction and Quality assessment

Risk of bias was assessed using the PEDro scale [22]. The Intra-class Correlation Coefficient for consensus ratings is 0.68 (95% confidence interval 0.57-0.76) executed by experienced assessors; therefore ratings from the physiotherapy evidence database (www.pedro.org.au) were used [23]. We considered a cut-off score of ≥ 6 as high quality [24]. When no score was available in the PEDro database, two authors independently assessed the risk of bias.

Two a-priori data extraction forms were developed for this review. One form to score patient and study characteristics of the RCTs **(Appendix 2)** and the other to score the HOAC II based clinical reasoning process rating scale **(Table 1)**. To determine the completeness of the clinical reasoning process a 6-item scale was developed based on the HOAC II **(Table 1)**. Two independent raters scored the RCTs on this scale. Differences were discussed until consensus was reached.

Data synthesis and analysis

We rated a clinical reasoning process complete if 1. an experienced problem was described, 2. a cause was 'diagnosed' or 'argued', 3. the main goal of the intervention was related to the 'cause', 4. the intervention matched the main goal, 5. the intervention related outcome measure matched the main goal of the physiotherapist and 6. the problem related outcome measure matched the patient-experienced problem **(Table 1)**. The rating scale is described in **Table 1**.

For each score on the HOAC II based clinical reasoning process rating scale, there was a prerequisite: there had to be a "+" score on the preceding item.
Without a clearly defined cause, it is not possible to define a clear goal and for that reason it is not possible to match the intervention with intervention related outcome measures. Therefore, all 6 items should be scored with at least "+" or "?" before we scored the clinical reasoning process as complete.

Spearmans rho was calculated, to determine the association between PEDro scores and the number of positive items on the HOAC II based clinical reasoning process rating scale, using the software package of IBM SPSS Statistics 22.0 (SPSS Inc., Chicago, Il).

Table 1: *HOAC II based clinical reasoning process rating scale.*

	Items		Score
1	Is a patient-experienced problem described?	+	A patient-experienced problem, for example pain or activity limitation must be described as an inclusion criterion.
		-	A patient-experienced problem is not described as an inclusion criterion.
2	Is the cause of the problem diagnosed or argued?	++	A cause is 'diagnosed' if a test is used to determine the cause of the patient-experienced problem and that this test is described as an inclusion criterion.
		+	A cause is 'argued' if the argumentation is described in the introduction section but no further objectification took place as an inclusion criterion
		?	A cause is unclear if the argumentation described in the introduction is multi interpretable.
		-	A cause is not described.
3	Is the main goal of the intervention(s) related to the cause ? (as described in 2)	+	The main goal of the intervention should be to eliminate the 'argued' or 'diagnosed' cause.
		-	The main goal is not to eliminate the argued or diagnosed goal.
4	Does the intervention(s) match the main goal ? (as described in 3)	+	The intervention should be aimed at achieving the main goal.
		-	The intervention is not focused on the main goal.
5	Does the intervention related outcome measure match the direct goal? (as described in 3)	+	The intervention related outcome measure should measure the change of the cause.
		-	There is no outcome measure that measure the change of the cause
6	Does the problem related outcome measure match the patient experienced problem ? (as described in 1)	+	The problem related outcome measure should measure the change of the experienced problem by the patient.
		-	There is no outcome measure that measure the chance in the patient-experienced problem.

Results

The literature search retrieved 2799 studies. After removing the duplicates, 2331 remained for further screening. **Figure 1** describes the screening process. One hundred and twenty-two studies were included [25 -146].

Appendix 2 gives an overview of the participant and study characteristics. Sample sizes varied from 9 [74] to 393 [137] participants. Recruitment took place in various ways, for example by newspaper advertisement or recruitment from different kind of clinics. There were more female than male participants in the study populations. Twenty-three (19%) studies included only females pursuing a homogeneous study population. Ninety-eight (80%) studies included participants with chronic neck pain [25, 27-32, 34-36, 38, 39, 41-43, 47-50, 53-64, 67, 69-74, 76, 77, 79-87, 90, 92-94, 97, 99-108,110-116, 119-121, 123-135, 137-146].

All RCTs reported pain as the most experienced problem by the participants [25-146]
Of the 122 studies thirty-seven studies, (30%) scored a complete clinical reasoning process **(Figure 2)**. Fifty-six studies (46%) scored "-" on item 2 (cause) of the rating scale and therefore, the problem related outcome, matched intervention and intervention related outcome measures also scored negative [25-27, 32, 35, 36, 43, 45, 47-49, 54-57, 65, 67, 68, 71-74, 76, 78, 81-83, 91-95, 103,105,108, 110,114,117, 119, 120, 124, 125, 129,130, 132, 134, 135,137,138,140, 142, 144-146].
Sixty-six RCTs (54%) described a cause of the experienced problem.[29-31, 33, 34,37-42, 44, 45, 50-53, 58-64, 66, 67, 69, 71, 75, 77, 79, 80, 84-90, 96-100, 102, 104, 107, 111-113, 115, 116 ,118, 121-123, 126-128, 131, 133, 136, 139, 141, 143].
Forty-six studies (38%) had an 'argued' or 'unclear' cause (the argued cause pathway) [29, 34, 37-42, 44, 51, 53, 58-64, 66, 67, 69, 70, 75, 77, 84, 85, 88, 89, 98, 100, 102, 107, 112, 113, 116, 121-123,127,128, 131,133, 136,139,141,142].
Twenty studies (16%) scored a 'diagnosed' cause (the diagnosed cause pathway)[30 ,31, 33, 46, 50, 52, 79, 80, 86, 87, 90, 96, 97, 99, 104 ,106, 111, 115, 118, 126].

The researched population in these 20 (16%) studies with a "diagnosed cause" actually had the impairment or activity limitation that the intervention intended to improve. However 5 (4%) RCTs had no cause related goal and thereafter, 4 (3%) no intervention related outcome measures. Therefore 11 (9%) of the included studies had a diagnosed cause with at least one intervention related outcome measure [30, 31, 46, 50, 79, 80, 86, 96, 97, 104, 111]. Of these 11 studies 8 (6%) presented also problem outcome measures and therefore completed the entire clinical reasoning process [30,31,46,50,79,80,86,97]. The detailed score of the components of the clinical reasoning process is described in **Appendix 3**.

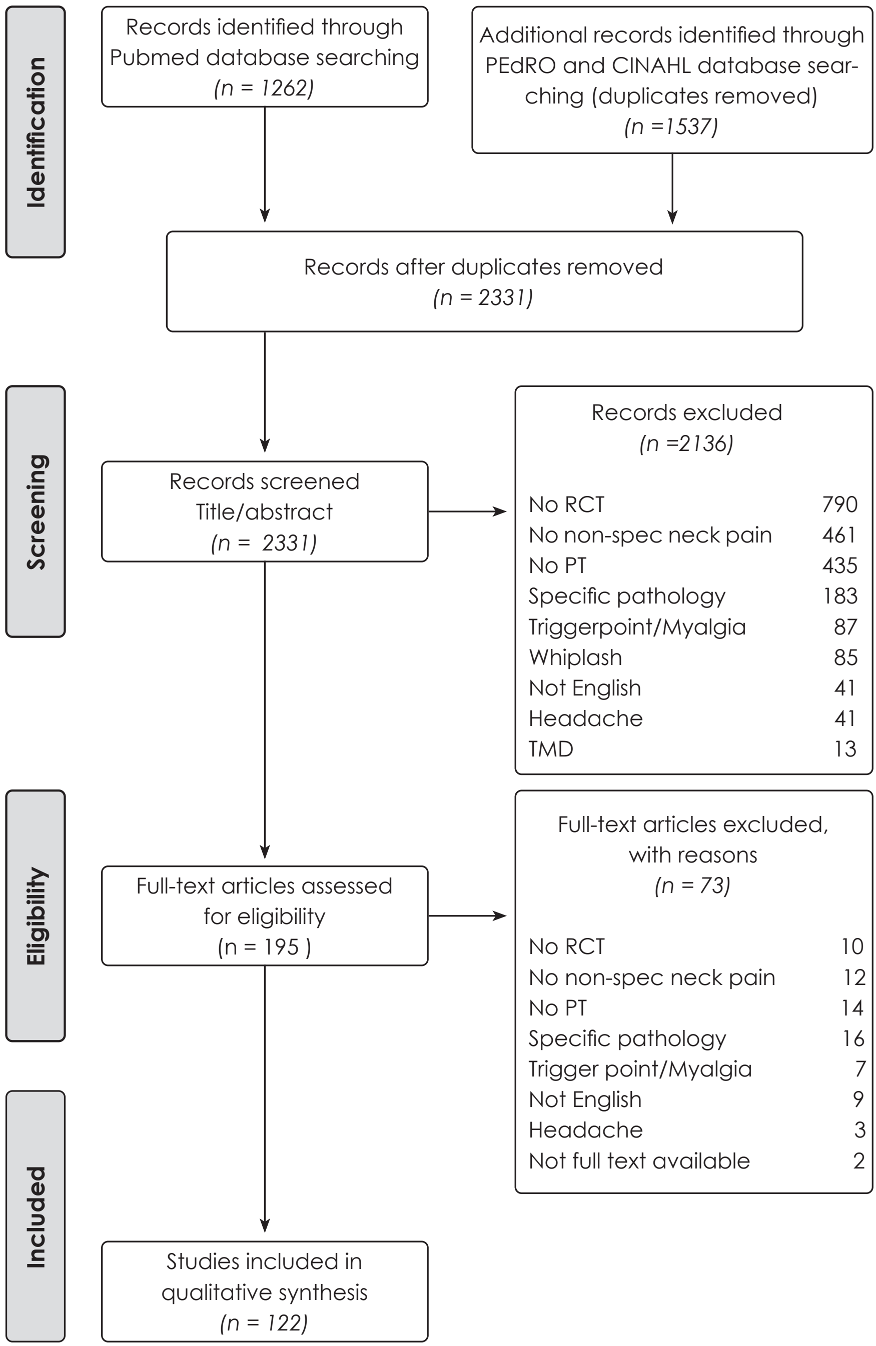


Figure 1: *Flowchart of articles reviewed.*

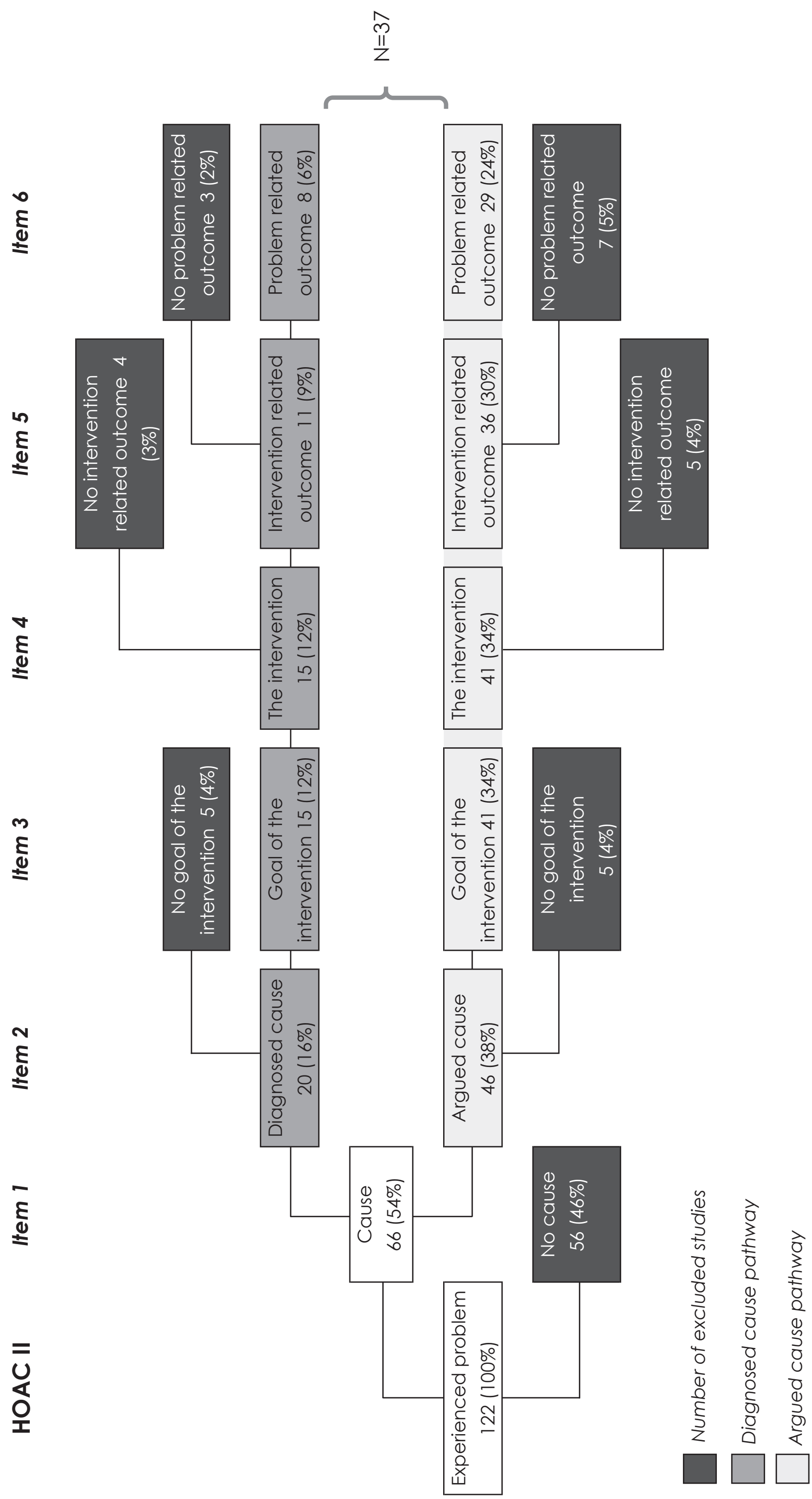


Figure 2: *HOAC II clinical reasoning process rating outcome*

The PEDro scores ranged from 2 to 10 with a median of seven. Because the majority, 87 (71%) of the studies, scored ≥ 6, the overall methodological quality was high. Of five studies no score was available in the PEDro database [70, 71,81 82, 124]. Therefore the first two authors assessed the Risk of Bias. Finally, there was a small negative correlation between PEDro scores and the number of positive items on the HOAC II based clinical reasoning process rating scale (spearmans rho -0.2).

Discussion

This review illustrates that the minority of studies (n=37; 30%) describe the complete clinical reasoning process, and that only a very small proportion of these studies with a complete clinical reasoning process (n=8; 6%) had a ‘diagnosed cause’. In fact, the HOAC II key-component most frequently missing was the “cause” (either diagnosed or argued), with nearly half of the studies not describing any cause at all. It could be argued that these are the ultimate “trial and error” RCTs because even an argued cause, that is, an argued reason why the intervention could be effective, is missing. The HOAC II key-component most frequently described is the “intervention”. This means that in all the included RCTs with a cause, the interventions were described in terms of a cause matching the predefined goals.

Only 11 (9%) of the included studies had a diagnosed cause with at least one intervention related outcome measure [30, 31, 46, 50, 79, 80, 86, 96, 97, 104, 111]. These studies make it possible to understand the clinical reasoning process used for the choice of the intervention and what the intervention aimed to achieve (the goal). In contrast to studies with a ‘diagnosed’ cause, in studies with “an argued cause” it remained unclear what the impairment, activity limitation or restriction in participation was. Thus, it is possible that in these studies the population did not have an impairment, activity limitation or restriction in participation at all. To illustrate; there were 5 studies aiming to improve neck Range of Motion (ROM), but the authors did not find any improvement in ROM [70, 71, 81, 82, 124]. However, their conclusion that the intervention had no effect on ROM can be questioned as ROM at baseline was equal to norm values [147]. This could occur because a diagnosed ROM limitation was not used as an inclusion criterion. Although some participants could have a ROM limitation, the possibility remains limited to achieve a good result if norm values are measured at baseline. This example clearly emphasises the need to define and measure specific impairments, activity limitations or restriction in participation as inclusion criteria for participants.

Recently, Hoffmann et al made recommendations to enhance the usability of systematic reviews[148]. Within the PICO (Patient/Intervention/Comparison/Outcome) format the intervention should be given as much consideration as the other components. They recommend the use of their Template for Intervention Description and Replication (TIDierR) checklist [149]. The TIDierR checklist and guide was published with the specific aim of improving the completeness of reporting and ultimately the replicability of interventions. The authors included an item into the TIDieR checklist to describe any rationale, theory, or goal of the elements essential to the intervention [149]. This is to gain insight into the working mechanism of the intervention. They also state that "the known or supposed mechanism of action of the active components of the intervention should be described because if the active components of the intervention were omitted, the intervention would be ineffective" as demonstrated in our 'ROM' example.

Added to this, there is now a consensus statement about reporting spinal manipulative therapy including the item "rationale of the therapy" [150]. This also underpins the need for a diagnosed impairment or activity limitation with matching goal of the intervention and intervention related outcome to understand if an intervention is to be effective and further understand its working mechanism. This knowledge is of the utmost importance for the physiotherapist to make evidence based decisions during the clinical reasoning process and, this knowledge is lacking in 91% of the RCTs included in this review.

Finally we assessed if the risk of bias and clinical reasoning were correlated. There was a small negative correlation of -0.2 of the PEDro scores with extent of the clinical reasoning process. The negative score implies that lower risk of bias is associated with lower complete clinical reasoning. These finding indicates that a better methodological quality does not necessarily imply a better clinical reasoning process. As stated earlier, the updated method guideline for systematic reviews in the Cochrane back and neck group strongly advises the use of the TIDieR checklist for describing the intervention [19]. However, the clinical reasoning process is broader and was more optimally represented in the previous edition in the Cochrane back and neck group guideline [9]. For the next version of the guideline we strongly advise to consider incorporating assessment of the clinical reasoning process, or otherwise to at least include a description of the diagnostic process, so it becomes possible to asses if the population under research had the impairment or activity limitation that the intervention intended to improve.

This systematic review (SR) has limitations. Firstly, we did not request additional information from authors. Authors may not have reported clinical reasoning while in fact it did take place. Therefore it is possible that a negative score was given despite the fact that clinical reasoning has taken place. However, the main omission in the scored clinical reasoning processes was the diagnostic process. Diagnostic inclusion criteria objectify the assumed cause of the experienced problem and were used in only 16% of the RCTs. In addition, it is unlikely that authors forget to mention inclusion criteria. Hence, it is not expected that an unfair negative score, due to unreported clinical reasoning, will often occur.

Secondly, It is possible that researchers have adopted a different framework or model that underpins the choice of the intervention. However many other models of clinical reasoning in physiotherapy all use a diagnostic process to substantiate the choice of intervention[12,151-153]. Furthermore, the WCPT policy statement: "Description of physical therapy" stated that: "physiotherapist are professionally required to undertake a comprehensive examination/assessment of the patient/client", thereby clearly illustrating that a diagnostic process is a conditional part of the physiotherapeutic process [17]. Despite the importance of the diagnostic process, our review highlights that the absence of a diagnostic process is the main omission in the included studies.

Thirdly, we realise that there is no Gold Standard for clinical reasoning. We developed a scoring list by using the HOAC II steps. The HOAC II has two advantages. First, it is compatible with "the guide to physiotherapist practice's [11]. This ensures that the HOAC is in line with daily physiotherapy practice. Second, in the structure of the HOAC II the hypothetico-deductive reasoning model is incorporated. The advantage here is that this model has its roots in the empirical-analytical research paradigm matching the RCT methodology [12]. In summary, the HOAC II is consistent with the physiotherapy process and in line with the RCT methodology. In addition, although scoring the clinical reasoning process is subjective, by using the HOAC as a scoring tool we are confident that the scorings system is at least more transparent. Finally, we only judged whether the key components were present, not whether the components were valid. This could be subject to further study.

A strength of this study is the large number of included studies. As we anticipated finding a large body of RCTs as we used a sensitive search strategy, strict inclusion and exclusion criteria were applied and myalgia, whiplash and headache were excluded in order to include a homogeneous population with non-specific neck pain. Because of the sensitive search it is not expected that many studies have been missed, or that these missed studies (if any) will have a substantial impact on our main findings. Another strength was the use of the PEDro ratings. The reliability of PEDro scores is known for trained raters [23]. Therefore, we adopted the scores from the PEDro organisation website, because trained raters performed their ratings. This way we made sure that the listed scores are of sufficient reliability.

To our knowledge, this is the first systematic review that provides an overview of the completeness of the clinical reasoning process of RCT's in patients with non-specific neck. In the non-specific neck pain literature we found only one review of Kjellmann et al with similarities to this study[150]. It concerned patients with neck pain but they also included specific pathology. In contrast to our study, they evaluate the inclusion criteria, intervention and outcome measures. They reported that no study used functional limitations as an inclusion criterion. In fact, none of their included RCTs had a diagnosed problem as an inclusion other than a specific pathology. They also found a great diversity in interventions and that mostly PROMS were used as outcome measures with the exception of ROM as a regularly used impairment outcome measure. Our study more or less confirms these findings. Despite a different research population, a study of Hoogeboom et al also shows similarities with our study[154]. In contrast to our study they scored part of the clinical reasoning process where they specifically targeted the validity of the intervention. The best comparable item was the match between the diagnosed cause and the intervention. They scored a match in 8% of the studies, which is quiet comparable with our score of 12%.

Future research should focus on all key elements discussed in this review. Diagnostic tests should be reported as inclusion criteria with their matching interventions. In addition, measurement properties of these tests should be reported. This is equally important for the reporting of appropriate outcome measures, which should include both intervention and patient related outcomes. For example; two studies with a complete clinical reasoning process about endurance training showed good results on intervention related outcome measures however, poor results on problem related outcome measures [30,80]. The use of problem related outcome measures could have led, unjustly, to the conclusion that this intervention had no effect. This underpins the importance of using both types of outcomes measures.

The outcome of diagnostic tests should lead to relevant subgroups matching the chosen intervention. This fits within the current discussion about subgroups and classification systems and the need to develop targeted treatments for known impairments and activity limitations or developed classification systems for patients with non-specific neck pain [155-158]. We hope that this review contributes to the subgroup discussion.

In summary a complete line of clinical reasoning appears to be of paramount importance for the examination of a specific intervention with its matching specific effect in order to understand working mechanisms of interventions. In general, this study was a first step to provide insight in the completeness of the clinical reasoning process within RCTs on non-specific neck pain.

In conclusion: In the majority of studies no complete clinical reasoning process was described, therefore lacking, to a large extent, the external validity. A very small proportion (9%) had an diagnosed cause with a matching intervention and intervention related outcome measures, thereby determining what needs to be treated and if the goal of the intervention was reached. Finally, the small negative correlation between the extent of the clinical reasoning process and the risk of bias, indicates that a better methodological quality does not necessarily imply a better clinical reasoning process.

2

References

1. Hogg-Johnson S, van der Velde G, Carroll LJ, Holm LW, Cassidy JD, Guzman J, 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. Spine 2008;33(4 Suppl):S39-51.
2. Korthals-de Bos IB, Hoving JL, van Tulder MW, Rutten-van Molken MP, Ader HJ, de Vet HC, et al. Cost effectiveness of physiotherapy, manual therapy, and general practitioner care for neck pain: economic evaluation alongside a randomised controlled trial. BMJ 2003;326(7395):911.
3. Greenhalgh T, Howick J, Maskrey N, Evidence Based Medicine Renaissance Group. Evidence based medicine: a movement in crisis? BMJ 2014;348:g3725.
4. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials 2017;18(1):122-017-1870-2.
5. Balague F, Mannion AF, Pellise F, Cedraschi C. Non-specific low back pain. Lancet 2012;379(9814):482-491.
6. Tsakitzidis G, Remmen R, Dankaerts W, van Royen P. Non-specific neck pain and evidence-based practice. ESJ 2013;9(3):1-19.
7. Eldridge S, Ashby D, Bennett C, Wakelin M, Feder G. Internal and external validity of cluster randomised trials: systematic review of recent trials. BMJ 2008;336(7649):876-880.
8. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar S, Grimmer KA. A systematic review of the content of critical appraisal tools. BMC Med Res Methodol 2004;4:22.
9. van Tulder M, Furlan A, Bombardier C, Bouter L, Editorial Board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine 2003;28(12):1290-1299.
10. Herbert RD, Bo K. Analysis of quality of interventions in systematic reviews. BMJ 2005;331(7515):507-509.
11. Rothstein JM, Echternach JL, Riddle DL. The Hypothesis-Oriented Algorithm for Clinicians II (HOAC II): a guide for patient management. Phys Ther 2003;83(5):455-470.
12. Edwards I, Jones M, Carr J, Braunack-Mayer A, Jensen GM. Clinical reasoning strategies in physical therapy. Phys Ther 2004;84(4):312-30; discussion 331-5.
13. (13) Norman G. Research in clinical reasoning: past history and current trends. Med Educ 2005;39(4):418-427.
14. Rushton A, Lindsay G. Defining the construct of masters level clinical practice in manipulative physiotherapy. Man Ther 2010;15(1):93-99.
15. Lee H, Mansell G, McAuley JH, Kamper SJ, Hubscher M, Moseley GL, et al. Causal mechanisms in the clinical course and treatment of back pain. Best Pract Res Clin Rheumatol

2016:30(6):1074-1083.
16. Guccione AA. Physical therapy diagnosis and the relationship between impairments and function. Phys Ther 1991;71(7):499-503; discussion 503-4.
17. World Confederation for Physical Therapy. Position statement: Description of Physical Therapy. 2011. Retrieved from www.wcpt.org/policy/ps-descriptionPT
18. Oxford University Press. Oxford learner's dictionaries. 2017; Available at: http://www.oxfordlearnersdictionaries.com/.
19. Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. Spine 2015;40(21):1660-1673.
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open Med 2009;3(3):e123-30.
21. Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. J Clin Epidemiol 2015;68(6):617-626.
22. de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. Aust J Physiother 2009;55(2):129-133.
23. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther 2003;83(8):713-721.
24. Veerbeek JM, Koolstra M, Ket JC, van Wegen EE, Kwakkel G. Effects of augmented exercise therapy on outcome of gait and gait-related activities in the first 6 months after stroke: a meta-analysis. Stroke 2011;42(11):3311-3315.
25. Akhter S, Khan M, Ali SS, Soomro RR. Role of manual therapy with exercise regime versus exercise regime alone in the management of non-specific chronic neck pain. Pak J Pharm Sci 2014;27(6 Suppl):2125-2128.
26. Ali A, Shakil-Ur-Rehman S, Sibtain F. The efficacy of Sustained Natural Apophyseal Glides with and without Isometric Exercise Training in Non-specific Neck Pain. Pak J Med Sci 2014;30(4):872-874.
27. Andrade Ortega JA, Ceron Fernandez E, Garcia Llorent R, Ribeiro Gonzalez M, Delgado Martinez AD. Microwave diathermy for treating nonspecific chronic neck pain: a randomized controlled trial. The Spine Journal 2014;14(8):1712-1721.
28. Aquino RL, Caires PM, Furtado FC, Loureiro AV, Ferreira PH, Ferreira ML. Applying joint mobilization at different cervical vertebral levels does not influence immediate pain reduction in patients with chronic neck pain: a randomized clinical trial. J Man Manipulative Ther 2009;17(2):95-100.
29. Bakar Y, Sertel M, Ozturk A, Yumin ET, Tatarli N, Ankarali H. Short term effects of classic massage compared to connective tissue massage on pressure pain threshold and muscle relax-

ation response in women with chronic neck pain: a preliminary study. J Manipulative Physiol Ther 2014;37(6):415-421.
30. Beer A, Treleaven J, Jull G. Can a functional postural exercise improve performance in the cranio-cervical flexion test?--a preliminary study. Man Ther 2012;17(3):219-224.
31. Beinert K, Taube W. The effect of balance training on cervical sensorimotor function and neck pain. J Mot Behav 2013;45(3):271-278.
32. Beltran-Alacreu H, Lopez-de-Uralde-Villanueva I, Fernandez-Carnero J, La Touche R. Manual Therapy, Therapeutic Patient Education, and Therapeutic Exercise, an Effective Multimodal Treatment of Nonspecific Chronic Neck Pain: A Randomized Controlled Trial. Am J Phys Med Rehabil 2015;94(10 Suppl 1):887-897.
33. Bid D, Ramalingam A, T., Bhatt J, A., Rathod P, N., Tandel K, V., Tandel S, S. The effectiveness of Mechanical Cervical Traction on Patients with Unilateral Mechanical Neck Pain. Indian J Physiother Occup Ther 2014;8(3):97-103.
34. Borisut S, Vongsirinavarat M, Vachalathiti R, Sakulsriprasert P. Effects of Strength and Endurance Training of Superficial and Deep Neck Muscles on Muscle Activities and Pain Levels of Females with Chronic Neck Pain. J Phys Ther Sci 2013;25(9):1157-1162.
35. Borman P, Keskin D, Ekici B, Bodur H. The efficacy of intermittent cervical traction in patents with chronic neck pain. Clin Rheumatol 2008;27(10):1249-1253.
36. Brage K, Ris I, Falla D, Sogaard K, Juul-Kristensen B. Pain education combined with neck- and aerobic training is more effective at relieving chronic neck pain than pain education alone--A preliminary randomized controlled trial. Man Ther 2015;20(5):686-693.
37. Briem K, Huijbregts P, Thorsteinsdottir M. Immediate effects of inhibitive distraction on active range of cervical flexion in patients with neck pain: a pilot study. J Man Manip Ther 2007;15(2):82-92.
38. Casanova-Mendez A, Oliva-Pascual-Vaca A, Rodriguez-Blanco C, Heredia-Rizo AM, Gogorza-Arroitaonandia K, Almazan-Campos G. Comparative short-term effects of two thoracic spinal manipulation techniques in subjects with chronic mechanical neck pain: a randomized controlled trial. Man Ther 2014;19(4):331-337.
39. Celenay ST, Akbayrak T, Kaya DO. A Comparison of the Effects of Stabilization Exercises Plus Manual Therapy to Those of Stabilization Exercises Alone in Patients With Nonspecific Mechanical Neck Pain: A Randomized Clinical Trial. J Orthop Sports Phys Ther 2016;46(2):44-55.
40. Celenay ST, Kaya DO, Akbayrak T. Cervical and scapulothoracic stabilization exercises with and without connective tissue massage for chronic mechanical neck pain: A prospective, randomised controlled trial. Man Ther 2016;21:144-150.
41. Chiu TT, Lam TH, Hedley AJ. A randomized controlled trial on the efficacy of exercise for patients with chronic neck pain. Spine 2004;30(1):E1-7.
42. Chiu TT, Hui-Chan CW, Chein G. A randomized clinical trial of TENS and exercise for patients

with chronic neck pain. Clin Rehabil 2005;19(8):850-860.

43. Chiu TT, Ng JK, Walther-Zhang B, Lin RJ, Ortelli L, Chua SK. A randomized controlled trial on the efficacy of intermittent cervical traction for patients with chronic neck pain. Clin Rehabil 2011;25(9):814-822.
44. Cleland JA, Childs JD, McRae M, Palmer JA, Stowell T. Immediate effects of thoracic manipulation in patients with neck pain: a randomized clinical trial. Man Ther 2005;10(2):127-135.
45. Cleland JA, Glynn P, Whitman JM, Eberhart SL, MacDonald C, Childs JD. Short-term effects of thrust versus nonthrust mobilization/manipulation directed at the thoracic spine in patients with neck pain: a randomized clinical trial. Phys Ther 2007;87(4):431-440.
46. Cleland JA, Mintken PE, Carpenter K, Fritz JM, Glynn P, Whitman J, et al. 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 2010;90(9):1239-1250.
47. Cook AJ, Wellman RD, Cherkin DC, Kahn JR, Sherman KJ. Randomized clinical trial assessing whether additional massage treatments for chronic neck pain improve 12- and 26-week outcomes. Spine J 2015;15(10):2206-2215.
48. Cunha AC, Burke TN, Franca FJ, Marques AP. Effect of global posture reeducation and of static stretching on pain, range of motion, and quality of life in women with chronic neck pain: a randomized clinical trial. 2008;63(6):763-770.
49. David J, Modi S, Aluko AA, Robertshaw C, Farebrother J. Chronic neck pain: a comparison of acupuncture treatment and physiotherapy. Br J Rheumatol 1998;37(10):1118-1122.
50. Dawood RS, Kattabei OM, Nasef SA, battarjee KA, Abdelraouf OR. Effectiveness of Kinesio Taping versus Cervical Traction on Mechanical Neck Dysfunction. Int J Therap Rehabil Res 2013;2(2).
51. de Camargo VM, Alburquerque-Sendin F, Berzin F, Stefanelli VC, de Souza DP, Fernandez-de-las-Penas C. Immediate effects on electromyographic activity and pressure pain thresholds after a cervical manipulation in mechanical neck pain: a randomized controlled trial. J Manipulative Physiol Ther 2011;34(4):211-220.
52. Deepa A, Dabholkar T, Y., Yardi S. Comparison of the efficacy of Maitland Thoracic Mobilization and Deep Neck Flexor Endurance Training Versus Only Deep Neck Flexor Endurance Training in Patients with Mechanical Neck Pain. Indian J Physiother Occup Ther 2014;8(3):77-82
53. Dunning JR, Cleland JA, Waldrop MA, Arnot CF, Young IA, Turner M, et al. Upper cervical and upper thoracic thrust manipulation versus nonthrust mobilization in patients with mechanical neck pain: a multicenter randomized clinical trial. J Orthop Sports Phys Ther 2012;42(1):5-18.
54. Dusunceli Y, Ozturk C, Atamaz F, Hepguler S, Durmaz B. Efficacy of neck stabilization exercis-

es for neck pain: a randomized controlled study. J Rehabil Med 2009;41(8):626-631.
55. Dziedzic K, Hill J, Lewis M, Sim J, Daniels J, Hay EM. Effectiveness of manual therapy or pulsed shortwave diathermy in addition to advice and exercise for neck disorders: a pragmatic randomized controlled trial in physical therapy clinics. Arthritis Rheum 2005;53(2):214-222.
56. Escortell-Mayor E, Riesgo-Fuertes R, Garrido-Elustondo S, Asunsolo-Del Barco A, Diaz-Pulido B, Blanco-Diaz M, et al. Primary care randomized clinical trial: manual therapy effectiveness in comparison with TENS in patients with neck pain. Man Ther 2011;16(1):66-73.
57. Evans R, Bronfort G, Schulz C, Maiers M, Bracha Y, Svendsen K, et al. Supervised exercise with and without spinal manipulation performs similarly and better than home exercise for chronic neck pain: a randomized controlled trial. Spine 2012;37(11):903-914.
58. Falla D, Jull G, Hodges P, Vicenzino B. An endurance-strength training regime is effective in reducing myoelectric manifestations of cervical flexor muscle fatigue in females with chronic neck pain. Clin Neurophysiol 2006;117(4):828-837.
59. 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 2007;87(4):408-417.
60. Falla D, Jull G, Hodges P. Training the cervical muscles with prescribed motor tasks does not change muscle activation during a functional activity. Man Ther 2008;13(6):507-512.
61. Falla D, Lindstrom R, Rechter L, Boudreau S, Petzke F. Effectiveness of an 8-week exercise programme on pain and specificity of neck muscle activity in patients with chronic neck pain: a randomized controlled study. European Journal of Pain 2013;17(10):1517-1528.
62. Gallego Izquierdo T, Pecos-Martin D, Lluch Girbes E, Plaza-Manzano G, Rodriguez Caldentey R, Mayor Melus R, et al. Comparison of cranio-cervical flexion training versus cervical proprioception training in patients with chronic neck pain: A randomized controlled clinical trial. J Rehabil Med 2016;48(1):48-55.
63. Ganesh GS, Mohanty P, Pattnaik M, Mishra C. Effectiveness of mobilization therapy and exercises in mechanical neck pain. Physiother Theory Pract 2014:1-8.
64. Giombini A, Di Cesare A, Quaranta F, Giannini S, Di Cagno A, Mazzola C, et al. Neck balance system in the treatment of chronic mechanical neck pain: a prospective randomized controlled study. Eur J Phys Rehabil Med 2013;49(3):283-290.
65. Gonzalez-Iglesias J, Fernandez-de-las-Penas C, Cleland JA, Gutierrez-Vega Mdel R. Thoracic spine manipulation for the management of patients with neck pain: a randomized clinical trial. J Orthop Sports Phys Ther 2009;39(1):20-27.
66. Gonzalez-Iglesias J, Fernandez-de-las-Penas C, Cleland JA, Alburquerque-Sendin F, Palomeque-del-Cerro L, Mendez-Sanchez R. Inclusion of thoracic spine thrust manipulation into an electro-therapy/thermal program for the management of patients with acute mechanical neck pain: a randomized clinical trial. Man Ther 2009;14(3):306-313.
67. Griffiths C, Dziedzic K, Waterfield J, Sim J. Effectiveness of specific neck stabilization exercises

or a general neck exercise program for chronic neck disorders: a randomized controlled trial. J Rheumatol 2009;36(2):390-397.

68. Griswold D, Learman K, O'Halloran B, Cleland J. A preliminary study comparing the use of cervical/upper thoracic mobilization and manipulation for individuals with mechanical neck pain. J Man Manip Ther 2015;23(2):75-83.
69. Haas M, Groupp E, Panzer D, Partna L, Lumsden S, Aickin M. Efficacy of cervical endplay assessment as an indicator for spinal manipulation. Spine 2003;28(11):1091-6; discussion 1096.
70. Hakkinen A, Salo P, Tarvainen U, Wiren K, Ylinen J. Effect of manual therapy and stretching on neck muscle strength and mobility in chronic neck pain. J Rehabil Med 2007;39(7):575-579.
71. Hakkinen A, Kautiainen H, Hannonen P, Ylinen J. Strength training and stretching versus stretching only in the treatment of patients with chronic neck pain: a randomized one-year follow-up study. Clin Rehabil 2008;22(7):592-600.
72. Helewa A, Goldsmith CH, Smythe HA, Lee P, Obright K, Stitt L. Effect of therapeutic exercise and sleeping neck support on patients with chronic neck pain: a randomized clinical trial. J Rheumatol 2007;34(1):151-158.
73. Hoving JL, Koes BW, de Vet HC, van der Windt DA, Assendelft WJ, van Mameren H, et al. Manual therapy, physical therapy, or continued care by a general practitioner for patients with neck pain. A randomized, controlled trial. Ann Intern Med 2002;136(10):713-722.
74. Hudson JS, Ryan CG. Multimodal group rehabilitation compared to usual care for patients with chronic neck pain: a pilot study. Man Ther 2010;15(6):552-556.
75. Humphreys BK, Irgens PM. The effect of a rehabilitation exercise program on head repositioning accuracy and reported levels of pain in chronic neck pain subjects. Journal of Whiplash & Related Disorders 2002;1(1):99-112.
76. Izquierdo Perez H, Alonso Perez JL, Gil Martinez A, La Touche R, Lerma-Lara S, Commeaux Gonzalez N, et al. Is one better than another?: A randomized clinical trial of manual therapy for patients with chronic neck pain. Man Ther 2014;19(3):215-221.
77. Javanshir K, Amiri M, Mohseni Bandpei MA, De las Penas CF, Rezasoltani A. The effect of different exercise programs on cervical flexor muscles dimensions in patients with chronic neck pain. J Back Musculoskelet Rehabil 2015;28(4):833-840.
78. Jordan A, Bendix T, Nielsen H, Hansen FR, Host D, Winkel A. Intensive training, physiotherapy, or manipulation for patients with chronic neck pain. A prospective, single-blinded, randomized clinical trial. Spine 1998;23(3):311-8; discussion 319.
79. Jull G, Falla D, Treleaven J, Hodges P, Vicenzino B. Retraining cervical joint position sense: the effect of two exercise regimes. J Orthop Res 2007;25(3):404-412.
80. Jull GA, Falla D, Vicenzino B, Hodges PW. The effect of therapeutic exercise on activation of the deep cervical flexor muscles in people with chronic neck pain. Man Ther 2009;14(6):696-701.

81. Kanlayanaphotporn R, Chiradejnant A, Vachalathiti R. The immediate effects of mobilization technique on pain and range of motion in patients presenting with unilateral neck pain: a randomized controlled trial. Arch Phys Med Rehabil 2009;90(2):187-192.
82. Kanlayanaphotporn R, Chiradejnant A, Vachalathiti R. Immediate effects of the central posteroanterior mobilization technique on pain and range of motion in patients with mechanical neck pain. Disabil Rehabil 2010;32(8):622-628.
83. Karlsson L, Takala EP, Gerdle B, Larsson B. Evaluation of pain and function after two home exercise programs in a clinical trial on women with chronic neck pain - with special emphasises on completers and responders. BMC Musculoskelet Disord 2014;15:6-2474-15-6.
84. Khan M, Soomro RR, Ali SS. The effectiveness of isometric exercises as compared to general exercises in the management of chronic non-specific neck pain. Pakistan Journal of Pharmaceutical Sciences 2014;27(5 Suppl):1719-1722.
85. Kim JH, Lee HS, Park SW. Effects of the active release technique on pain and range of motion of patients with chronic neck pain. J Phys Ther Sci 2015;27(8):2461-2464.
86. Kim JY, Kwag KI. Clinical effects of deep cervical flexor muscle activation in patients with chronic neck pain. J Phys Ther Sci 2016;28(1):269-273.
87. Kjellman G, Oberg B. A randomized clinical trial comparing general exercise, McKenzie treatment and a control group in patients with neck pain. J Rehabil Med 2002;34(4):183-190.
88. Klaber Moffett JA, Jackson DA, Richmond S, Hahn S, Coulton S, Farrin A, et al. Randomised trial of a brief physiotherapy intervention compared with usual physiotherapy for neck pain patients: outcomes and patients' preference. BMJ 2005;330(7482):75.
89. Ko T, Jeong U, Lee K. Effects of the inclusion thoracic mobilization into cranio-cervical flexor exercise in patients with chronic neck pain. J Phys Ther Sci2010;22(1):87-91.
90. Krauss J, Creighton D, Ely JD, Podlewska-Ely J. The immediate effects of upper thoracic translatoric spinal manipulation on cervical pain and range of motion: a randomized clinical trial. J Man Manipul Ther 2008;16(2):93-99.
91. Kumar D, Sandhu JS, Broota A. Efficacy of Mulligan concept (NAGs) on pain at available end range in cervical spine: a randomised controlled trial. Indian J Physiother Occup Ther 2011;5(1):154-158.
92. Lansinger B, Larsson E, Persson LC, Carlsson JY. Qigong and exercise therapy in patients with long-term neck pain: a prospective randomized trial. Spine 2007;32(22):2415-2422.
93. Lansinger B, Carlsson JY, Kreuter M, Taft C. Health-related quality of life in persons with long-term neck pain after treatment with Qigong and exercise therapy respectively. European Journal of Physiotherapy 2013;15(3):111-117.
94. Lau HM, Wing Chiu TT, Lam TH. The effectiveness of thoracic manipulation on patients with chronic mechanical neck pain - a randomized controlled trial. Man Ther 2011;16(2):141-147.
95. Leaver AM, Maher CG, Herbert RD, Latimer J, McAuley JH, Jull G, et al. A randomized con-

trolled trial comparing manipulation with mobilization for recent onset neck pain. Arch Phys Med Rehabil 2010;91(9):1313-1318.

96. Lee J, Lee Y, Kim H, Lee J. The effects of cervical mobilization combined with thoracic mobilization on forward head posture of neck pain patients. Journal of Physical Therapy Science 2013;25(1):7-9.
97. Lee KW, Kim WH. Effect of thoracic manipulation and deep craniocervical flexor training on pain, mobility, strength, and disability of the neck of patients with chronic nonspecific neck pain: a randomized clinical trial. J Phys Ther Sci 2016;28(1):175-180.
98. Lee MY, Kim SG, Lee HY. The effect of cervical stabilization exercise on active joint position sense: A randomized controlled trial. J Back Musculoskelet Rehabil 2016;29(1):85-88.
99. Lluch E, Arguisuelas MD, Calvente Quesada O, Martinez Noguera E, Peiro Puchades M, Perez Rodriguez JA, et al. Immediate effects of active versus passive scapular correction on pain and pressure pain threshold in patients with chronic neck pain. J Manipulative Physiol Ther 2014;37(9):660-666.
100. Lluch E, Schomacher J, Gizzi L, Petzke F, Seegar D, Falla D. Immediate effects of active cranio-cervical flexion exercise versus passive mobilisation of the upper cervical spine on pain and performance on the cranio-cervical flexion test. Man Ther 2014;19(1):25-31.
101. Lopez-Lopez A, Alonso Perez JL, Gonzalez Gutierez JL, La Touche R, Lerma Lara S, Izquierdo H, et al. Mobilization versus manipulations versus sustain appophyseal natural glide techniques and interaction with psychological factors for patients with chronic neck pain: Randomized control Trial. Eur J Phys Rehabil Med 2014;51(2):121-132
102. Maayah M, Al-Jarrah M. Evaluation of Transcutaneous Electrical Nerve Stimulation as a Treatment of Neck Pain due to Musculoskeletal Disorders. J Clin Med Res 2010;2(3):127-136.
103. Madson TJ, Cieslak KR, Gay RE. Joint mobilization vs massage for chronic mechanical neck pain: a pilot study to assess recruitment strategies and estimate outcome measure variability. J Manipulative Physiol Ther 2010;33(9):644-651.
104. Mansilla-Ferragut P, Fernandez-de-Las Penas C, Alburquerque-Sendin F, Cleland JA, Bosca-Gandia JJ. Immediate effects of atlanto-occipital joint manipulation on active mouth opening and pressure pain sensitivity in women with mechanical neck pain. J Manipulative Physiol Ther 2009;32(2):101-106.
105. Martel J, Dugas C, Dubois JD, Descarreaux M. A randomised controlled trial of preventive spinal manipulation with and without a home exercise program for patients with chronic neck pain. BMC Musculoskelet Disord 2011;12:41-2474-12-41.
106. Martinez-Segura R, Fernandez-de-las-Penas C, Ruiz-Saez M, Lopez-Jimenez C, Rodriguez-Blanco C. Immediate effects on neck pain and active range of motion after a single cervical high-velocity low-amplitude manipulation in subjects presenting with mechanical neck pain: a randomized controlled trial. J Manipulative Physiol Ther 2006;29(7):511-517.

107. Marti¬nez-Segura R, Isabel De-La-Llave-Rincon A, Ortega-Santiago R, Cleland J, A., Fernandez-De-Las-Penas C. Immediate Changes in Widespread Pressure Pain Sensitivity, Neck Pain, and Cervical Range of Motion After Cervical or Thoracic Thrust Manipulation in Patients With Bilateral Chronic Mechanical Neck Pain : A Randomized Clinical Trial. J Orthop Sports Phys Ther 2012;42(9):806-814.
108. Masaracchio M, Cleland JA, Hellman M, Hagins M. Short-term combined effects of thoracic spine thrust manipulation and cervical spine nonthrust manipulation in individuals with mechanical neck pain: a randomized clinical trial. J Orthop Sports Phys Ther 2013;43(3):118-127.
109. McLean SM, Klaber Moffett JA, Sharp DM, Gardiner E. A randomised controlled trial comparing graded exercise treatment and usual physiotherapy for patients with non-specific neck pain (the GET UP neck pain trial). Man Ther 2013;18(3):199-205.
110. Monticone M, Baiardi P, Vanti C, Ferrari S, Nava T, Montironi C, et al. Chronic neck pain and treatment of cognitive and behavioural factors: results of a randomised controlled clinical trial. Eur Spine J 2012;21(8):1558-1566.
111. O'Leary S, Jull G, Kim M, Vicenzino B. Specificity in retraining craniocervical flexor muscle performance. J Orthop Sports Phys Ther 2007;37(1):3-9.
112. O'Leary S, Falla D, Hodges PW, Jull G, Vicenzino B. Specific therapeutic exercise of the neck induces immediate local hypoalgesia. J Pain 2007;8(11):832-839.
113. O'Leary S, Jull G, Kim M, Uthaikhup S, Vicenzino B. Training mode-dependent changes in motor performance in neck pain. Arch Phys Med Rehabil 2012;93(7):1225-1233.
114. Paoloni M, Tavernese E, Cacchio A, Tattoli M, Melis L, Ronconi R, et al. Patient-oriented rehabilitation in the management of chronic mechanical neck pain: a randomized controlled trial. Eur J Phys Rehabil Med 2013;49(3):273-281.
115. Pillastrini P, de Lima E Sa Resende,F., Banchelli F, Burioli A, Di Ciaccio E, Guccione AA, et al. Effectiveness of Global Postural Re-education in Patients With Chronic Nonspecific Neck Pain: Randomized Controlled Trial. Phys Ther 2016;96(9):1408-1416.
116. Pires PF, Packer AC, Dibai-Filho AV, Rodrigues-Bigaton D. Immediate and Short-Term Effects of Upper Thoracic Manipulation on Myoelectric Activity of Sternocleidomastoid Muscles in Young Women With Chronic Neck Pain: A Randomized Blind Clinical Trial. J Manipulative Physiol Ther 2015;38(8):555-563.
117. Pool JJ, Ostelo RW, Knol DL, Vlaeyen JW, Bouter LM, de Vet HC. Is a behavioral graded activity program more effective than manual therapy in patients with subacute neck pain? Results of a randomized clinical trial. Spine 2010 1;35(10):1017-1024.
118. Puentedura EJ, Landers MR, Cleland JA, Mintken PE, Huijbregts P, Fernandez-de-Las-Penas C. Thoracic spine thrust manipulation versus cervical spine thrust manipulation in patients with acute neck pain: a randomized clinical trial. J Orthop Sports Phys Ther 2011;41(4):208-220.

119. Puntumetakul R, Suvarnnato T, Werasirirat P, Uthaikhup S, Yamauchi J, Boucaut R. Acute effects of single and multiple level thoracic manipulations on chronic mechanical neck pain: a randomized controlled trial. Neuropsychiatr Dis Treat 2015;11:137-144.
120. Rendant D, Pach D, LÃ¼dtke R, Reisshauer A, Mietzner A, Willich SN, et al. Qigong versus exercise versus no therapy for patients with chronic neck pain: a randomized controlled trial. Spine 2011;36(6):419-427.
121. Revel M, Minguet M, Gregoy P, Vaillant J, Manuel JL. Changes in cervicocephalic kinesthesia after a proprioceptive rehabilitation program in patients with neck pain: a randomized controlled study. Arch Phys Med Rehabil 1994;75(8):895-899.
122. Rolving N, Christiansen DH, Andersen LL, Skotte J, Ylinen J, Jensen OK, et al. Effect of strength training in addition to general exercise in the rehabilitation of patients with non-specific neck pain. A randomized clinical trial. Eur J Phys Rehabil Med 2014;50(6):617-626.
123. Rudolfsson T, Djupsjobacka M, Hager C, Bjorklund M. Effects of neck coordination exercise on sensorimotor function in chronic neck pain: a randomized controlled trial. J Rehabil Med 2014;46(9):908-914.
124. Saavedra-Hernandez M, Castro-Sanchez AM, Arroyo-Morales M, Cleland JA, Lara-Palomo IC, Fernandez-de-Las-Penas C. Short-term effects of kinesio taping versus cervical thrust manipulation in patients with mechanical neck pain: a randomized clinical trial. J Orthop Sports Phys Ther 2012;42(8):724-730.
125. Saavedra-Hernandez M, Arroyo-Morales M, Cantarero-Villanueva I, Fernandez-Lao C, Castro-Sanchez A, M, Puentedura E, J., et al. Short-term effects of spinal thrust joint manipulation in patients with chronic neck pain: a randomized clinical trial. Clin Rehabil 2013;27(6):504-512.
126. Saayman L, Hay C, Abrahamse H. Chiropractic manipulative therapy and low-level laser therapy in the management of cervical facet dysfunction: a randomized controlled study. J Manipulative Physiol Ther 2011;34(3):153-163.
127. Salom-Moreno J, Ortega-Santiago R, Cleland J, Aland, Palacios-CeÃƒÂ±a M, Truyols-DomÃƒÂ¬nguez S, FernÃƒÂ¡ndez-de-las-PeÃƒÂ±as C. Immediate Changes in Neck Pain Intensity and Widespread Pressure Pain Sensitivity in Patients With Bilateral Chronic Mechanical Neck Pain: A Randomized Controlled Trial of Thoracic Thrust Manipulation vs Nonâ€“Thrust Mobilization. J Manipulative Physiol Ther 2014;37(5):312-319.
128. Sarig Bahat H, Takasaki H, Chen X, Bet-Or Y, Treleaven J. Cervical kinematic training with and without interactive VR training for chronic neck pain - a randomized clinical trial. Man Ther 2015;20(1):68-78.
129. Schomacher J. The effect of an analgesic mobilization technique when applied at symptomatic or asymptomatic levels of the cervical spine in subjects with neck pain: a randomized controlled trial. J Man Manip Ther 2009;17(2):101-108.

130. Sherman KJ, Cook AJ, Wellman RD, Hawkes RJ, Kahn JR, Deyo RA, et al. Five-week outcomes from a dosing trial of therapeutic massage for chronic neck pain. Ann Fam Med 2014;12(2):112-120.
131. Sillevis R, Cleland J, Hellman M, Beekhuizen K. Immediate effects of a thoracic spine thrust manipulation on the autonomic nervous system: a randomized clinical trial. J Man Manip Ther 2010;18(4):181-190.
132. Snodgrass S, J., Rivett D, A., Sterling M, Vicenzino B. Dose Optimization for Spinal Treatment Effectiveness: A Randomized Controlled Trial Investigating the Effects of High and Low Mobilization Forces in Patients With Neck Pain. J Orthop Sports Phys Ther 2014;44(3):141-152.
133. Sterling M, Jull G, Wright A. Cervical mobilisation: concurrent effects on pain, sympathetic nervous system activity and motor activity. Man Ther 2001;6(2):72-81.
134. Taimela S, Takala EP, Asklof T, Seppala K, Parviainen S. Active treatment of chronic neck pain: a prospective randomized intervention. Spine (Phila Pa 1976) 2000;25(8):1021-1027.
135. Thompson DP, Oldham JA, Woby SR. Does adding cognitive-behavioural physiotherapy to exercise improve outcome in patients with chronic neck pain? A randomised controlled trial. Physiotherapy 2016;102(2):170-177.
136. Vernon HT, Aker P, Burns S, Viljakaanen S, Short L. Pressure pain threshold evaluation of the effect of spinal manipulation in the treatment of chronic neck pain: a pilot study. J Manipulative Physiol Ther 1990;13(1):13-16.
137. Viljanen M, Malmivaara A, Uitti J, Rinne M, Palmroos P, Laippala P. Effectiveness of dynamic muscle training, relaxation training, or ordinary activity for chronic neck pain: randomised controlled trial. BMJ 2003;327(7413):475.
138. von Trott P, Wiedemann AM, Ludtke R, Reishauer A, Willich SN, Witt CM. Qigong and exercise therapy for elderly patients with chronic neck pain (QIBANE): a randomized controlled study. J Pain 2009;10(5):501-508.
139. Vonk F, Verhagen AP, Twisk JW, Koke AJ, Luiten MW, Koes BW. Effectiveness of a behaviour graded activity program versus conventional exercise for chronic neck pain patients. Eur J Pain 2009;13(5):533-541.
140. Walker MJ, Boyles RE, Young BA, Strunce JB, Garber MB, Whitman JM, et al. The effectiveness of manual physical therapy and exercise for mechanical neck pain: a randomized clinical trial. Spine 2008;33(22):2371-2378.
141. Yang J, Lee B, Kim C. Changes in proprioception and pain in patients with neck pain after upper thoracic manipulation. J Phys Ther Sci 2015;27(3):795-798.
142. Ylinen J, Takala E, Nykänen M, Häkkinen A, Mälkiä E, Pohjolainen T, et al. Active neck muscle training in the treatment of chronic neck pain in women: a randomized controlled trial. JAMA 2003;289(19):2509-2516.
143. Ylinen J, Takala EP, Kautiainen H, Nykanen M, Hakkinen A, Pohjolainen T, et al. Effect of long-

term neck muscle training on pressure pain threshold: a randomized controlled trial. Eur J Pain 2005;9(6):673-681.

144. Ylinen J, Hakkinen A, Nykanen M, Kautiainen H, Takala EP. Neck muscle training in the treatment of chronic neck pain: a three-year follow-up study. Eura Medicophys 2007;43(2):161-169.
145. Ylinen J, Kautiainen H, Wiren K, Hakkinen A. Stretching exercises vs manual therapy in treatment of chronic neck pain: a randomized, controlled cross-over trial. J Rehabil Med 2007;39(2):126-132.
146. Zaproudina N, Hanninen OO, Airaksinen O. Effectiveness of traditional bone setting in chronic neck pain: randomized clinical trial. J Manipulative Physiol Ther 2007;30(6):432-437.
147. Swinkels RA, Swinkels-Meewisse IE. Normal values for cervical range of motion. Spine 2014;39(5):362-367.
148. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687.
149. Groeneweg R, Rubinstein SM, Oostendorp RA, Ostelo RW, van Tulder MW. Guideline for Reporting Interventions on Spinal Manipulative Therapy: Consensus on Interventions Reporting Criteria List for Spinal Manipulative Therapy (CIRCLe SMT). J Manipulative Physiol Ther 2017;40(2):61-70.
150. Kjellman GV, Skargren EI, Oberg BE. A critical analysis of randomised clinical trials on neck pain and treatment efficacy. A review of the literature. Scand J Rehabil Med 1999;31(3):139-152.
151. Elven M, Hochwalder J, Dean E, Soderlund A. A clinical reasoning model focused on clients' behaviour change with reference to physiotherapists: its multiphase development and validation. Physiother Theory Pract 2015;31(4):231-243.
152. Jones LE, O'Shaughnessy DF. The pain and movement reasoning model: introduction to a simple tool for integrated pain assessment. Man Ther 2014;19(3):270-276.
153. Jones MA. Clinical reasoning in manual therapy. Phys Ther 1992;72(12):875-884.
154. Hoogeboom TJ, Oosting E, Vriezekolk JE, Veenhof C, Siemonsma PC, de Bie RA, et al. Therapeutic validity and effectiveness of preoperative exercise on functional recovery after joint replacement: a systematic review and meta-analysis. PLoS One 2012;7(5):e38031.
155. Childs JD, Fritz JM, Piva SR, Whitman JM. Proposal of a classification system for patients with neck pain. J Orthop Sports Phys Ther 2004;34(11):686-96; discussion 697-700.
156. Clair DA, Edmondston SJ, Allison GT. Physical therapy treatment dose for nontraumatic neck pain: a comparison between 2 patient groups. J Orthop Sports Phys Ther 2006;36(11): 867- 875.
157. Liu R, Kurihara C, Tsai HT, Silvestri PJ, Bennett MI, Pasquina PF, et al. Classification and

Treatment of Chronic Neck Pain: A Longitudinal Cohort Study. Reg Anesth Pain Med 2017;42(1): 52-61.

158. Wang WT, Olson SL, Campbell AH, Hanten WP, Gleeson PB. Effectiveness of physical therapy for patients with neck pain: an individualized approach using a clinical decision-making algorithm. Am J Phys Med Rehabil 2003;82(3):203-18; quiz 219-21.

Appendix 1: *Medline search*

((("Neck Pain"[Mesh] OR "neck pain"[tiab] OR neckache*[tiab] OR "neck ache"[-tiab] OR "neck aches"[tiab] OR cervicodynia*[tiab] OR cervicalgia*[tiab]) AND ("Physical therapy Modalities"[Mesh] OR "physical therapy"[tiab] OR "physical therapy"[tiab] OR "physical therapies"[tiab] OR "manual therapy"[tiab] OR "manual therapies"[tiab])) OR ("Neck Pain/rehabilitation"[Mesh] OR "Neck Pain/therapy-"[Mesh] OR "Neck Pain/prevention and control"[Mesh])) AND (randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab]) NOT (animals[mh] NOT (animals[mh] AND humans[mh]))

Appendix 2: *Patient and study characteristics*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
Akther [25]	62	From a institute of physical medicine and rehabilitation	63	39	7.5	4,5 months	Exp gr = HVT + exercises Interv gr = exercises	5
Ali [26]	102	From an out-patient department of Physical ther-apy	Not reported	Not reported	7.5	Not reported	Exp gr = SNAG + Isometric exercises Interv gr = SNAG	4
Andrade Ortega [27]		From a primary care setting	76	44	5.4	> 3 months	Exp gr1 = continuous microwave + TENS + exercises Exp gr2 = pulsed microwave +TENS + exercises Plac gr = unplugged microwave + TENS + exercises	8
Aquino [28]	48	University teaching clinic + private clinics	17.5	34.1	6.2	> 3 months	Exp gr = mobilizations over a symptomatic level Plac gr = mobilizations applied over a randomly chosen level	8
Bakar [29]	45	From a hospital neurosurgery polyclinic	100	40	Not reported	3-6 months	Exp gr = Classic massage Interv gr = Connective tissue massage	6
Beer [30]	20	Advertisement	50	29	2.8	6 years	Exp gr = Postural exercise intervention Contr gr = No treatment	5
Beinert [31]	34	Not reported	Not reported	23	4.6	27 months	Exp gr = Balance training Contr gr = no training, remain active	6
Beltran-Alacreu [32]	45	By referral from hospital	11	42	Not reported	90 months	Exp 1 gr = SMT + patient education Exp 2 gr = SMT + patient education + exercise Interv gr = SMT	6
Bid [33]	40	Not reported	53	42	8.2	Not reported	Exp gr = conventional therapy + Cervical traction Interv gr = conventional therapy	5

Appendix 2: *Patient and study characteristics (continued)*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
Borisut [34]	100	Not reported	100	30.7	5.8	> 6 months	Exp gr1 = strength exercises Exp gr2 = cranio cervical flexion exercises Exp gr3 = combined exercises Cont gr = no exercises	5
Borman [35]	42	Outpatient service of physical medicine and rehabilitation department	65.2	49.3	5.4	4.1 year	Exp gr = PT with traction Interv gr = Pt without traction	5
Brage [36]	20	By local advertisements	100	41	5.5	172 months	Exp gr = specific neck/shoulder and aerobic training + pain education Interv gr = pain education	7
Briem [37]	40	By referral	78	35	5	Not reported	Exp gr = Inhibitive distraction Plac gr = sham distraction	5
Casanova-Mendez [38]	64	Database of the principal researcher's clinic	72	37	2.4	> 3 months	Exp gr = HVT Dog technique Interv gr = HVT Toggle Recoil	8
Celenay [39]	60	By referral by their physician	65	48	4	> 3 months	Exp gr = Cervical and scapulothoracic stabilization exercise + connective tissue massage Interv gr = Cervical and scapulothoracic stabilization exercise	6
Celenay [40]	102	By referral by their physician	62	45	3.9	> 3 months	Exp gr = Cervical and scapulothoracic stabilization exercise + MT Interv gr = Cervical and scapulothoracic stabilization exercise	7
Chui [41]	145	Physical therapy outpatient department	69.9	43.8	4.5	18 % 3-6 month; 21 % 6-12 month; 61 % > 1year	Exp gr = Exercise + infrared irradiation Plac gr = Non exercise + infrared irradiation	7

Chiu [42]	218	From the physical therapy outpatient departments of two hospitals	68	44	4.5	3-6 months 19%, 6-12 months 20%, > 12 months 61%	Exp gr1 = TENS + infrared irradiation Exp gr2 = exercises + infrared Irradiation Plac gr = Infrared irradiation	6
Chui [43]	79	Physical therapy outpatient department	70.8	48.9	5.5	> 3 months	Exp gr = Traction + infrared irradiation Plac gr = non traction + infrared irradiation	7
Cleland [44]	52	Outpatient orthopedic physical therapy clinic	75	35.5	4.2	Not reported	Exp gr = SMT Plac gr = placebo SMT	8
Cleland [45]	60	Outpatient orthopedic physical therapy clinic	55	43.3	4.9	57.6 days	Exp gr = SMT Interv gr = mobilization	7
Cleland [46]	140	Physical therapy clinics	69	39.9	4.2	63.5 days	Exp gr1 = SMT+ exercise Interv gr1= Exercise only Exp gr2 = + CPR Interv gr2= - CPR Exp gr3 = SMT+ exercise + CPR Interv gr3= SMT+ exercise – CPR	7
Cook [47]	179	From an integrated healthcare system and from a general population	73	47	5.8	41% > 5 years	Exp gr = massage + booster treatment Interv gr = Massage	7
Cunha [48]	33	Outpatient orthopedics and rheumatology clinic	100	46.6	6,6	> 3 months	Exp gr = Global posture reeducation Interv gr = stretching group	10
Dawood [50]	54	referred from orthopedic consultants	28	22-36	7.1	> 3 months	Exp gr 1 = Kinesio tape + exercises Exp gr 2 = cervical traction posture pump + exercises Interv gr = exercises	5

Appendix 2: *Patient and study characteristics (continued)*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
de Camargo [51]	37	Workers of the University of Campinas	43.2	30.1	Not reported	Not reported	Exp gr = SMT (C5-6) Cont gr = no therapy in a similar tim	5
Deepa [52]	30	Not reported	63	37	4.6	Not reported	Exp gr = Thoracic mobilization + Deep neck flexor endurance training Interv gr = Deep neck flexor endurance training	4
Dunning [53]	107	Outpatient physical therapy clinic	68.5	41.8	5.3	1 year	Exp gr = SMT (C1-2, T1-2) Interv gr = mobilization (C1-2, T1-2)	8
Dusunceli [54]	60	Not reported	66.7	52.0	6.7	4.2 year	Exp gr1 = PT + isometric and stretching Exp gr2 = PT + stabilization exercises Interv gr = PT	6
Dziedzic [55]	350	Outpatient physical therapy facilities	63.3	51.2	5.0	77.3 % > 3 month	Exp gr1 = addition of manual therapy Exp gr2 = addition of PSWD Interv gr = advise and exercise	8
Escortell-Mayor [56]	90	Primary health care physical therapy units	88.9	40.1	5.6	147 days	Exp gr = SMT Intev gr = TEN	8
Evans [57]	270	Newspaper advertisement, community posters, mass mailing	72.2	46.3	5.6	9.4 year	Exp gr1 = exercise Exp gr2 = with addition of SMT Interv gr = Home exercise and advice	8
Falla [58]	58	Newspaper advertisement	100	37.9	4.1	7.9 year	Exp gr = endurance-strength training Interv gr = low load CCF training	8
Falla [59]	58	Newspaper advertisement	100	37.9	4.1	7.9 year	Exp gr = endurance-strength training Interv gr = low load CCF training	7

Falla [60]	58	Newspaper advertisement	100	37.9	4.1	7.9 year	Exp gr = endurance-strength training Interv gr = low load CCF training	6
Falla [61]	46	Pain management center, general practitioners, newspaper advertisement	100	38.9	5.2	9.2 year	Exp gr = progressive exercise programme Cont gr = no exercises	8
Gallego Izquierdo [62]	28	By advertisements	64	29	3.7	> 3 months	Exp gr = CCF exercises Interv gr = proprioception exercises	8
Ganesh [63]	60	Printed advertisement in the institution	37	42	6.1	> 3 months	Exp gr1 = Maitland mob + exercises Exp gr2 = Mulligan mob + exercises Interv gr = Exercises	6
Giombini [64]	45	Outpatient clinics	68.9	42.5	6.1	3 months - 2 years	Exp gr1 = Neck balance system Exp gr2 = Sham Neck Balance system Plac gr = Pulsed electro-magnetic fields	8
Gonzalez-Iglesias [65]	45	Referred by primary health physician to a physical therapy clinic	46.7	34.5	5.4	19 days	Exp gr = addition of thoracic SMT Interv gr = electro + thermal therapy	7
Gonzalez-Iglesias [66]	45	Referred by primary health physician to a physical therapy clinic	55.6	34	5.5	17.5 days	Exp gr = addition of thoracic SMT Interv gr = electro + thermal therapy	9
Griffiths [67]	71	4 outpatient sites	64.8	51.3	5.2	27 month	Exp gr = stabilization exercise Interv gr= general exercise	8
Griswold [68]	20	From 2 clinic sites and a university through direct solicitation, direct access, and physician referral	55	39	5.4	Not reported	Exp gr = Cervical and thoracic SMT Interv gr = Cervical and thoracic mobilization	6

Appendix 2: *Patient and study characteristics (continued)*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
Haas [69]	104	Therapist referral or advertisement	63	42.6	4.1	57 % > 3month	Exp gr = SMT according to endplay findings Plac gr = SMT according to sham findings	8
Häkkinen [70]	125	Employers of the town	100	42.5	2.3	3.4 year	Exp gr = MT followed by stretching Interv gr = Stretching followed by MT	6
Häkkinen [71]	101	Occupational health care centers	45.5	40.5	6.2	5.6 year	Exp gr = additional strength training Interv gr = stretching	7
Helewa [72]	128	Not reported	62.7	52.4	2.8	2-12 months	Exp gr1 = addition support pillow Exp gr2 = addition exercises Exp gr3 = addition pillow + exercises Plac gr = hot or cold packs + massage	6
Hoving [73]	183	Referred by general practitioners	60.8	45.5	7.9	2-6 week 48%, 7-12 week 26%, ≥ 13 week 26%	Exp gr1 = MT mobilization Exp gr2 = PT Interv gr = GP	8
Hudson [74]	12	From the Physical therapy Department at Hospital	77	42	6.5	6.4 months	Exp gr = Multimodal group rehabilitation Interv gr = Usual care physical therapy	5
Humphreys [75]	63	Not reported	50	22.6	5.2	19 month	Exp gr1 = symptomatic non exercise Exp gr2 = symptomatic exercise Exp gr3 = asymptomatic exercise Cont gr = asymptomatic non-exercise	5
Izquierdo Perez [76]	69	Referred by primary care physician	57	37	3.0	90 months	Exp 1 gr = SNAG Exp 2 gr = HVT Interv gr = mobilization	9
Javanshir [77]	60	Referred to the physical therapy clinic	66	38	5	3.3 years	Exp gr1 = cranio cervical flexion exercises Interv gr = cervical flexion exercises	7

Jordan [78]	119	Referral by general practitioners and from a large public works company	73.9	36	4	3month-3 year 35%, 3-5 year 34%, 5-10 year 24%, >10 year 7%	Exp gr1 = intensive training Exp gr2 = PT Exp gr3 = SMT Cont gr = Neck school	5
Jull [79]	64	Advertisement in a local newspaper	100	40.9	6.5	9.7 year	Exp gr = craniocervical flexion training Interv gr = proprioceptive training	6
Jull [80]	46	Advertisement in a local newspaper	100	38.4	4.4	9.7 year	Exp gr = craniocervical flexion training Interv gr = high load strength training	7
Kanlayanaphotporn [81]	60	Outpatient physical therapy clinic	60	42.3	6.1	902 days	Exp gr = preferred mobilization Plac gr = random mobilization	7
Kanlayanaphotporn [82]	60	Physical therapy clinic	70	42.2	6	1694 days	Exp gr = specific mobilization Plac gr = random mobilization	7
Karlsson [83]	57	Through advertisement in local papers	100	44	5.5	8 years	Exp gr = Strength training Interv gr = Stretching	5
Khan [84]	68	From a physical medicine and rehabilitation university clinic	60	34	8.3	3 months - 3 years	Exp gr = Isometric neck exercises Interv gr = general neck exercises	5
Kim [85]	24	By referral from hospital	Not reported	42	6	> 3 months	Exp gr 1= Active release technics Exp gr 2= Mobilization Cont gr = No treatment	4
Kim [86]	28	By an independent researcher	Not reported	45	5.1	> 3 months	Exp gr = DCF exercises Interv gr = Strength exercises	5

Appendix 2: *Patient and study characteristics (continued)*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
Kjellman [87]	70	3 physical therapy units and a physical therapy practice	75.7	44.9	6.7	≤ 1 week 5%, 1-4 week 33%, ≥1>3 month 19%, ≥3 month 43%	Exp gr1 = McKenzie treatment Exp gr2 = General exercise Plac gr = Ultra sound low dose	6
Klaber-Moffett [88]	268	8 outpatient physical therapy clinics	64	48.3	4.5	<6 month 65.4 %	Exp gr = brief PT Interv gr = Usual PT	6
Ko [89]	53	From a physical therapy training center	100	37	6.5	Not reported	Exp gr = Cranio cervical flexor exercises + Thoracic mobilisation Cont gr = Cranio cervical flexor exercises	5
Krauss [90]	32	3 outpatient physical therapy clinics	81.3	34.6	2.9	Not reported	Exp gr = SMT upper thoracic spine Interv gr = No treatment	6
Kumar [91]	100	Not reported	Not reported	Not reported	Not reported	Not reported	Exp g1r = additional NAG for 1-12 days Exp gr2 = additional NAG for 1-6 days Exp gr3 = additional NAG from 7 till 12 days Interv gr = hotpack, exercise for ROM and strength	2
Lansinger [92]	122	Advertisement in a local newspaper	70	44	5.3	55% 1-5 year, 45% > 5 year	Exp gr = Exercise Interv gr = Qigong	7
Lansinger [93]	139	Advertisement in a local newspaper	70	43.8	5.3	3 month-1 year 17%, > 1 year 38%, >5 year 23 %, >10 year 22%	Exp gr = dynamic exercises Interv gr = Qigong	6

Lau [94]	120	Outpatient clinic	50	44	5.0	Not reported	Exp gr = Infrared irradiation + thoracal HVT + educational materials + exercises Plac gr = Infrared irradiation + educational materials	8
Leaver [95]	182	Private physical therapy, chiropractic, and osteopathy clinics	64.8	38.9	6	19 days	Exp gr = SMT Interv gr = Mobilizations	8
Lee [96]	30	From a hospital	Not reported	Not reported	Not reported	Not reported	Exp gr = Cervical and thoracic Mobilization Interv gr = Cervical mobilization	5
Lee [97]	46	Not reported	50	Not reported	5.2	17 months	Exp gr 1 = Thoracic SMT + deep craniovervical flexor training Exp gr 2 = Deep craniocervical flexor training Interv gr = Active self-exercise	6
Lee [98]	24	Not reported	50	21	Not reported	Not reported	Exp gr = Cervical stabilization exercise Cont gr = No treatment	5
Lluch [99]	23	By notices in different hospitals and universities	87	38.9	6.5	> 3 months	Exp gr = Active scapular correction Interv gr = Passive scapular correction	5
Lluch [100]	18	Pain clinic of a University hospital	Not reported	42	4.5	> 3 months	Exp gr = Active exercises Interv gr = Passive mobilizations	7
Lopez-Lopez [101]	48	From a primary health care center	87	36	2.8	> 3 months	Exp gr1 = HVT Exp gr2 = SNAG Interv gr = Mobilization	8
Maayah [102]	30	Outpatient department of University hospital	50	56	Not reported	1-7 months 3%, 5 years 20%, 48% > 5 years, vague history of pain 29%	Exp gr = TENS Plac gr = TENS stimulator	4
Madson [103]	23	By physicians	70	50	3.4	38 months	Exp gr = Postural education + home exercises + sedative massage Interv gr = Postural education + home exercises + mobilization	6

Appendix 2: *Patient and study characteristics (continued)*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
Mansilla-Ferragut [104]	37	Advertisement in a local newspaper	100	35	Not reported	> 6 months	Exp gr = SMT Plac gr = Sham SMT	4
Martel [105]	98	Radio and printed advertisement	65.8	41.1	3.6	<6 month 1%, 6 month-1 year 6%, 1-2 years 15%, 2-3 year 19%, 3-5 year 17%, 5-10 years 19%, >10 year 23 %	Exp gr1 = SMT Exp gr2 = SMT + home exercise Cont gr = no treatment	6
Martinez-Segura [106]	71	Referred by a primary care physician to a private clinic of physical therapy and osteopathy	63.3	37	5.6	4.3 month	Exp gr = SMT Interv gr = mobilizations	5
Martinez-Segura [107]	90	Referred by a primary care physician to physical therapy	51.1	37	5.6	3.7 year	Exp gr1 = SMT right side Exp gr2 = SMT left side Interv gr = Thoracic trust	7
Masaracchio [108]	66	Presented to physical therapy or volunteers	78	32.5	5	36 days	Exp gr = Cervical spine mobilizations + thoracic spine SMT + home exercises Interv gr = Cervical spine mobilizations + home exercises	7

McLean [109]	151	Waiting list of four secondary care physical therapy departments	59.6	53.8	Not reported	Not reported	Exp gr = Graded exercises Interv gr = PT	6
Monticone [110]	80	Outpatients referred to the physical medicine and rehabilitation unit	75	49	5.2	15 months	Exp gr = PT + cognitive behavioral Therapy Interv gr = PT (mobilization, postural control and strength exercises	7
O'Leary [111]	50	Not reported	100	38.5	1.5	> 3 months	Exp gr = Craniocervical flexion training Interv gr = Conventional cervical exercises	7
O'Leary [112]	48	Not reported	100	41	0.9	> 3 months	Exp gr = Cranio cervical Coordination exercises Contr gr = Cervical flexion endurance exercises	6
O'Leary [113]	60	Recruited from the university and general community	58	37.9	2.4	6.8 year	Exp gr1 = Endurance/strenth training Exp gr2 = Coordination training Interv gr = Mobility training	8
Paoloni [114]	220	Outpatient clinic of a physical medicine and rehabilitation university department	74.5	Not reported	6.9	> 3 months	Exp gr = Patient-oriented therapeutic approach Interv gr = Prescription-oriented therapeutic approach	7
Pillastrini [115]	94	From a University hospital	76	47	4.5	Not reported	Exp gr = Global postural re-education Interv gr = Manual therapy	5
Pires [116]	32	From University	100	25	3.1	> 6 months	Exp gr = Thoracic SMT Plac gr = Sham SMT	7
Pool [117]	146	Not reported	61	45.1	6.1	4-12 weeks	Exp gr = BGA Interv gr = Mobilizations	7
Puentedura [118]	24	Patients who presented to physical therapy	67	33.7	4.1	15.3 days	Exp gr = Cervical thrust + exercise Interv gr = Thoracic thrust + exercise	7

2

Appendix 2: *Patient and study characteristics (continued)*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
Puntumetakul [119]	48	From a physical therapy unit	Not reported	27	5.2	> 3 months	Exp gr 1 = Multi-level thoracic SMT Exp gr 2 = Single level thoracic SMT Contr gr = No therapy	6
Rendant [120]	123	Not reported	88	46	5.6	6 months - 5 years	Exp gr1 = Qigong therapy Exp gr2 = Exercise therapy Contr gr = Waiting list	7
Revel [121]	60	Outpatient department of rheumatology	85	48	4.8	30 month	Exp gr = Rehabilitation therapy +drugs Interv gr = Drugs	4
Rolving [122]	83	From general practitioners	72	39	7	Not reported	Exp gr = General physical activity + specific strength training Interv gr = General physical activity	8
Rudolfsson [123]	108	Through advertisement in the local paper and by invitation s at primary and occupational healthcare units	100	51	5.5 (median)	109 months (median)	Exp 1 gr = Neck coordination exercises Exp 2 gr = Strength training Interv gr = Massage	6
Saavedra-Hernandez [124]	80	Physical therapy private clinic	46.5	45	5.1	79 month	Exp gr = SMT Interv gr = Kinesio tape	8
Saavedra-Hernandez [125]	82	Referred for physical therapy at a private clinic	50	44.5	4.9	80 month	Exp gr = Mid-cervical spine SMT Intervgr = Mid-cervical spine + cervico-thoracic junction + upper thoracic spine SMT	8
Saayman [126]	60	By advertisement	Not reported	29	6	6 months	Exp gr1 = HVT Exp gr2 = Low level laser therapy Exp gr3 = HVT + Low Level Laser Therapy	7

Salom-Moreno [127]	52	Referred by their primary care physician	58	33	5.9	2.3 years	Exp gr = Thoracic HVT Interv gr = Thoracic mob	7
Sarig Bahat [128]	32	Not reported	66%	41	3,5	Not reported	Exp gr = Kinematic training + virtual reality Interv gr = Kinematic training	6
Schomacher [129]	126	Referral or advertisement in 3 private physical therapy practices	Not reported	49.6	3.6	75.4 month	Exp gr = Mobilization symptomatic levels Plac gr = Mobilization a-symptomatic levels	7
Sherman [130]	228	From an integrated health care system and from the general population	72	47	5.8	39% > 5 years	Exp gr = Five groups receiving various doses of massage Contr gr = Waiting list	7
Sillevis [131]	100	5 outpatient physical therapy clinic	77	44.8	3.6	24.3 month	Exp gr = SMT thoracic spine Plac gr = Placebo maneuver	6
Snodgrass [132]	64	Advertisement	75.3	33.5	3.2	3-6 month 8%, 6-12 month 11%, 1-2 year 19%, >2 year 62%	Exp gr1 = High mobilization force Exp gr2 = Low mobilization force Plac gr = Detuned laser	7
Sterling [133]	30	Not reported	54	36	Not reported	> 3 months	Exp gr1 = Passive mobilization Plac gr2 = Placebo condition Contr g = No contact at all	4
Taimela [134]	76	Various workplaces through their respective healthcare systems	Not reported	Not reported	5.1	> 3 months	Exp gr1 = Multimodal active treatment Exp gr2 = Neck lecture home exercise Interv gr = Neck lecture	7
Thompson [135]	56	From outpatient physical therapy departments	45	47	5.6	4.4 years	Exp gr = Progressive neck exercise Program Interv gr = Interactive behavioural modification therapy	5

Appendix 2: *Patient and study characteristics (continued)*

Study	Sample size (Total n)	Recruitment	Female %	Age (mean years)	Pain at baseline (score 0-10) *	Duration of pain (mean; %; >)	Group interventions	PEDro score
Vernon [136]	9	Not reported	33	32	Not reported	Not reported	Exp g = SMT Plac gr = Sham SMT	5
Viljanen [137]	393	Employers through their respective healthcare systems	100	44	4.5	11 year	Exp gr1 = Dynamic muscle training Exp gr2 = Relaxation training Cont gr = Ordinary activity	8
von Trott [138]	117	4 residences for elderly people	95	75.7	5	19.9 year	Exp gr1 = Qigong Exp gr2 = Exercise Cont gr = Waiting list	5
Vonk [139]	139	Referral by general practi-tioner	60	45.7	6.9	57 month	Exp gr = BGA Interv gr = Conventional exercise	5
Walker [140]	94	Referred by physicians	67	47.5	5.3	801 days	Exp gr = SMT+ exercise Plac gr = Minimal intervention	8
Yang [141]	30	Not reported	50	31	7	Not reported	Exp gr = Cervical stability training + upper thoracic MT Interv gr = Cervical stability training	4
Ylinen [142]	180	Referred by physicians	100	46	5.8	8.5 year	Exp gr1 = Strength training Exp gr2 = Endurance training Cont gr = Recreational activities	7
Ylinen [143]	180	Occupational health care services	100	46	5.8	> 6 months	Exp gr1 = Endurance training Exp gr2 = Strength training Interv gr = Stretching exercise	6
Ylinen [144]	177	From their respective occupational health care systems	100	46	5.7(medi-an)	8,5 years	Exp gr = Strength exercises Exp gr = Endurance exercises Cont gr = No therapy	5

Ylinen [145]	125	Occupational healthcare services and office workers	100	43	5	3.4 year	Exp gr = SMT + mobilization Interv gr = Stretching+ home exercise	7
Zaproudina [146]	105	Advertisement in a local newspaper	33.3	41.5	4.8	11.2 year	Exp gr1 = Traditional bone setting Exp gr2 = PT Interv gr = Massage	7

*BGA= Behavioral graded activity, BQ= Bournemouth questionnaire, CCF= Cranio-cervical flexion, CPR= Clinical prediction rules, DCF= Deep cervical flexor muscles, GP= General practitioner, HVT= High velocity trust, MT= Manual therapy, NAG= Natural apophyseal glide, PSWD= Pulsed Shortwave Diathermy, PT= Physical therapy, ROM= Range of Movement, SMT= Spinal Manipulation Therapy, SNAG = sustain apophyseal natural glide, TENS= Transcutaneous electrical nerve stimulation *All VAS pain scores are converted to a scale of 0-10*

Appendix 3: *HOAC II based clinical reasoning process outcome*

Study	Patient-experienced problem	Cause	Cause related goal	Matched Interven-tion	Goal related out-come	Problem related outcome*
Akther [25]	+ Neck pain > 3 months	-	-	-	-	+ Vas pain, NDI
Ali [26]	+ Neck pain	-	-	-	-	+ Vas pain, NDI
Andrade Orte-ga [27]	+ neck pain > 3 mont hs	-	-	-	-	+ Vas pain, NDI, SF36
Aquino [28]	+ Neck pain > 3 months	-	-	-	-	+ NPRS
Bakar [29]	+ Neck pain 3-6 months	?	+ To stimulate cutaneovis-ceral reflexes	+	+ PPT, EMG	-
Beer [30]	+ Neck pain	++ Relationship between neck pain and reduced muscle function and Poor performance on the CCF	+ To improve deep cervi-cal flexor muscle perfor-mance	+	+ EMG of each stage of CCFT	+ VAS pain, NDI, PSFS
Beinert [31]	+ Neck pain	++ Impaired joint sense	+ To improve balance performance	+	+ JPS	+ NPRS
Beltran-Alacreu [32]	+ Neck pain > 3 months	-	-	-	-	- NDI, TSK, FABQ, NFME, VAFS
Bid [33]	+ Neck pain < 4 weeks, NRS pain 4-9	++ Mechanical neck pain = restricted and painful movements, two positive orthopedic tests indicating facet joint dysfunction	+ To improve facet joint dysfunction	+	-	+ NPRS, ROM, NDI
Borisut [34]	+ Work related neck pain > 6 months; Vas pain>30	+ Reduced deep cervical flexor muscle activity	+ To improve muscle activity	+	+ EMG	+ VAS pain, NDI
Borman [35]	+ Neck pain >6 weeks	-	-	-	-	+ VAS pain, NDI, NHP

Brage [36]	+ Neck pain > 6 months, experienced reduced neck function (NDI> 10/50)	-	-	-	-	+ NPRS, NDI, CCFT, GPE, SF36
Briem [37]	+ Neck pain	+ Aberrant neurophysiology with hypertone muscles	+ Restore ROM	+	+ ROM	-
Casanova-Mendez [38]	+ neck pain > 3 month	+ Sensitized pathways	+ To provide an immediate hypoalgesic effect	+	+ PPT	+ VAS pain, ROM
Celenay [39]	+ Neck pain, >3 months, NDI ≥10/50	+ Aberrant neurophysiology with hypertone muscles	+ To stimulate cutaneovisceral reflexes	+	+ PPT	+ VAS pain, Level of Anxiety (STAI), SF36
Celenay [40]	+ Neck pain, >3 months	+ Reduced cervical and scapula-thoracic stability	-	-	-	+ VAS pain, NDI, PPT, ROM, SF36
Chui [41]	+ neck pain > 3 months	+ Muscle atrophy	+ To improve muscle strength	+	+ Strength	+ VNPS, PIS, NPQ
Chiu [42]	+ neck pain > 3 months	+ Sensitized pathways	+ To stimulate the large-diameter afferent fibers to reduce pain	+	-	+ NPRS, Strength, NPQ
Chui [43]	+ neck pain > 3 months	-	-	-	-	+ VNPS, ROM, NPQ
Cleland [44]	+ Neck pain	+ Relationship between the cervical and thoracic spine	+ Stimulation of descending inhibitory mechanisms	+	-	+ VAS pain
Cleland [45]	+ Neck pain, NDI ≥10	-	-	-	-	+ NPRS, NDI
Cleland [46]	+ Neck pain, NDI ≥10	++ CPR	+ To examine the validity of the CPR	+	+ NPRS	+ NDI
Cook [47]	+ Chronic non-specific neck pain	-	-	-	-	+ NPRS,NDI, Perceived stress scale
Cunha [48]	+ Neck pain > 12 weeks	-	-	-	-	+ VAS pain, ROM, SF 36
David [49]	+ Neck pain > 6 weeks	-	-	-	-	+ VAS pain, ROM, NPQ, GHQ28

Appendix 3: *HOAC II based clinical reasoning process outcome* *(continued)*

Study	Patient-experienced problem	Cause	Cause related goal	Matched Interven-tion	Goal related out-come	Problem related outcome*
Dawood [50]	+ Neck pain > 3 months, NDI > 5	++ Cervical lordotic curve less than 34°	+ Restore the mechanical neck dysfunction	+	+ Cervical curve lordosis test	+ VAS pain, NDI
de Camargo [51]	+ Mechanical neck pain	+ Sensitized pathways	+ To improve muscle func-tion and to provide an immediate hypoalgesic effect	+	+ EMG, PPT	-
Deepa [52]	+ Neck pain	++ The Grieves criteria, Upper thoracic stiffness and tenderness	-	-	-	+ NPRS, ROM, NDI
Dunning [53]	+ Neck pain > 6 weeks	+ Disturbance in joint mo-bility in the upper thoracic spine and cervical spine	+ To increase ROM	+	+ Flexion-rotation test	+ NPRS, CCFT, NDI
Dusunceli [54]	+ Neck pain > 6 weeks	-	-	-	-	+ VAS pain, ROM, NDI, BDS
Dziedzic [55]	+ Neck pain	-	-	-	-	+ NPRS, NPQ, SF12
Escortell-Mayor [56]	+ Neck pain	-	-	-	-	+ VAS pain, NDI, SF12
Evans [57]	+ Neck pain, NPRS ≥ 3	-	-	-	-	+ NPRS, NDI, SF36
Falla [58]	+ Neck pain > 3 months, NDI ≤ 15	+ Muscle fatigability	+To reduce muscle fati-gability	+	+ EMG	+ VAS pain, Strength, NDI
Falla [59]	+ Neck pain > 3 months, NDI ≤ 15	+ Relation between endur-ance-strength and cervi-cal posture	+ Improve endur-ance-strength of the cer-vical muscles to improve control of the cervical posture	+	+ Cervical posture measures	+ VAS pain, NDI
Falla [60]	Neck pain > 3 months, NDI ≤ 15	+ Abnormal muscle acti-vation	+ To change muscle acti-vation during a functional task	+	+ EMG	+ VAS pain, NDI

Falla [61]	+ Neck pain and disability > 1 year	+ Abnormal neural control	+ To improve the neural control of the neck muscles	+	+ EMG	+ VAS pain, NDI, PSFS, SF 36, FABQ
Gallego Izquierdo [62]	+ Chronic neck pain > 3 months, NDI ≤ 15/50	+ Chronic neck pain = reduced neck muscle function + signs of cervical movement control dysfunction, muscle tenderness	+ Restore neck muscle function	+	+ CCFT	+ VAS pain, PPT, NDI
Ganesh [63]	+ Neck pain > 12 weeks	+ Reduction of ROM, provocation during passive accessory movements	+ To improve ROM	+	+ ROM	+ VAS pain, NDI
Giombini [64]	+ Neck pain >3 - < 2 years	+ Impairment in muscle function leads to muscular tension	+ Reduction of muscular tension	+	-	+ VAS pain, NDI, NPDS
Gonzalez-Iglesias [65]	+ Neck pain < 1 month	-	-	-	-	+ VAS pain, ROM, NPQ
Gonzalez-Iglesias [66]	+ Neck pain < 1 month	+ Decreased mobility of the thoracic spine	+ To improve the mobility of the thoracic spine	+	-	+ VAS pain, ROM cervical spine, NPQ
Griffiths [67]	+ Neck pain ≥ 3 months	+ Pain causes inhibition of deep stabilizing muscles	-	-	-	+ NPDS, NPQ, SF36
Griswold [68]	+ Neck pain, NDI ≥ 20%	-	-	-	-	+ NPRS, AROM, DCF endurance, NDI, PSFS, GROC
Haas [69]	+ Neck pain VAS ≥10	+ Endplay restriction	+ Restore normal endplay	+	-	+ VAS pain, VAS stiffness
Häkkinen [70]	+ Neck pain > 6 months	+ Association between pain and reduced strength and neck mobility	+ Restore strength and ROM	+	+ ROM, Isometric neck strength	+ VAS pain
Häkkinen [71]	+ Neck pain > 6 months	-	-	-	-	+ VAS pain, isometric strength, ROM, NPDS, NDI

Appendix 3: *HOAC II based clinical reasoning process outcome (continued)*

Study	Patient-experienced problem	Cause	Cause related goal	Matched Interven-tion	Goal related out-come	Problem related outcome*
Helewa [72]	+ 2 months ≥ neck pain ≤ 12 months	-	-	-	-	+ VAS pain, grip and neck strength, Tender point count and score, NPQ, SF36
Hoving [73]	+ Neck pain or stiff-ness ≥ 2 weeks	-	-	-	-	+ NPRS, ROM, NDI, Perceived recovery scale, EUQLI
Hudson [74]	+ Recurrent or chronic neck pain > 3 months	-	-	-	-	+ NPRS, NDI
Humphreys [75]	+ Neck pain > 3 months	+ Proprioceptive dysfunc-tion	+ To improve head reposi-tioning accuracy	+	+ HRA	+ VAS pain
Izquierdo Perez [76]	+ Neck pain > 3 months	-	-	-	-	+ VAS pain, ROM, NDI, GROC
Javanshir [77]	+ neck pain > 3 months	?	+ To activate cervical flexor muscles	+	+ muscle size mea-surements	+ NPRS, NDI
Jordan [78]	+ Neck pain ≥ 3 months	-	-	-	-	+ NPRS, Doctor's GA, PPE, DI
Jull [79]	+ Neck pain > 3 months	++ Abnormal joint position sense (JPS)	+ To improve JPS	+	+ JPS	+ VAS pain, NDI
Jull [80]	+ Neck pain> 3 months, NDI< 15	++ Poor performance on the CCFT	+ To improve muscle ac-tivity during the CCFT	+	+ EMG	+ NPRS, NDI, ROM, PPE
Kanlayanaphot-porn [81]	+ Neck pain > 1 week, VAS pain > 20mm	-	-	-	-	+ VAS pain, ROM

Kanlayanaphot-porn [82]	+ Acute neck pain = < 30 days, subacute 30-90 days, chronic >90 days; VAS pain > 20mm	-	-	-	-	+ VAS pain, ROM
Karlsson [83]	+ Neck pain > 6 months, NRS pain > 3, NDI ≥ 10	-	-	-	-	+ NPRS, ROM, Strength, NDI
Khan [84]	+ Neck pain be-tween 3 months- 3 years	+ neck pain correlates with weakness in muscles	-	-	-	+ VAS pain, ROM, NPQ
Kim [85]	+ Neck pain > 3 months, NDI 5-14	+ Reduced soft tissue functioning	+ Restore ROM	+	+ ROM	+ VAS pain, NDI
Kim [86]	+ Neck pain > 3 months, NDI≤ 15	++ Reduced endurance	+ Restore endurance to restore forward head posture	+	+ Head tilt angle, neck flexion angle, forward shoulder angle	+ NPRS, NDI
Kjellman [87]	+ Neck pain	++ 1 of 4 pain provoking tests was positive	-	-	-	+ VAS pain, NDI, NRS general health, DRAM
Klaber-Moffett [88]	+ Neck pain > 2 weeks	+ Psychosocial factors	-	-	-	- Distress scale, TSK, NPQ, SF 36
Ko [89]	+ Neck pain	+ Reduced endurance, disturbance in joint mobil-ity in the upper thoracic spine	+ To improve muscular endurance	+	+ Endurance	+ Vas pain, NDI
Krauss [90]	+ Neck pain > 3 months, limited ROM	++ Thoracic segmental motion restriction	-	-	-	+ NPRS, cervical ROM
Kumar [91]	+ Neck pain or stiffness	-	-	-	-	+ VAS pain, ROM
Lansinger [92]	+ Neck pain > 3 months	-	-	-	-	+ VAS pain, ROM, Grip strength, NDI
Lansinger [93]	+ Neck pain ≥ 3 months, VAS pain ≥ 20 mm	-	-	-	-	- SF36

2

Appendix 3: *HOAC II based clinical reasoning process outcome* *(continued)*

Study	Patient-experienced problem	Cause	Cause related goal	Matched Interven-tion	Goal related out-come	Problem related outcome*
Lau [94]	+ Mechanical neck pain > 3 months	-	-	-	-	+ NPRS, ROM, cranio vertebral angle, NPQ, SF36
Leaver [95]	+ Neck pain < 3 months	-	-	-	-	+ NPRS, NDI, PSFS, PPE, SF12
Lee [96]	+ Neck pain	++ Forward head posture with > 15 mm anterior weight bearing	+ To reduce the forward head posture	+	+ CVA, CRA	-
Lee [97]	+ Chronic neck pain > 3 months, NDI > 20	++ Limited craniocervical and thoracic flexion and extension ROM	+ Restore ROM	+	+ Thoracic ROM	+ VAS pain, NDI, Strength, endur-ance,
Lee [98]	+ Neck pain	+ Relation between cervi-cal overload and reduced proprioception	+ Improve proprioception	+	+ HRA	-
Lluch [99]	+ Neck pain > 3 months, NDI ≥ 5	++ Scapular dyskinesis	+ To correct scapular position	+	-	+ NPRS, PPT
Lluch [100]	+ Neck pain ≥ 3 months, NPRS ≥ 3	+ Chronic neck pain leads to reduced motor control	+ To improve motor con-trol	+	+ EMG of each stage of CCFT	+ Resting pain, PPT, ROM
Lopez-Lopez [101]	+ Neck pain > 12 weeks	-	-	-	-	+ VAS pain, ROM by a CROM, PPT
Maayah [102]	+ Neck pain	?	+ To provide an immedi-ate hypoalgesic effect	+	+ Neck pain thresh-old myometer machine	-
Madson [103]	+ Chronic neck pain ≥ 3 months	-	-	-	-	+ VAS pain, NDI
Mansilla-Ferra-gut [104]	+ Mechanical bilat-eral neck pain > 6 months	++ Active mouth opening < 40 mm	+ To improve active mouth opening, to de-crease sensitivity of the trigeminal nerve	+	+ PMMO, PPT	-

Martel [105]	+ Neck pain > 12 weeks	-	-	-	-	+ VAS pain, ROM, NDI, BQ, FABQ, SF12
Martinez-Segura [106]	+ Neck pain > 1 month	++ joint dysfunction C3-C5	-	-	-	+ VAS pain, ROM
Martinez-Segura [107]	+ Neck pain	?	+ To provide an immediate hypoalgesic effect	+	+ PPT	- ROM
Masaracchio [108]	+ Neck pain < 3 months, NDI > 20%	-	-	-	-	+ NPRS, NDI, GROC
McLean [109]	-	-	-	-	-	- NPQ, DASH
Monticone [110]	+ Neck pain > 3 months	-	-	-	-	+ NPRS, NPDS, SF36
O'Leary [111]	+ Neck pain ≥ 3 months, NDI score 10-28	++ positive findings in the physical exam and on CCFT	+ To improve muscle performance	+	+ Maximal voluntary contraction (MVC), endurance with 50% MVC with the CCFT	-
O'Leary [112]	+ Neck pain > 3 months, score NDI ≥ 5	?	+ To provide an immediate hypoalgesic effect and sympathoexcitation	+	+ PPT, TPT, Neck blood flow, Hand blood flow, skin conductance left and right hand, Skin temperature neck and hand, Heart rate, Blood pressure systolic and diastolic	+ VAS pain-rest, VAS pain-active
O'Leary [113]	+ Neck pain >6 months, NDI between 10-15, positive findings in the physical exam	+ Association between chronic neck pain and motor performance	+ To improve motor performance	+	+ ROM, strength, endurance, EMG	+ VAS pain, NDI
Paoloni [114]	+ Chronic neck pain, VAS pain ≥ 40 mm	-	-	-	-	+ VAS pain, NPDS
Pillastrini [115]	+ Neck pain > 3 months	++ dysfunctional muscle chains	+ To normalize muscle chains	+	-	+ VAS pain, NPDS, ROM

Appendix 3: *HOAC II based clinical reasoning process outcome* *(continued)*

Study	Patient-experienced problem	Cause	Cause related goal	Matched Intervention	Goal related outcome	Problem related outcome*
Pires [116]	+ Chronic pain or fatigue in the cervical region ≥ 6 months, NDI ≥ 5	+ Increase in myoelectric activity by reduction in thoracic mobility	+ To reduce myoelectric activity	+	+ EMG	+ VAS pain
Pool [117]	+ 4weeks > neck pain < 12 weeks	-	-	-	-	+ NPRS, NDI
Puentedura [118]	+ NDI>10	++ 4 of 6 criteria used by Cleland et al	-	-	-	+ NPRS, FABQ, NDI
Puntumetakul [119]	Chronic mechanical neck pain ≥ 3 months, NDI ≥ 10%	-	-	-	-	+ VAS pain, NDI, ROM
Rendant [120]	+ Neck pain > VAS 40, duration between 6 months and 5 years, normal cervical spine flexibility	-	-	-	-	+ VAS pain, NPDS, General self-efficacy scale, SF36
Revel [121]	+ Neck pain >3 months	+ Reduced cervicocephalic kinesthetic sensibility	+ To improve cervicocephalic kinesthesia	+	+ HRA	+ VAS pain
Rolving [122]	+ Neck pain	+ Relation between chronic neck pain and reduced strength	+ To improve strength	+	+ Strength	+ NPRS, Fear avoidance beliefs
Rudolfsson [123]	+ Neck pain > 3 months, > 9 points of the first 19 items in the DASH	+ Relation between neck pain and sensorimotor impairments	+ To improve sensorimotor function	+	+ Cop-A, Ra-A, Tr-A, VE	+ NPRS, ROM
Saavedra-Hernandez [124]	+ Neck pain	-	-	-	-	+ NPRS, ROM, NDI
Saavedra-Hernandez [125]	+ Chronic neck pain	-	-	-	-	+ NPRS, ROM, NDI
Saayman [126]	+ Neck pain	++ Cervical facet dysfunction	+ Restore joint functioning	+	-	+ NPRS, ROM, NDI

Salom-Moreno [127]	+ Neck pain > 6 months, bilateral symptoms,	?	+ To activate descending systems from the dPAG	+	+ PPT	+ NPRS
Sarig Bahat [128]	+ Neck pain > 3 months, NDI >10%	+ People with chronic neck pain has kinematic impairments	+ To improve kinematics	+	+ Virtual reality kinematics	+ VAS, NDI, TSK, sensorimotor measurements
Schomacher [129]	+ Neck pain changed with movement between C2-7	-	-	-	-	+ NPRS, NRS SoM
Sherman [130]	+ Neck pain ≥ 3 months, NPRS > 4, NDI >5	-	-	-	-	+ NPRS, NDI
Sillevis [131]	+ Neck pain > 3 months	+ Central sensitization	+ To improve the autonomic nervous system functioning	+	+ pupil diameter	+ VAS pain
Snodgrass [132]	+ Neck pain > 3 months, NPRS > 3	-	-	-	-	+ VAS pain, PPT, ROM, spinal stiffness
Sterling [133]	+ Neck pain > 3 months, symptoms originating from the C5/6 segment	+ Neck pain leads to reduced muscle control	+ To activate descending systems from the dPAG and to restore muscle control	+	+ PPT, Thermal pain thresholds, EMG/ CCFT	+ VAS pain
Taimela [134]	+ Neck pain > 3 months, height >140 cm	-	-	-	-	+ VAS pain, ROM, PPT
Thompson [135]	+ Chronic neck pain ≥ 3 months	-	-	-	-	+ NPRS, NPQ, PCS, TSK, CPSS PVAQ
Vernon [136]	+ Neck pain	?	+ To increase pain thresholds	+	+ PPT	-
Viljanen [137]	+ Neck pain > 12 weeks	-	-	-	-	+ NPRS, dynamic muscle strength, ROM, NDI
von Trott [138]	+ Neck pain ≥ 6 months	-	-	-	-	+ VAS pain, NPDS, ADS, SF36

Appendix 3: *HOAC II based clinical reasoning process outcome* *(continued)*

Study	Patient-experienced problem	Cause	Cause related goal	Matched Interven-tion	Goal related out-come	Problem related outcome*
Vonk [139]	+ Neck pain > 3 months	+ psychological and social factors and maladaptive behaviors	-	-	-	+ NPRS, TSK, NDI, MPI, CPSS(PSE, FSE, CSE), PCS, CES-D, EQ
Walker [140]	+ Neck pain VAS≥30, NDI ≥ 10	-	-	-	-	+ VAS pain, NDI, PPE
Yang [141]	+ neck pain > 3 months	+ chronic neck pain is associated with instability which is accompanied by loss of somatosensory system function	+ Restore proprioception	+	+ cervical reposi-tion test	+ VAS pain
Ylinen [142]	+ Neck pain > 6 months	-	-	-	-	+ VAS pain, ROM, NDI, NPDS
Ylinen [143]	+ Neck pain > 6 months	+ Sensitized pathways	+ To provide an hypoalge-sic effect	+	+ PPT	+ VAS pain
Ylinen [144]	+ Neck pain > 6 months	-	-	-	-	+ VAS pain, ROM, Isometric neck strength, PPT, NDI, NSPI
Ylinen [145]	+ Neck pain > 6 months	-	-	-	-	+ VAS pain, NDI, NPDS
Zaproudina [146]	+ Neck pain VAS ≥ 30	-	-	-	-	+ VAS pain, FKD, ROM, NDI

*ADS= general scale of depression, BDI = beck depression inventory, BDS= Beck Depression Scale, BQ= Bournemouth questionnaire, CCFT= Cranio-cervical flexion test, CES-D= center for epidemiologic studies depression scale, CoP-A = Centre of pressure area, CPSS= chronic pain self-efficacy scale, CRA= cranial rotation angle, CROM= Cervical Range Of Motion device, CSE= Coping self-efficacy, CVA= cranial vertical angle, DASH= Disabilities of the arm, shoulder and hand, DCF= Deep cervical flexor muscle; DI= Disability index, dPAG = dorsal periaqueductal gray, DRAM= Distress and Risk assessment method, EMG= Electromyography, EQ= Euroquol-5D, EUQLI= Euro quality of life index, FABQ= Fear Avoidance Beliefs Questionnaire, FKD= Frons-knee distance, FSE= Functional self-efficacy, GA= Global assessment, GHQ28= General Health Questionnaire 28, GPE= General perceived effect, GROC = Global rating of change, HRA= Head replacement accuracy, JPS= Joint Position Sense, MPI= multi-dimensional pain inventory, NDI= Neck Disability Scale, NFME= Neck flexor muscle endurance; NHP= Nottingham Health Profile (Quality of life), NPQ= Northwick Park Neck Pain Questionnaire, NSPI = neck and shoulder pain and disability index, NPDS= Neck pain and disability scale, NRS= Numeric rating scale, NPRS= Numeric Pain Rating Scale, PCS= pain catastrophizing scale, PIS= peak isometric strength, PMMO= Pain-free maximal mouth opening, PPE= Patient perceived effect, PPT= pressure pain threshold, PSE= Pain self-efficacy, PSFS= Patient specific function scale, PVAQ= Pain vigilance and awareness questionnaire, Ra-A = Rambling area, ROM= Range of Movement, SF 12, 36= Short Form 12, 36, SoM= sensation of Movement, STAI= State trait anxiety inventory, TPT= Thermal pain threshold, Tr-A = Trembling area, TSK= Tampa scale for kinesiophobia, VAS= Visual Analogue Scale, VE = variability in end point precision, VNPS= Verbal Numeric Pain Scale *The patient-experienced problem related outcomes are underlined, the other outcomes not.*

3

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

François Maissan
Jan Pool
Edwin de Raaij
Harriët Wittink
Raymond Ostelo

Musculoskeletal Science and Practice 47 (2020) 102133

Abstract

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.

Keywords: Treatment-based classification system; Neck pain,; Physiotherapy; Systematic review

Introduction

Neck pain is the fourth major cause of disability worldwide [1]. 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 [2]. In 2015, more than a third of a billion people worldwide had neck pain of more than 3 months duration [3].

At least six Cochrane reviews focussing on physiotherapy interventions for patients with neck pain reported inconclusive evidence for their effectiveness [4-9]. 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 [10], has been a research priority for the last few years [11] as it might increase the effectiveness of the interventions [12]. 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)) [13-15].
The development of a TBCS can be achieved through a (clinical) judgement approach and/or a statistical approach [16]. 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.

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.

Methods

Design.
This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [17] and registered in the international prospective register of systematic reviews PROSPERO (CRD 42018087763).

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**.

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 (18). 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.

Data extraction and Quality assessment
Description
We used a framework, used in multiple reviews, to describe the characteristics of a classification system [19]. 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.

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 [19], see **Table 1.** The overall inter-rater reliability of the Buchbinder scale had an Intraclass Correlation Coefficient (ICC) of 0.82 [19]. 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 - 7.

Two authors (FM, JP) independently extracted the data, using the guidance as described previously [16]. 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"

Table 1:
Criteria used to appraise the quality of treatment based classification systems [19]

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?
Generalisability	Has it been used in other studies and/or settings?***

Operationalisation after pilot testing.
**Judgement based development. A Yes when more methods were used besides the judgement of one person or small group of physiotherapists such as reviewing the literature. If one method was used a score of Partial and if it was unclear how a judgement had been formed we scored No.*
***How long does the TBCS take to perform. Achievable in a standard physiotherapeutic examination of 30 minutes as Yes, if not as Partial and No if the amount of time remained unclear.*
****Has it been used in other studies and/or settings? We scored a Yes if the other study, that applied the TBCS, also included a non-specific neck pain population, a No if it has been used in other populations or has not been used at all in other studies.*

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)[20]. 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)[21].

Effectiveness

We assessed the risk of bias of the RCTs using the PEDro scale (www.pedro.org.au)[22]. The PEDro scale has moderate-to-good reliability with an ICC of 0.68 (95% confidence interval (CI) 0.57 to 0.76)[23]. We considered RCTs with a score of > six points on the PEDro scale as studies with a low risk of bias [24].

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.

Results

Search results for TBCSs

The literature search retrieved 7664 studies: after removing duplicates, 6051 remained for further screening. **Figure 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 [25-42].

We identified 13 different TBCSs [25,26,30,32-36,38-42]. Two TBCSs were very similar, but not identical [26,34]. Fritz et al. used the proposed classification system from Childs et al. to develop an algorithm to prioritize the findings and place each patient into a classification category. This algorithm is slightly different from that of Childs et al. [26] due to differences in criteria and interventions **(Table 3)**. Therefore, we included both as separate TBCSs and considered both studies as development studies.

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 [25]. 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 [30,33,35,39-41] and are all referred to as Clinical Prediction Rules (CPRs) [43,44] Seven TBCSs used a judgment-based approach [25,26,32,34,36,38,42].

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, Figure 2 and Appendix 2)**.

We included five reliability studies [25,27-29,34] for four TBCSs: STarT Back tool, McKenzie system, Cleland classification system and Fritz [25,30,34,36]. 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) [26,32,38,42], while one TBCS gained the highest score of 5 (out of 7) points [25]. 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 [31,37,45,46] supporting the generalizability of these TBCSs. **Figure 2** shows the

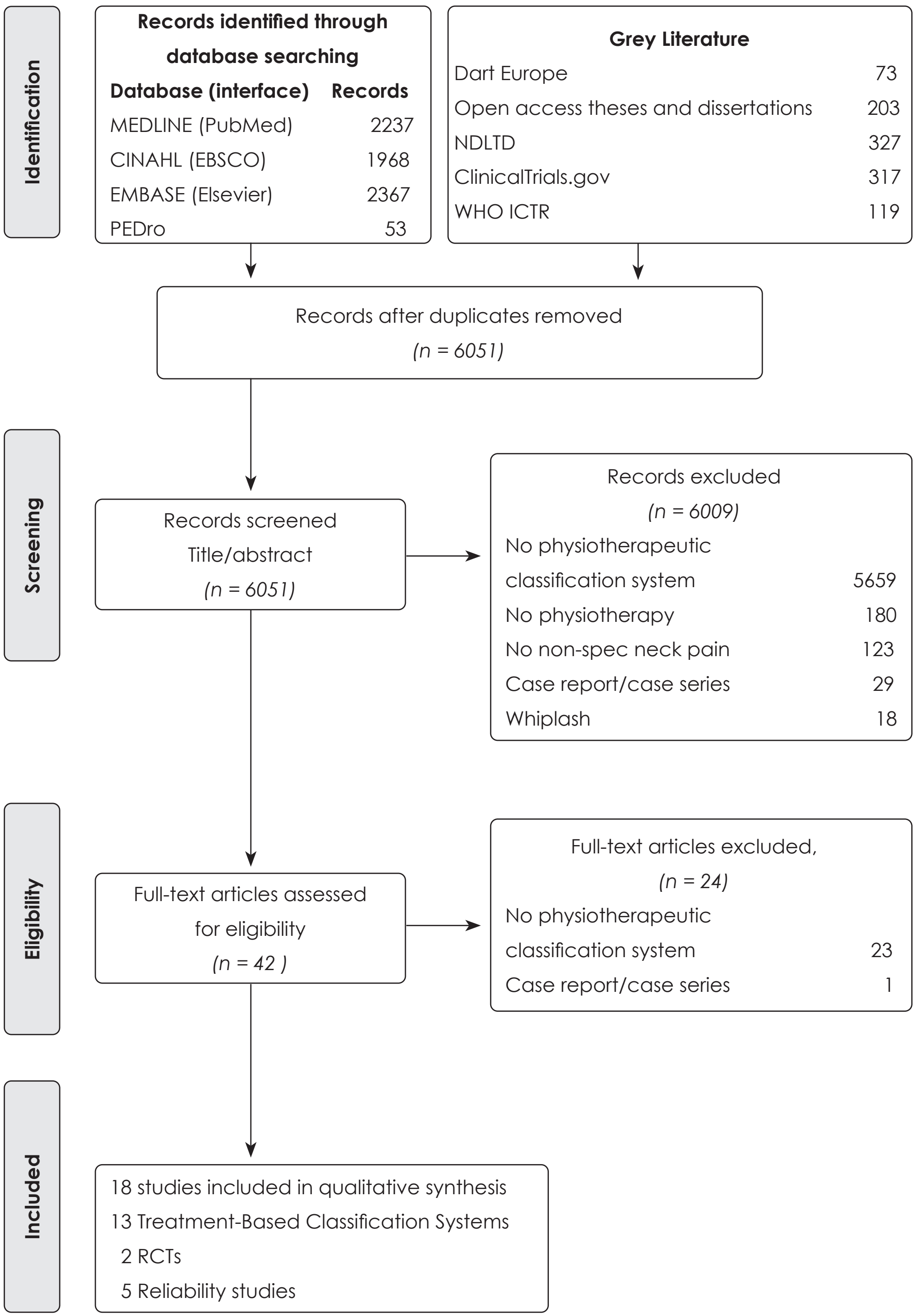


Figure 1: *Flowchart of articles reviewed*

summary of the quality of the 13 classification systems. The overall quality of the TBCSs ranged from low to moderate.

Effectiveness of TBCSs

Two RCTs investigated the effectiveness of two TBCSs: the Cleland classification system and the McKenzie system [31,37].

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 [47-49] 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 [31].

Concerning the McKenzie system; one study compared the effectiveness of three groups: McKenzie treatment, general exercise and a control group [37]. 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).

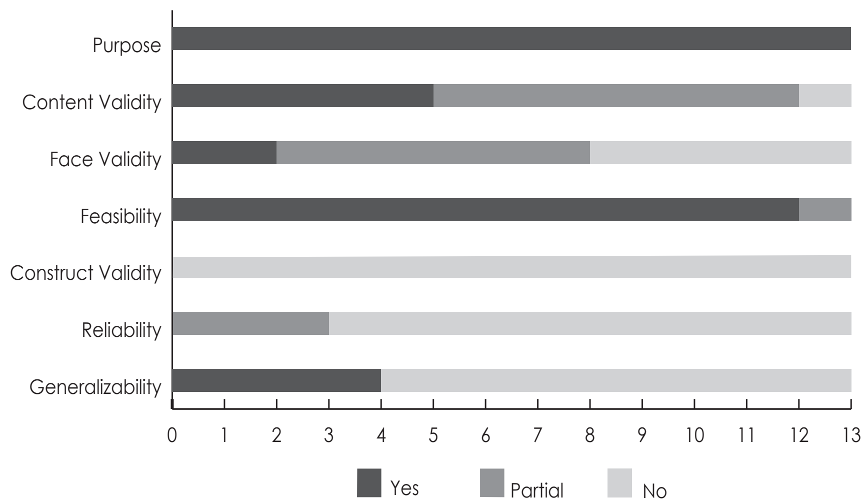


Figure 2: *Quality summary of the 13 classification systems, based on the appraisal tool*

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 (30)	CPR to identify patients with neck pain who are likely to benefit from thoracic spinethrust 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
de las Peñas (33)	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 (35)	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 (39)	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

Table 2 : *The general part of the framework to describe treatment-based classification systems and the critical appraisal score* *(continued)*

Primary author	Purpose	Method of development	Domain of interest	Specific exclusion	Critical appraisal
Raney (40)	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 (41)	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 (45)	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 (26)	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 (32)	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 (34)	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 (36)	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 (38)	A self-classification system for a smartphone-based exercise program feasible for office workers as a method of self-managing their neck pain.	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
Wang (42)	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.	Neck pain with or without radiating pain.	2.5

3

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

Primary author	Categories	Criteria used	Treatment
Cleland (30)	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.
de las Peñas (33)	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 (35)	1. CPR for a standardized program of stretching and muscle perfmance 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 (39)	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 (40)	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 (41)	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 (45)	1. Low risk for persisting disability	LR 1. Total StartBack score ≤ 3 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.
	2. Moderate risk for persisting disability	MR 1. Total StartBack score ≥4; Score question 5-9 ≤ 3 points	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.
	3. High risk for persisting disability	HR 1. Total StartBack score ≥ 4; Score question 5-9 ≥ 4 points	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.

Table 3 : *The specific part of the framework to describe classification systems (continued)*

Primary author	Categories	Criteria used	Treatment
Childs (26)	1. Mobility	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	MB 1. Cervical and thoracic spine mobilization/manipulation. 2. Active range of motion exercises.
	2. Centralization	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	CZ 1. Mechanical/manual cervical traction. 2. Repeated movements to centralize symptoms.
	3. Conditioning and increase exercise tolerance	CD 1. Lower pain and disability scores 2. Longer duration of symptoms 3. No signs of nerve root compression 4. No peripheralization/centralization during range of motion	CD 1. Strengthening and endurance exercises for the muscles of the neck and upper quarter. 2. Aerobic conditioning exercises.
	4. Pain control	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	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.
	5. Reduce head-ache	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	RH 1. Cervical spine manipulation /mobilization. 2. Strengthening of neck and upper quarter muscles. 3. Postural education.

Dewitte (32)	1. Cervical spine convergence 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	CCP 1. Distraction technique; gapping Technique. 2.Translatoric technique - indirect upslope technique or direct downslope technique.
	2. Cervical spine divergence pattern	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	CDP 1. Distraction technique. 2. Translatoric upslope technique - focus or locking approach.

3

Table 3 : *The specific part of the framework to describe classification systems (continued)*

Primary author	Categories	Criteria used	Treatment
Fritz (34)	1. Mobility	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	MOB 1. Cervical or thoracic mobilization or manipulation. 2. Strengthening exercises for the deep neck flexor muscles.
	2. Centralization	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	CEN 1. Mechanical or manual cervical traction (at least 50% of the sessions). 2. Cervical retraction exercises (at least 50% of the sessions).
	3. Exercise and conditioning	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	EAC 1. Strengthening exercises for the upper quarter muscles. 2. Strengthening exercises for the neck or deep neck flexor muscles.
	4. Pain control	PC 1. Mode of onset no whiplash mechanism 2. Duration of symptoms < 30 days 3. NPRS > 7 or NDI score > 52/100	PC 1. Cervical spine mobilization. 2. Cervical range-of-motion exercises.
	5. Headache	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	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 (36)	1. Posture sydrome	PS 1. pain arising as a result of mechanical deformation of normal soft tissues from prolonged end range loading of periarticular structures	PS 1. Posture correction.
	2. Dysfunction syndrome	DyS 1. pain occurring as a result of mechanical deformation of structurally impaired tissues (such as tissue which is scarred, adhered or adaptively shortened).	DyS 1. Exercise into the direction of the dysfunction with the aim of remodeling the tissue.
	3. Derangement syndrome	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.	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.
	4. Other	Ot 1. Those who do not fit the mechanical syndromes but who exhibit signs and symptoms of other known pathology	

3

Table 3 : *The specific part of the framework to describe classification systems (continued)*

Primary author	Categories	Criteria used	Treatment
Lee (38)	1. Exercise and conditioning	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 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	EaC 1. Strengthening exercise for deep neck muscles and upper-quarter muscles.
	2. Mobility	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	MoB 1. Stretching exercise. 2. Strengthening exercise for deep neck muscles.
	3. Centralization	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	CeN 1. Cervical retraction exercise and repeated neck extension exercise. 2. Strengthening exercise for deep neck muscles.
	4. Reduce headache	RH 1. Headache triggered by neck movement or position 2. Headache elicited by pressure on the ipsilateral posterior neck	RH 1. Self-myofascial release technique. 2. Strengthening exercise for deep neck and upper-quarter muscles.

Wang (42)	1: Radicular arm pain or neck pain; 3 patterns	RAD-1 1. Positive neurologic signs 2. Centralization or distal symptom reduction occurred with repeated movements	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. Positive neurologic signs 2. Centralization or distal symptom reduction did not occur with repeated movements 3. Manual traction decreased symptoms	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. Positive neurologic signs 2. Centralization or distal symptom reduction did not occur with repeated movements 3. Manual traction did not decrease symptoms	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.

3

Table 3 : *The specific part of the framework to describe classification systems* *(continued)*

Primary author	Categories	Criteria used	Treatment
Wang (42) (continued)	2: Referred arm pain or neck pain; 6 patterns	REF-1 1. Negative neurologic signs 2. Referred arm pain with or without neckpain 3. Centralization or distal symptom reduction occurred with repeated movements	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. 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-2 1. Mechanical positional cervical traction 2. Specific level manual traction 3. Postural exercise 4. Education
		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-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. 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-4 1. Thoracic outlet release techniques (tissue- specific). 2. Specific mobilization and stretching. 3. Postural exercise. 4. Activity tolerance training. 5. Education.

Wang (42) (continued)		REF-5 1. Negative neurologic signs 2. Referred arm pain with or without neckpain 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-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. 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	REF-6 1. Specific joint mobilization. 2. Mechanical traction.
	3: cervicogenic headaches: 4 patterns	HA-1 1. Negative neurologic signs 2. Referred pain into the temporal/facial area	HA-1 1. Temporomandibular joint treatment protocol.
		HA-2 1. Negative neurologic signs 2. Referred pain in non-facial areas 3. Occipito-atlantal (OA) joint distraction provoked or reduced symptoms	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. 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-3 1. Specific joint mobilization or muscle energy techniques to AA joint. 2. Postural exercise.
		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	HA-4 1. Specific joint mobilization or muscle energy techniques to the involved level(s). 2. Mechanical traction in the absence of specific joint mobilization technique.

Table 3 : *The specific part of the framework to describe classification systems (continued)*

Primary author	Categories	Criteria used	Treatment
Wang (42) (continued)	4: Neck Pain only; 5 patterns	NP-1 1. Negative neurologic signs 2. Neck pain only 3. Gross movement tests showed capsular pattern restriction	NP-1 1. Mechanical traction. 2. Manual traction. 3. Neck range of motion exercise. 4. Postural exercise.
		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-2 1. Specific joint mobilization (muscle energy or dorsal gliding techniques) to the involved level(s). 2. Postural exercise.
		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-3 1. Specific joint mobilization (muscle energy or gliding techniques) to the adjacent level(s). 2. Stabilization exercise.
		NP-4 1. Negative neurologic signs 2. Neck pain only 3. Gross movement tests showed non-ca sular 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-4 1. Specific joint mobilization (muscle energy or ventral gliding techniques) to the involved level(s). 2. Postural exercise.
		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	NP-5 1. Specific joint mobilization (muscle energy or gliding techniques) to the adjacent level(s). 2. Stabilization exercise.

Discussion

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.

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" [50].

Judgement based TBCSs also had poor face validity. Exemplary for 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[51]. 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 [52].

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.

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 [53]. Another review that critically appraised statistically-derived TBCSs focused on musculoskeletal conditions[54] concluded that "at present, there is little evidence that statistically-derived TBCSs can be used to predict effects of treatment for musculoskeletal conditions".

Strengths and weaknesses

As far as we know, this review is the first review focusing 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 [55].

Implications

One important feature of a TBCS is the clinical relevance [50]. 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 [56,57].
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.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. 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 2012;380(9859):2163-2196.
2. March L, Smith EU, Hoy DG, Cross MJ, Sanchez-Riera L, Blyth F, et al. Burden of disability due to musculoskeletal (MSK) disorders. Best Pract Res Clin Rheumatol 2014;28(3):353-366.
3. Hurwitz EL, Randhawa K, Yu H, Cote P, Haldeman S. The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. Eur Spine J 2018;27 (suppl6):796-801.
4. Ezzo J, Haraldsson BG, Gross AR, Myers CD, Morien A, Goldsmith CH, et al. Massage for mechanical neck disorders: a systematic review. Spine 2007;32(3):353-362.
5. Graham N, Gross A, Goldsmith CH, Klaber Moffett J, Haines T, Burnie SJ, et al. Mechanical traction for neck pain with or without radiculopathy. Cochrane Database Syst Rev 2008;(3):CD006408.
6. Gross A, Forget M, St George K, Fraser MM, Graham N, Perry L, et al. Patient education for neck pain. Cochrane Database Syst Rev 2012;3:CD005106.
7. Gross AR, Hoving JL, Haines TA, Goldsmith CH, Kay T, Aker P, et al. A Cochrane review of manipulation and mobilization for mechanical neck disorders. Spine 2004;29(14):1541-1548.
8. Kay TM, Gross A, Goldsmith CH, Rutherford S, Voth S, Hoving JL, et al. Exercises for mechanical neck disorders. Cochrane Database Syst Rev 2012;8:CD004250.
9. Monticone M., Ambrosini E., Cedraschi C., Rocca B., Fiorentini R., Restelli M., et al. Cognitive-behavioral Treatment for Subacute and Chronic Neck Pain: A Cochrane Review. Spine 2015;40(19):1495-1504.
10. Linton SJ, Nicholas M, Shaw W. Why wait to address high-risk cases of acute low back pain? A comparison of stepped, stratified, and matched care. Pain 2018;159(12):2437-2441.
11. Foster NE, Hill JC, Hay EM. Subgrouping patients with low back pain in primary care: are we getting any better at it? Man Ther 2011;16(1):3-8.
12. Coupe VMH, van Hooff ML, de Kleuver M, Steyerberg EW, Ostelo RWJG. Decision support tools in low back pain. Best Pract Res Clin Rheumatol 2016;30(6):1084-1097.
13. Damgaard P, Bartels EM, Ris I, Christensen R, Juul-Kristensen B. Evidence of Physiotherapy Interventions for Patients with Chronic Neck Pain: A Systematic Review of Randomised Controlled Trials. ISRN Pain 2013;2013:567175.
14. Tsakitzidis G, Remmen R, Dankaerts W, van Royen P. Non-specific neck pain and evidence-based practice. ESJ 2013;9(3):1-19.
15. Fairbank J, Gwilym SE, France JC, Daffner SD, Dettori J, Hermsmeyer J, et al. The role of classification of chronic low back pain. Spine (Phila Pa 1976) 2011;36(21 Suppl):S19-42.
16. Riddle DL. Classification and low back pain: a review of the literature and critical analysis of selected systems. Phys Ther 1998;78(7):708-737.

17. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open Med 2009;3(3):e123-30.
18. Hogg-Johnson S, van der Velde G, Carroll LJ, Holm LW, Cassidy JD, Guzman J, 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. Spine (Phila Pa 1976) 2008;33(4 Suppl):S39-51.
19. Buchbinder R, Goel V, Bombardier C, Hogg-Johnson S. Classification systems of soft tissue disorders of the neck and upper limb: do they satisfy methodological guidelines? J Clin Epidemiol 1996;49(2):141-149.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-174.
21. Portney LG, Watkins MP. Foundations of clinical research. second edition ed. New Jersey: Prentice Hall Health; 2000; ISBN 0-8385-2695-0.
22. de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. Aust J Physiother 2009;55(2):129-133.
23. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther 2003;83(8):713-721.
24. Veerbeek JM, Koolstra M, Ket JC, van Wegen EE, Kwakkel G. Effects of augmented exercise therapy on outcome of gait and gait-related activities in the first 6 months after stroke: a meta-analysis. Stroke 2011;42(11):3311-3315.
25. Bier JD, Ostelo RWJG, Koes BW, Verhagen AP. Validity and reproducibility of the modified STarT Back Tool (Dutch version) for patients with neck pain in primary care. Musculoskelet Sci Pract 2017;31:22-29.
26. Childs JD, Fritz JM, Piva SR, Whitman JM. Proposal of a classification system for patients with neck pain. J Orthop Sports Phys Ther 2004;34(11):686-96; discussion 697-700.
27. Clare H.A., Adams R., Maher CG. Reliability of the McKenzie spinal pain classification using patient assessment forms. Physiotherapy 2004;90:114-119.
28. Clare HA, Adams R, Maher CG. Reliability of McKenzie classification of patients with cervical or lumbar pain. J Manipulative Physiol Ther 2005;28(2):122-127.
29. Cleland JA, Childs JD, Fritz JM, Whitman JM. Interrater reliability of the history and physical examination in patients with mechanical neck pain. Arch Phys Med Rehabil 2006;87(10):1388-1395.
30. Cleland JA, Childs JD, Fritz JM, Whitman JM, Eberhart SL. 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 2007;87(1):9-23.
31. Cleland JA, Mintken PE, Carpenter K, Fritz JM, Glynn P, Whitman J, et al. 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 2010;90(9):1239-1250.
32. Dewitte V, Beernaert A, Vanthillo B, Barbe T, Danneels L, Cagnie B. Articular dysfunction patterns in patients with mechanical neck pain: a clinical algorithm to guide specific mobilization and manipulation techniques. Man Ther 2014;19(1):2-9.
33. Fernandez-de-las-Penas C, Cleland JA, Palomeque-del-Cerro L, Caminero AB, Guillem-Mesado A, Jimenez-Garcia R. 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 2011;51(2):246-261.
34. Fritz JM, Brennan GP. Preliminary examination of a proposed treatment-based classification system for patients receiving physical therapy interventions for neck pain. Phys Ther 2007;87(5):513-524.
35. Hanney WJ, Kolber MJ, George SZ, Young I, Patel CK, Cleland JA. 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 2013;8(6):756-776.
36. Hefford C. McKenzie classification of mechanical spinal pain: profile of syndromes and directions of preference. Man Ther 2008;13(1):75-81.
37. Kjellman G, Oberg B. A randomized clinical trial comparing general exercise, McKenzie treatment and a control group in patients with neck pain. J Rehabil Med 2002;34(4):183-190.
38. Lee M, Lee SH, Kim T, Yoo HJ, Kim SH, Suh DW, et al. 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 2017;98(1):80-87.
39. Puentedura EJ, Cleland JA, Landers MR, Mintken PE, Louw A, Fernandez-de-Las-Penas C. 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 2012;42(7):577-592.
40. Raney NH, Petersen EJ, Smith TA, Cowan JE, Rendeiro DG, Deyle GD, et al. Development of a clinical prediction rule to identify patients with neck pain likely to benefit from cervical traction and exercise. Eur Spine J 2009;18(3):382-391.
41. Saavedra-Hernandez M, Castro-Sanchez AM, Fernandez-de-Las-Penas C, Cleland JA, Ortega-Santiago R, Arroyo-Morales M. 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 Manipulative Physiol Ther 2011;34(3):144-152.
42. Wang WT, Olson SL, Campbell AH, Hanten WP, Gleeson PB. Effectiveness of physical therapy for patients with neck pain: an individualized approach using a clinical decision-making algorithm. Am J Phys Med Rehabil 2003;82(3):203-18; quiz 219-21.

43. Beattie P, Nelson R. Clinical prediction rules: what are they and what do they tell us? Aust J Physiother 2006;52(3):157-163.
44. Randolph AG, Guyatt GH, Calvin JE, Doig G, Richardson WS. Understanding articles describing clinical prediction tools. Evidence Based Medicine in Critical Care Group. Crit Care Med 1998;26(9):1603-1612.
45. Bier JD, Sandee-Geurts JJW, Ostelo RWJG, Koes BW, Verhagen AP. 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 2018;99(1):65-71.
46. Farrell KP, Lampe KE. Patient outcomes with and without implementation of a neck pain classification system: a preliminary analysis. Orthopaedic Physical Therapy Practice 2018;30(2):82-90.
47. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, et al. 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 Manage 2011;41(6):1073-1093.
48. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs 2005;14(7):798-804.
49. Schellingerhout JM, Verhagen AP, Heymans MW, Koes BW, de Vet HC, Terwee CB. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. Qual Life Res 2012;21(4):659-670.
50. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. JAMA 2000;284(1):79-84.
51. Jull GA, Falla D, Vicenzino B, Hodges PW. The effect of therapeutic exercise on activation of the deep cervical flexor muscles in people with chronic neck pain. Man Ther 2009;14(6):696-701.
52. Maissan F, Pool JJM, Raaij de E, Mollema J, Ostelo RWJG, Wittink H. The clinical reasoning process in randomised clinical trials with patients with non-specific neck pain is incomplete: A systematic review. Musculoskeletal Science and Practice 2018;35:8-17.
53. Kelly J, Ritchie C, Sterling M. Clinical prediction rules for prognosis and treatment prescription in neck pain: A systematic review. Musculoskelet Sci Pract 2017;27:155-164.
54. Stanton TR, Hancock MJ, Maher CG, Koes BW. Critical appraisal of clinical prediction rules that aim to optimize treatment selection for musculoskeletal conditions. Phys Ther 2010;90(6):843-854.
55. Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. J Clin Epidemiol 2015;68(6):617-626.

56. van Giessen A, Peters J, Wilcher B, Hyde C, Moons C, de Wit A, et al. Systematic Review of Health Economic Impact Evaluations of Risk Prediction Models: Stop Developing, Start Evaluating. Value Health 2017;20(4):718-726.
57. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagn Progn Res 2018;2:11-018-0033-6.

Appendix 1: *Search strategies*

Pubmed
("Neck Pain"[Mesh] OR "neck pain"[tiab] OR neckache*[tiab] OR "neck ache"[tiab] OR "neck aches"[tiab] OR cervicodynia*[tiab] OR cervicalgia*[tiab])
AND
(algorithm*[Title/Abstract] OR "Algorithms"[Mesh] OR "Classification"[Mesh] OR "Prognosis"[Mesh:NoExp] OR "Physical Examination"[Mesh] OR "Clinical Decision-Making"[Mesh] OR clinical decision making[Title/Abstract] OR predict*[Title/Abstract] OR taxonom*[Title/Abstract] OR classif*[Title/Abstract] OR prognos*[Title/Abstract] OR subgroup*[Title/Abstract])
AND
("Treatment Outcome"[Mesh] OR "Rehabilitation"[Mesh] OR "Physical Therapy Modalities"[Mesh] OR physiotherap*[Title/Abstract] OR physical therap*[Title/Abstract] OR treatment[Title/Abstract] OR rehabilitation[Title/Abstract])
Cinahl
(MH "Neck Pain") OR "neck pain" OR neckache* OR "neck ache" OR "neck aches" OR cervicodynia OR cervicalgia)
AND
(MH "Physical Examination+") OR (MH "Physical Therapy Assessment") OR (MH "Predictive Value of Tests") OR (MH "Prognosis+") OR (MH "Classification") OR (MH "Decision Making, Clinical") OR MH "Algorithms" OR "clinical decision making" OR "physical examination" OR predict* OR taxonom* OR classif* OR prognos* OR subgroup* OR algorithm*)
AND
(MH "Physical Therapy+" OR MH "Treatment Outcomes+" MH "Rehabilitation+" OR physiotherap* OR physical therap* OR treatment OR rehabilitation)
Embase
('neck pain'/exp OR 'neck pain':ab,ti OR neckache*:ab,ti OR 'neck ache':ab,ti OR 'neck aches':ab,ti OR cervicodynia*:ab,ti OR cervicalgia*:ab,ti)
AND
('classification algorithm'/exp OR 'clinical decision making'/exp OR 'prognosis'/exp OR 'prognosis':ab,ti OR 'classification'/exp OR 'classification':ab,ti OR 'prediction'/exp OR 'prediction':ab,ti OR 'clinical decision making':ab,ti OR predict*:ab,ti OR taxonom*:ab,ti OR classif*:ab,ti OR prognos*:ab,ti OR subgroup*:ab,ti OR algorithm*:ab,ti)
AND
('rehabilitation'/exp OR 'physiotherapy'/exp OR 'treatment outcome'/exp OR 'physical therap*':ab,ti OR physioterap*:ab,ti OR treatment:ab,ti OR rehabilitation:ab,ti)
PEDro
Abstract and Title: Classif*
Problem: Pain
Body part: Head or neck
Grey literature databases

- **Dart Europe:** “neck” AND “classification”
- **Open access Theses and Dissertations:** “neck” AND “classification”
- **NDLTD:** “neck” AND classification”
- **Clinical trials.gov:** “neck” AND “classification”
- **WHO ICTRP:** “neck” AND “classification”

Appendix 2: *Summery of the critical appraisal scores for all papers*

	Cleland, 2007	de las Peñas 2011	Hanney, 2013	Puentedura, 2012	Raney, 2009	Saavedra-Hernandez, 2011	Bier, 2017	Childs, 2014	DeWitte, 2014	Fritz, 2007	Hefford, 2008	Lee, 2017	Wang, 2003
Purpose													
Is purpose, population and setting clearly specified?	y	y	y	y	y	y	y	y	y	y	y	y	y
Content validity													
(i) Is the domain and all specific exclusions from the domain clearly specified?	y	y	y	y	y	y	y	y	p	y	y	y	y
(ii) Are all relevant categories included?	y	y	y	y	y	y	y	y	y	y	y	y	y
(iii) Is the breakdown of categories appropriate, considering the purpose?	y	y	y	y	y	y	y	y	y	y	y	y	y
(iv) Are the categories mutually exclusive?	y	y	y	y	y	y	p	p	p	p	p	p	p
(v) Was the method of development appropriate?	y	y	y	y	n	y	y	y	y	y	y	y	p
(vi) If multiaxial, are criteria of content validity satisfied for each additional axis?	na	na	na	na	na	na	na	na	na	na	na	na	y
Face Validity													
(i) Is the nomenclature used to label the categories satisfactory?	y	y	y	y	y	y	y	y	y	y	y	y	y
(ii) Are the terms used based upon empirical (directly observable) evidence?	y	y	y	y	y	y	y	y	y	y	y	y	y
(iii) Are the criteria for determining inclusion into each category clearly specified?	y	y	y	y	y	y	y	y	y	y	y	p	y
(iv) If yes do these criteria appear reasonable?	p	p	y	y	y	p	y	n	y	n	y	n	n
(v) Have the criteria been demonstrated to have reliability or validity?	p	p	p	p	p	p	y	p	n	p	y	p	p
(vi) Are the definitions of criteria clearly specified?	y	y	y	y	y	y	y	n	y	y	y	y	y
(vi) If multiaxial are criteria of face validity satisfied for each additional axis?	na	na	na	na	na	na	na	na	na	na	na	na	p

Appendix 2: *Summery of the critical appraisal scores for all papers (continued)*

	Cleland, 2007	de las Peñas 2011	Hanney, 2013	Puentedura, 2012	Raney, 2009	Saavedra-Hernandez, 2011	Bier, 2017	Childs, 2014	DeWitte, 2014	Fritz, 2007	Hefford, 2008	Lee, 2017	Wang, 2003
Feasibility													
(i) Is the classification simple to understand?	y	y	y	y	y	y	y	y	y	y	y	y	y
(ii) Is the classification easy to perform?	p	y	y	y	y	y	y	y	y	y	y	y	y
(iii) Does it rely on clinical examination alone?	y	y	y	y	y	y	y	y	y	y	y	y	y
(iv) No special skills/tools or training required?	y	y	y	y	y	y	y	y	y	y	y	y	y
(v) How long does it take to perform?	y	y	y	y	y	y	y	y	y	y	y	y	y
Construct Validity													
(i) Does it discriminate between entities thought to be different in a way appropriate for the purpose?	y	y	y	y	y	y	y	y	y	y	y	dk	dk
(ii) Does it perform satisfactorily compared to other systems classifying the same domain?	dk	dk	dk	dk	dk	dk	dk	dk	dk	dk	dk	dk	dk
Reliability													
(i) Does the system provide consistent results when classifying the same conditions?	y	dk	dk	y	dk	dk	y	dk	dk	dk	dk	dk	dk
(ii) Is the intraobserver and interobserver reliability satisfactory?	p	dk	dk	p	dk	dk	p	dk	dk	y	y	dk	dk
Generalisability													
(i) Has it been used in other studies &/or settings?	y	n	n	n	n	n	y	n	n	y	y	n	n
TOTAL OVERALL SCORE	4.5	3.5	3.5	4	3.5	3.5	5	2.5	2.5	3.5	4.5	2.5	2.5

y = yes; n = no; p = partial; dk = don't know; na = non-applicable

4

Clinical reasoning in unimodal interventions in patients with non-specific neck pain in daily physiotherapy practice, a Delphi study

François Maissan
Jan Pool
Erik Stutterheim
Harriët Wittink
Raymond Ostelo

Musculoskeletal Science and Practice 37 (2018) 8–16

Abstract

Background: Neck pain is the fourth major cause of disability worldwide but sufficient evidence regarding treatment is not available. This study is a first exploratory attempt to gain insight into and consensus on the clinical reasoning of experts in patients with non-specific neck pain.

Objective: First, we aimed to inventory expert opinions regarding the indication for physiotherapy when, other than neck pain, no positive signs and symptoms and no positive diagnostic tests are present.
Secondly, we aimed to determine which measurement instruments are being used and when they are used to support and objectify the clinical reasoning process.
Finally, we wanted to establish consensus among experts regarding the use of uni-modal interventions in patients with non-specific neck pain, i.e. their sequential linear clinical reasoning.

Study design: A Delphi study.

Methods: A Web-based Delphi study was conducted. Fifteen experts (teachers and researchers) participated.

Results: Pain alone was deemed not be an indication for physiotherapy treatment. PROMs are mainly used for evaluative purposes and physical tests for diagnostic and evaluative purposes. Eighteen different variants of sequential linear clinical reasoning were investigated within our Delphi study. Only 6 out of 18 variants of sequential linear clinical reasoning reached more than 50% consensus.

Conclusion: Pain alone is not an indication for physiotherapy. Insight has been obtained into which measurement instruments are used and when they are used. Consensus about sequential linear lines of clinical reasoning was poor.

Keywords: Non-specific neck pain, Physiotherapy, Evidence based medicine,
Delphi study

Introduction

Neck pain is the fourth major cause of disability worldwide[1]. The prevalence for neck pain in the world was 4.9% in 2010. The total disability burden from musculoskeletal disorders (MSK) measured as Years Lived with Disability (YLDs), was 21.3% of which 20.1% as a result of neck pain [2]. Moreover, from 1990 to 2010, the disability burden attributable to MSK disorders increased by 46%. It is further expected that this burden will increase in the coming years.[2] Therefore, effective treatment methods are necessary. There is no conclusive evidence regarding specific pathology in the majority of cases of acute or chronic neck pain, therefore, most cases are labeled as nonspecific neck pain or neck pain of unknown origin, without further subdivision into subgroups [3].

The most frequently used interventions for the treatment of neck pain are exercises, manipulative therapies, mobilization, massage, and multidisciplinary biopsychosocial rehabilitation [4]. The evidence regarding the effectiveness of physiotherapy for neck pain is emerging [5-8]. However, sufficient evidence for application of a specific physiotherapy modality or therapy aiming at a specific patient subgroup is not available(9). The main recommendation in a review of physiotherapy interventions for patients with chronic neck pain was to identify relevant subgroups with matching treatments among patients with non-specific neck pain [9].

This matches the need in clinical practice for forms of "targeted treatment", or personalized treatment [10]. Personalized treatment is tailoring therapy to specific subgroups in order to optimize effectiveness [11]. There are indications in patients with non-specific lower back pain that subgrouping is effective and cost effective [12]. There is a need to do the same for patients with non-specific neck pain.

In clinical practice, physiotherapists first determine if physiotherapy is indicated for a patient. If so, they then, as a part of the clinical reasoning process, subgroup their patients aiming to match their treatment to the signs and symptoms and results of the diagnostic tests. Another important part of the clinical reasoning process is the use of measurement instruments. Measurement instruments, such as Patient Reported Outcome Measures (PROMs) and physical tests are used to support and objectify the clinical reasoning process. Which measurement instruments are most appropriate when they are used and how they support the clinical reasoning process in patients with non-specific neck pain is unclear.

The Hypothesis-Oriented Algorithm for Clinicians II (HOAC II)[13] provides an algorithm to describe the clinical reasoning process, and it combines the diagnostic process (the 'why') with the therapeutic process (the 'what'). This helps the physical therapist to decide "why" to do "what" as the "why" describes the specific diagnosed group within a population of patients with non-specific neck pain and the "what" describes the matched treatment. In addition, the HOAC II recommends matching outcome measures to the goals of treatment in order to evaluate the effectiveness of treatment. This way this linear clinical reasoning process consists of three sequential phases: the diagnostic, the therapeutic and, the evaluative phase. Following the HOAC II, sequential linear clinical reasoning in the present study is defined as the transition from signs and symptoms to diagnostic tests, from diagnostic tests to an intervention with matching treatment goal and evaluation based on outcome measurements related to the matched goals.

A recent review of the literature on the clinical reasoning process in research on patients with non-specific neck pain showed a lack of a complete clinical reasoning process with matching outcome measures. Only 11 (9%) out of 122 randomized controlled trials (RCT) described a complete clinical reasoning process whereby the diagnostic process ("the why"), i.e. other signs or symptoms or diagnostic tests in their inclusion criteria, could be linked to the therapeutic process ("the what") [14]. A remarkable outcome was that 46% of the 122 included RCT's described no impairment or activity limitation at all, with inclusion criteria limited to age and (duration of) pain. It can be questioned whether subjects having pain without any other signs/symptoms or positive diagnostic tests have an indication for physiotherapy treatment.

In conclusion, there is insufficient scientific evidence to form subgroups with matching, uni- or multimodal, interventions within patients with non-specific neck pain. Campbell et al [15] indicated that we first must understand working mechanisms of unimodal interventions before combining them into multimodal interventions. Therefore, it is sensible to first reach consensus on the various aspects of the clinical reasoning process when using unimodal interventions in patients with non-specific neck pain.

In this study we aimed to 1. describe expert opinion on the indication for physiotherapy when a patient's only problem is pain without other signs or symptoms or positive diagnostic tests 2. explore which measurement instruments are being used by experts to support their clinical reasoning process; when they are being used and for which purpose, 3. to establish consensus regarding the use of unimodal interventions, i.e. sequential linear clinical reasoning.

Methods

A Web-based Delphi study was conducted. The first round included questions to achieve the first two goals. In the second and third round an attempt was made to reach consensus regarding the use of unimodal interventions (to achieve the third goal).

To get expert opinion on the indication for physiotherapy when patients' only problem is pain without other signs or symptoms or positive diagnostic tests, we asked: "Suppose you have a patient with non-specific neck pain. Other than pain, there are no other signs or symptoms and no positive diagnostic test(s). There are no contra-indications for physiotherapy. Do you think there is an indication for physiotherapy?" Secondly, we wanted information on the type of measurement instruments being used for diagnostic and/or evaluative purposes. For this purpose, we offered a list with the most frequently used measurement instruments in patients with non-specific neck pain selected from the 122 RCTs included in the review of Maissan et al.[14]. We asked the experts which measurement instruments they use and whether they use additional measurement instruments. We also gathered information about the timing of the evaluative tests; i.e. only by start and finish of the treatment or also during the treatment. We made a distinction between patients with acute/sub-acute and chronic neck pain. In this way, we assessed whether the duration of the presence of pain influences the timing of measurements. Regarding the diagnostic process we assessed the extent to which physical impairments were pragmatically diagnosed (i.e. a test developed in their own practice without evidence of the psychometric properties) or with valid tests (with known psychometric properties). If a valid test was used, we asked them to specify the test.

The Delphi method is appropriate to reach consensus in a field where a lack of agreement or incomplete knowledge is evident [16]. The Delphi technique is a widely used and accepted method for achieving convergence of opinion concerning real-world knowledge solicited from experts within certain areas of interest [17]. Therefore, the Delphi method creates the opportunity to gather information from a group of international experts in treating patients with non-specific neck pain, without the need of a meeting [18]. In this method, experts independently and anonymously answer a range of questions. During several rounds these experts get insight into group opinions and have the possibility to reconsider their own opinion as the results of the earlier rounds are returned until they achieve consensus [19]. The Delphi study consisted of three rounds as described by Hsu et al [17]. For the Delphi study "Formdesk" software was used and invitations to participate were sent by email.

Preparation phase

For this exploratory Delphi study a convenience sample of twenty-four experts from the expert network of the second author were invited to participate. The expert group consisted of teachers and researchers in the field of Orthopedic Manipulative Therapy **(Table 1)** and, were all experts in treatment of patients with non-specific neck pain. Most experts were members of the standard committee of wherein the second author (JP) also participated. **Table 1** shows their current main job and other characteristics. Although it was not part of the inclusion, all participants also met the four criteria set by Jensen et al. for being an expert in physiotherapy. These criteria are: knowledge, clinical reasoning skills, examination and evaluation skills of movement and virtues [20]. We chose experts because we assumed that they were most likely to reach consensus based on their knowledge from scientific research and the ability to translate this knowledge into practice. By completing the questionnaire of round 1, the participants confirmed their participation in this study.

Procedure

The workgroup of this Delphi study consisted of the first two authors (FM, JP) who designed the three rounds and summarized the returned data. We considered more than 50% consensus in responses as the consensus cut-off point [16].

First round Delphi

The first round addressed a linear clinical reasoning in physiotherapy based on the HOAC II. It consisted of open-ended questions starting with questions about the use of measurement instruments. As a starting point, the most frequently used measurement instruments in patients with non-specific neck pain were presented to the experts. We asked the experts for which signs and/or symptoms they would use a specific diagnostic test to determine the hypothesized cause of the patient experienced problem. Then we asked about the relationship between the diagnostic test (cause) and a chosen intervention. In other words, which diagnostic tests lead to which specific intervention. Finally, we wanted to determine which outcome measures the experts use to evaluate the effect of a specific intervention. In this way, we aimed to get an overview of the match between history taking (signs and symptoms), physical examination (diagnostic test), intervention, and the use of evaluative outcome measures. According to HOAC II this sequence describes the entire linear clinical reasoning process of the physical therapist [13]. In addition to the clinical reasoning process, we inventoried which interventions the experts use regularly. Interventions used by 3 or less experts were not considered “regular treatment” in patients with non-specific neck pain.

Second round Delphi

In the second round each expert was asked to review and reflect on the items summarized by the investigators based on the information provided in round one[17]. The experts were asked to rate the importance of the signs and symptoms in relation to the given diagnostic test, to rate the importance of the diagnostic test in relation to the given intervention and finally, rate the importance of the outcome measurement instruments in relation to the given intervention. The rating scale ranged from 1 (very important) to 5 (not important). This way preliminary priorities among items were established. In the second round the experts were also asked which physical or mental function or activity of the patient they wanted to improve (goal of the intervention) in relation to the chosen interventions.

Third round Delphi

The third round consisted of the summarized items and ratings of the importance of the previous round. In this final round, the complete sequence of a linear clinical reasoning process was presented in a table for each intervention. Each line in **Table 5** represents such a sequence. For each included diagnostic test a different sequence of linear clinical reasoning was added to the questionnaire. The reason for this approach was that multiple diagnostic tests could lead to the same intervention, however, with possible different goals and evaluative measurement instruments. The signs and symptoms with more than 50% consensus on the score "very important" or "important" were combined because multiple signs and symptoms could lead to one diagnostic test. When there was no consensus after round two on "signs and symptoms" and "direct goal of intervention" and "evaluation test", summarized results of round two were offered as final choice options in round three per sequence of linear clinical reasoning.

Results

Round 1 of the Delphi study began with 15 participants. There were 4 drop outs in round 2. The drop outs in round 2 were due to technical problems logging in the system. Data of these four experts were only partly restored. In round 3 fourteen experts participated. There was only one drop out. Upon inquiry no reason could be determined.

Table 1: *Participant characteristics (n=15)*

Gender	9 female
Age; mean (range)	49.9 (39-65)
Nationality	Canada 3, Australia 3, New Zealand 2, USA 1, the Netherlands 1, Belgium 1, Portugal 1, South Africa 1, Denmark 1, Spain 1
Highest level of education	Phd 10, Msc 5
Current main job functions	Private practice 6, education 10, research 11, consultant 1
Years of experience in physiotherapy practice; mean (range)	20.3 (6-40)
Present-day work time as a physical therapist in hours a week; mean (range)	9.9 (1-35)

Table 1 describes the characteristics of the participants. All experts considered physiotherapy treatment not indicated in patients with non-specific neck pain without any positive signs and/or symptoms or diagnostic tests. However, six out of fourteen (43%) experts named one possible treatment, namely pain education.

Table 2 and 3 show which measurement instruments the experts use in daily practice. Table 2 shows which PROMs the experts use in their daily practice and when the experts use the PROMs in their clinical reasoning process. The use is explicitly expressed as either diagnostic or evaluative use or both. **Table 3** shows the use of measurement instruments to measure physical constructs (physical tests). The timing “regular during treatment” of the use of these physical tests, to guide the intervention, was higher (81%) than the use of PROMs (39%) for patients with acute/sub-acute non-specific neck pain than in the group of patients with chronic non-specific neck pain namely, 73% and 30%. The PROMs were mainly used at the beginning and end of the treatment, except for pain measurements and the Neck Disability Index (NDI), which were also considered to be able to guide an intervention.

Table 2: *Use of Patient Reported Outcome Measures (n =15)*

				Acute-sub acute neck pain		Chronic neck pain	
Patient Reported Outcome Measure	**Total use (n)**	**Diagnotic use (n)**	**Evaluative use (n)**	**Start and finish (n)**	**Regular during**	**Start and finish (n)**	**Regular during**
Visual analogue pain scale	8	6	7	7	5	6	6
Numeric pain rating scale	10	5	9	9	8	9	3
Neck disability index	14	5	13	13	13	14	3
Short form 12	4	0	3	3	0	3	0
Short form 36	4	0	4	4	0	4	0
Patient perceived effect	2	1	3	1	1	2	3
Patient specific function scale	6	1	6	6	3	6	0
Neck pain and disability scale	1	0	1	1	0	1	0
Tampa scale of Kinesiophobia	6	4	4	4	0	4	0
Fear avoidance beliefs questionnaire	6	3	5	5	1	5	1
Borg scale	3	0	3	6	2	7	1
McGill pain questionnaire	7	2	5	8	4	9	4
Headache disability index	7	0	7	9	4	8	4
Impact of event scale (IES)	6	2	4	6	0	6	0
Pain catastrophizing scale (PCS)	11	2	9	9	0	9	1
Dizziness handicap inventory (DHI)	7	0	7	6	4	6	4
Illness perception questionnaire (IPQ)	8	2	6	7	0	7	1
Pain self-efficacy questionnaire (PSEQ)	9	3	6	10	0	10	1
Whiplash disability questionnaire (WDQ)	5	0	5	6	2	6	2
Leeds Assessment of Neuropathic Symptoms and Signs pain scale (SLANSS)	5	4	5	5	2	5	5

4

Table 3: *Use of measurement instruments to measure physical constructs (n = 15)*

				Acute-sub acute neck pain		Chronic neck pain	
Constructs	**Total use (n)**	**Diagnotic use (n)**	**Evaluative use (n)**	**Start and finish (n)**	**Regular during treament**	**Start and finish (n)**	**Regular during treatment**
Range of motion	14	13	13	10	13	10	10
Muscle force	9	8	4	4	4	4	3
Pressure pain threshold	5	3	5	4	2	4	2
Electromyography	1	0	1	1	0	1	0
Pinch force	2	2	2	2	0	0	0
Joint play	10	10	7	5	7	4	5
Joint mobility	15	15	10	10	7	8	7
Endurance	15	5	4	4	4	3	3
Coordination	5	5	5	5	4	4	4
Active stability	8	7	7	7	4	7	3
Passive instability	10	10	4	4	2	4	1
Upper cervical instability	12	12	3	3	2	3	1
Reflexes	13	13	7	5	5	7	5
Sensibility	13	13	8	7	5	7	5
Neurodynamics	13	13	9	9	6	9	6

Table 4: *Measurement instruments to measure physical constructs (n = 15)*

Constructs	Pragmatic test (n)	Valid test (n)	Do not use (n)	Name of valid test or measurement instrument
Range of motion	6	5		Cervical range of motion (CROM) device, Goniometer, Inclinometer
Muscle force	5	3	3	Dynamometer; Medical Research Council (MRC) Scale for Muscle Strength
Pressure pain threshold	2	4	5	(digital) Pressure algometer
Pinch force	2	1	8	Jamar pinch force device
Joint play	11			
Joint mobility	10	1	1	Cervical range of motion (CROM) device
Endurance	5	4	2	Cranio-cervical flexion test (CCFT) , Neck extensor test
Coordination	9	1	1	Joint position sense (JPS)
Active stability	7		4	
Passive stability	8		3	
Upper cervical instability	8	4		UC stability tests, sharp purser
Neurodynamics	6	5		ULTT, SLUMP, SLR
Tenderness	8	1	2	Quantitative sensory testing (QST)
Postural control	8		3	
Proprioception	5	3	3	Joint position sense (JPS)
Motor control	5	4	2	Cranio-cervical flexion test (CCFT)/stabilizer, the Fly
Movement pattern	8		3	
Muscle spasm	8		3	
Muscle length	9			
Muscle tension	8		3	

Table 5: *Consensus in sequential linear clinical reasoning (n =14)*

sign and/or symptom => =>	diagnostic test => =>	intervention => =>	Direct goal of intervention => =>	Evaluation test	Consensus % (J/N)
Muscle weakness in history	Strength test	Strength exercises	increase strength	Endurance test	50% (N)
Reduction of movement	ROM	Stretching	increase ROM	ROM	64% (J)
Pain, reduced ROM, presence of stiffness	Joint mobility assessment	Traction	reduce pain	Pain VAS/NPRS	29% (N)
Pain	Pain VAS/NPRS	Massage	reduce muscle tension	Palpation of tenderness	36% (N)
Muscle tension	Muscle tension/ spasm test	Massage	Improve muscle tissue mobility	Palpation of tenderness	43% (N)
Pain	TP palpation	Dry needling	Reduce pain	TP palpation	43% (N)
Fear of movement	FABQ	Relaxation therapy	Reduce anxiety	Pain VAS/NPRS	50% (N)
Fear of movement	Coping test	Relaxation therapy	reduce anxity	Pain VAS/NPRS	57% (J)
Fear of movement	Stress test	Relaxation therapy	reduce anxity	Pain VAS/NPRS	43% (N)
Decrease maladaptive cognitions and behaviors	Central sensitization test	Pain education	decrease maladaptive cognitions and behaviors	NDI (neck disability index)	62% (J)

Movement dysfunction, reduced ROM, presence of stiffness	ROM	Mobilization	improve quality of movement	ROM	36% (N)
Movement dysfunction, reduced ROM, presence of stiffness	Joint mobility assessment	Mobilization	Improve quality of movement	Joint mobility assessment	43% (N)
Movement dysfunction, reduced ROM, presence of stiffness	End feel	Mobilization	Improve quality of movement	Joint mobility assessment	43% (N)
Reduced control/ endurance, symptoms of motor control impairment	Motor control test	Endurance exercises	increase endurance	Neck endurance test	79% (J)
Postural dysfunction, symptoms of imbalance	Postural control test	Endurance exercises	increase postural cotrol	Neck endurance test	50% (N)
Weakness, postural dysfunction, coordination difficulties	Active stability test	Stabilization exercises	Improve movement control	Active stability tests	36% (N)
Weakness, postural dysfunction, coordination difficulties	Proprioception test/CCFT	Coordination exercises	Improve motor control	Proprioceptive tests/CCFT	79% (J)
Dizziness, Unsteadiness, coordination difficulties, symptoms of motor control impairment	Movement pattern assessment	Coordination exercises	Improve motor control	Movement pattern assessment	79% (J)

4

Table 4 shows whether a construct was measured in a pragmatic or valid manner. Multiple valid measurement instruments were named, however, most experts measured physical constructs pragmatically.

The interventions used by less than 3 experts were: transcutaneous electrical nerve stimulation, electro thermal therapy, low level laser, Ultra Sound and taping. These interventions were excluded in this study. **Table 5** describes the degree of consensus among the participating experts concerning the remaining interventions. The sequential linear clinical reasoning process is divided into 5 steps: - signs and/or symptoms, - diagnostic test, - intervention, - direct goal of the intervention and, - evaluation test. For example: "movement dysfunction/presence of stiffness" as sign and/or symptom leads to the diagnostic test for "range of movement" (ROM), which leads to the intervention "mobilization" which leads to the direct goal of "improve quality of movement", which leads to the evaluation test "ROM". The last column represents the degree of consensus of that specific sequential line of linear clinical reasoning. Only 6 out of 18 lines of sequential linear clinical reasoning reached more than 50% consensus. In addition to the consensus sequence as shown in table 5, an overview of all other additional information given by the experts in round 3 is shown in **Appendix 1**.

Discussion

The experts state that pain alone is not an indication for physiotherapy, and that there must be other signs/symptoms present and/or at least one or more positive diagnostic test to substantiate the indication for physiotherapy. However, the review of Maissan et al [14]. reported that only 16% of the 122 randomized controlled trials (RCTs) had a diagnosed cause, i.e. at least one diagnostic test was used as an inclusion criterion. Therefore, one could argue that RCTs, to examine the effect of a physiotherapy intervention, were regularly conducted without first determining whether or not there was an indication for physiotherapy, and if this was done, then this was not explicitly described in those RTCs. This discrepancy between what experts deemed to be important and the absence, or lack of a clear description, of an indication for physiotherapy underlines the need to apply recognisable clinical reasoning within the methodology of RCTs to enhance transferability to daily practice, especially the translation of the diagnostic process into the in/exclusion criteria for subject recruitment.

Our Delphi illustrates that experts used a wide range of measurement instruments. The most used PROMs were pain questionnaires (Visual Analogue Scale pain and Numeric Pain Rating Scale) and a questionnaire for physical functioning (NDI). Also PROMs regarding psychosocial topics, like catastrophizing or illness perceptions, were used. This indicates that the experts measure physical functions as well as mental functions. However, it is notable how rarely questionnaires about mental functioning are used for diagnostic purposes. For example, the Tampa scale of Kinesiophobia is used by 4 experts, Fear avoidance beliefs questionnaire by 3 and Impact of event scale (IES) by 2. On the other hand, the Visual analogue pain scale is used by 6 experts and the NDI by 5. This finding corresponds with other scientific research [21,22]. Indeed, for diagnostic purposes the experts use measurement instruments for constructs representing physical functioning 4 times more than PROMs.

We acknowledge that physiotherapists use multiple interventions within one treatment session i.e. complex or multimodal interventions [15] and that focusing on unimodal interventions is a simplification of clinical practice. To illustrate, one expert commented: “clinical practice is fluid and an intervention is not delivered in isolation” thereby underlining the fact that daily practice is more complex. However, this study shows that achieving consensus on unimodal clinical reasoning proved to be difficult enough. Despite the lenient boundary of consensus at > 50%, only 6 out of 16 linear lines of clinical reasoning reached consensus of > 50%. After round 1 the most frequently used interventions were further explored. After round two and three inconsistency was mainly on the items “Direct goal of the intervention” and “Evaluation test”. In hindsight, this was to be expected as interventions can pursue different goals with different evaluative outcome measures. An explanation for the inconsistency among expert regarding the item “Evaluation test” could be that different experts may have different preferences for certain outcome measures.

In addition to the aforementioned items “Direct goal of the intervention” and “Evaluation test”, there was also inconsistency on the other two items “sign and/or symptom” and “diagnostic test”. A last explanation for overall inconsistency could be the lack of translatability of the results from scientific research [23]. We hypothesized that if scientific research includes a clear clinical reasoning process, it should be easier to translate this reasoning into daily practice, especially by experts. In our study we found that in interventions to improve motor control or endurance consensus was reached. This can be explained when we look in more detail at a review of Maissan et al. This review showed that 4 RCTs to improve motor control [24,25] or endurance [26,27] were part of the 11 out of 122 RCTs with a complete clinical reasoning process. In other words, these RCTs included a diagnostic criterion like potential impairments, limitations in activities or restrictions in participation, to get a sub-group that matched with the unimodal intervention. Hence, it looks as if these research findings were easy to translate to clinical practice. Therefore, this seems to confirm the need to perform future research with a more clinically relevant focus.

A limitation of the study was the technical problems in round two; because of this technical problem some data have been missed. However, because in round three all the experts of round one had the opportunity to participate and the opportunity for additional suggestions, the loss of information was minimal. A second possible limitation is the relatively small number of experts. However, they were consistent in their answers. Still, this was an exploratory study and further research is needed to confirm our findings.

A strength of our study was the diversity of nationalities **(Table1)** which ensured that the results of clinical reasoning are transboundary and not biased by habits of a single country. We also consider a strength the use of a transparent framework for clinical reasoning, namely the HOAC.

We already stated that future research must organize itself in a more clinically relevant manner, thus include a diagnostic process prior to or part of inclusion in a RCT. Per intervention, experts appointed signs/symptoms and diagnostic tests **(Table 5)** which can be the basis of a diagnostic process leading to a matching intervention which then can be incorporated into daily practice or into clinical trials. We do want to emphasize the importance of evaluating the direct effect of the intervention in addition to patient experienced effects in order to determine if there is a (causal) relation between them, both in scientific research and in daily practice. To do so, measurement instruments to "diagnose" physical or mental constructs should also be used in the evaluative process to determine if the physical or mental construct under treatment has improved. However, appendix 1 shows, in addition to a wide variety of measurement instruments or diagnostic tests, also an inconsistency in the use of the same measurement instruments before and after the intervention. Only by using the same measurement instrument before and after the intervention more clarity can be obtained if the perceived effect by the patient is due because, or in spite of the intervention.

In conclusion: Pain alone, without other signs/symptoms present and/or at least one or more positive diagnostic test does not substantiate the indication for physiotherapy. Insight has been obtained into which measurement instruments are used and when they are used. Consensus about sequential linear lines of clinical reasoning was poor.

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. 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 2012;380(9859):2163-2196.
2. March L, Smith EU, Hoy DG, Cross MJ, Sanchez-Riera L, Blyth F, et al. Burden of disability due to musculoskeletal (MSK) disorders. Best Pract Res Clin Rheumatol 2014;28(3):353-366.
3. Bogduk N BL. Back Pain and Neck Pain: an evidence-based update. In: Devor M, Rowbotham MC, Wiesenfeld-Hallin Z, eds., editor. Progress in Pain Research and Management. 16th ed. Seattle: IASP Press; 2000; 371-7.
4. Pool JJ, Ostelo RW, Knol D, Bouter LM, de Vet HC. Are psychological factors prognostic indicators of outcome in patients with sub-acute neck pain? Man Ther 2010;15(1):111-116.
5. Bertozzi L, Gardenghi I, Turoni F, Villafane JH, Capra F, Guccione AA, et al. Effect of therapeutic exercise on pain and disability in the management of chronic nonspecific neck pain: systematic review and meta-analysis of randomized trials. Phys Ther 2013;93(8):1026-1036.
6. Cheng YH, Huang GC. Efficacy of massage therapy on pain and dysfunction in patients with neck pain: a systematic review and meta-analysis. Evid Based Complement Alternat Med 2014;2014:204360.
7. Gross A, Kay TM, Paquin JP, Blanchette S, Lalonde P, Christie T, et al. Exercises for mechanical neck disorders. Cochrane Database Syst Rev 2015;1:CD004250.
8. Monticone M, Cedraschi C, Ambrosini E, Rocca B, Fiorentini R, Restelli M, et al. Cognitive-behavioural treatment for subacute and chronic neck pain. Cochrane Database Syst Rev 2015;5:CD010664.
9. Damgaard P, Bartels EM, Ris I, Christensen R, Juul-Kristensen B. Evidence of Physiotherapy Interventions for Patients with Chronic Neck Pain: A Systematic Review of Randomised Controlled Trials. ISRN Pain 2013;2013:567175.
10. Godman B, Finlayson AE, Cheema PK, Zebedin-Brandl E, Gutierrez-Ibarluzea I, Jones J, et al. Personalizing health care: feasibility and future implications. BMC Med 2013;11:179-7015-11-179.
11. Bates S. Progress towards personalized medicine. Drug Discov Today 2010;15(3-4):115-120.
12. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet 2011;378(9802):1560-1571.
13. othstein JM, Echternach JL, Riddle DL. The Hypothesis-Oriented Algorithm for Clinicians II (HOAC II): a guide for patient management. Phys Ther 2003;83(5):455-470.
14. Maissan F, Pool JJM, Raaij de E, Mollema J, Ostelo RWJG, Wittink H. The clinical reasoning process in randomised clinical trials with patients with non-specific neck pain is incomplete:

4

A systematic review. Musculoskeletal Science and Practice 2018;35:8-17.
15. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000;321(7262):694-696.
16. Giannarou L, Zervas E. Using Delphi technique to build consensus in practice. Int. journal ofbusiness science and applied management 2014;9(2):66-82.
17. Hsu C, Sandford BA. The Delphi technique: Making sense of consensus. Practical assessment, research & evaluation 2007;12(10):1-7.
18. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, et al. Consensus development methods, and their use in clinical guideline development. Health Technol Assess 1998;2(3):i-iv, 1-88.
19. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000;32(4):1008-1015.
20. Jensen GM, Gwyer J, Shepard KF, Hack LM. Expert practice in physical therapy. Phys Ther 2000;80(1):28-43; discussion 44-52.
21. Synnott A, O'Keeffe M, Bunzli S, Dankaerts W, O'Sullivan P, O'Sullivan K. Physiotherapists may stigmatise or feel unprepared to treat people with low back pain and psychosocial factors that influence recovery: a systematic review. J Physiother 2015;61(2):68-76.
22. Emilson C, Asenlof P, Pettersson S, Bergman S, Sandborgh M, Martin C, et al. Physical therapists' assessments, analyses and use of behavior change techniques in initial consultations on musculoskeletal pain: direct observations in primary health care. BMC Musculoskelet Disord 2016;17:316-016-1173-x.
23. Maher CG, Sherrington C, Elkins M, Herbert RD, Moseley AM. Challenges for evidence-based physical therapy: accessing and interpreting high-quality evidence on therapy. Phys Ther 2004;84(7):644-654.
24. Jull G, Falla D, Treleaven J, Hodges P, Vicenzino B. Retraining cervical joint position sense: the effect of two exercise regimes. J Orthop Res 2007;25(3):404-412.
25. Beinert K, Taube W. The effect of balance training on cervical sensorimotor function and neck pain. J Mot Behav 2013;45(3):271-278.
26. Jull GA, Falla D, Vicenzino B, Hodges PW. The effect of therapeutic exercise on activation of the deep cervical flexor muscles in people with chronic neck pain. Man Ther 2009;14(6):696-701.
27. Beer A, Treleaven J, Jull G. Can a functional postural exercise improve performance in the cranio-cervical flexion test?--a preliminary study. Man Ther 2012;17(3):219-224.

Appendix 1: *Additional information besides the highest consensus in clinical reasoning* **(in bolt letters)**

sign and/or symptom => =>	diagnostic test => =>	intervention => =>	Direct goal of intervention	Evaluation test	Consensus % (J/N)
Muscle weakness in history - Chronic pain	**Strength test** - Plus a functional test - Muscle control test	**Strength exercises** - Muscle control task	**Increase strength** - Control pain	**Endurance test** - Functional test that integrates strength - Reassess test movement, in combination with subjective findings and activity levels of patient - Strength test - Strength test or endurance test depending on patient	**50% (N)**
Reduction of movement - Reduction in movement means what? Quality of ROM; Pain at end of range…	**ROM**	***Stretching*** - Combined intervention: muscle length, pain inhibition, specific joint motion, willingness to move - Mobilization with Movement; PNF	**Increase ROM** - Reduce pain - Increase muscle length - Achieve meaningful functional task	**ROM** - Meaningful immediate task if applicable followed by meaningful larger functional task - Main parameter out of physical and subjective examinations	**64% (J)**
Pain, reduced ROM, presence of stiffness - Extremity Symptoms Centralized with Traction - Neuropathic pain, paresthesia or numbness	**Joint mobility assessment** - Plus meaningful functional task - Test Item Cluster for Cervical Radiculopathy - NPRS	**Traction** - Mobilization/ manipulation	**Reduce pain** - Reduce neuropathic pain - Increase strength - Is pain dominant or not? If dominant: first short term goal will be pain reduction	**Pain VAS/NPRS**	**29% (N)**
Pain - Pain on palpation	**Pain VAS/NPRS** - Movement quality with pain - NPRS-surrogate measure in absence of reliable measure of tension - Pain provoking movement	**Massage**	**Reduce muscle tension** - Improve muscle tissue mobility - Initiation of movement - pain relief	**Palpation of tenderness** - Movement quality with pain - Pain provoking movement - NPRS	**36% (N)**

Appendix 1: *Additional information besides the highest consensus in clinical reasoning* **(in bolt letters)** *(continued)*

sign and/or symptom => =>	diagnostic test => =>	intervention => =>	Direct goal of intervention	Evaluation test	Consensus % (J/N)
Muscle tension - Pain - Limited ROM - Pale skin - Loss of specific functional movement	**Muscle tension/ spasm test**	**Massage**	**Improve muscle tissue mobility** - Initialization of movement - Reduce muscle tension - Improved quality of movement including initialization of movement	**Palpation of tenderness** - Motor control task	**43% (N)**
Pain - Muscle tension - Decreased ROM - Pain at rest and with movement - Local and referred pain changes in skin color, texture and mobility - Reduce muscle mobility pain linked to movement restriction and movement incoordination - It is a combination of interacting symptoms and signs.	**TP palpation** - Reduces ROM - NPRS	**Dry needling**	**Reduce pain** - Relieve Triggerpoint - Reduce myofascial tightness	**TP palpation** - NPRS	**43% (N)**
Fear of movement - Poor symptom management - Muscle tension	**FABQ** - Self efficacy questionnaire - Checklist of non-verbal pain indicator (CNPI)	**Relaxation therapy** - Graded exercises - Why relaxation therapy if someone is afraid to move? Graded exposure	**Reduce anxiety** - Reduce pain - Reduce muscle tone - Reduce fear of re-injury - Improve self-efficacy	**Pain VAS/NPRS** - Self efficacy questionnaire - Brief Pain Inventory and CNPI - Ability to perform fearful movements	**50% (N)**
Fear of movement - Anxiety, inability to relax - Poor symptom management	**Coping test** - NPRS - Depends on combination of symptoms	**Relaxation therapy**	**Reduce anxiety** - Reduce pain - Reduce muscle tone - Self management of symptoms	**Pain VAS/NPRS** - Brief Pain Inventory - Coping test, ability to relax - FABQ - Movement Test	**57% (J)**

Fear of movement - Anxiety - Inability to relax	**Stress test** - Coping test - DASS - Functional movement test	**Relaxation therapy** - Plus movement retraining - Relaxation-Interdisciplinary care with psychologist	**Reduce anxiety** - Reduce pain - Reduce muscle tone - Response to stress via movement	**Pain VAS/NPRS** - Self-evaluation of identification of environmental stressors - DASS-monitored by psychologist - Coping test, ability to relax - FABQ - Functional movement tests	**43% (N)**
Decrease maladaptive cognitions and behaviors - Decrease fear avoidance and wrong beliefs - Decreased self-efficacy	**Central sensitization test** - Central sensitization inventory - Not always formally tested. Depends entirely on the construct	**Pain education** - Plus graded exposure	**Decrease maladaptive cognitions and behaviors** - Decrease fear avoidance and wrong beliefs - Decreased self-efficacy - and increase function and decrease pain - Reframe belief or behavior into something positive	**NDI (neck disability index)** - Questionnaires on knowledge and beliefs - CSI - FABQ - Pain catastrophizing test	**62% (J)**
Movement dysfunction, reduced ROM, presence of stiffness	**ROM**	**Mobilization**	**Improve quality of movement** - Increase ROM - Relaxation - Reduce pain	**ROM** - and pain response	**36% (N)**
Movement dysfunction, reduced ROM, presence of stiffness - Pain	**Joint mobility assessment** - NPRS	**Mobilization**	**Improve quality of movement** - Increase ROM - Reduce pain - Relaxation	**Joint mobility assessment**	**43% (N)**
Movement dysfunction, reduced ROM, presence of stiffness	**End feel**	**Mobilization**	**Improve quality of movement** - Increase ROM - Reduce pain - Relaxation	**mobility assessment 6x** - ROM - End feel - VAS/NPRS - Movement coordination test and functional tasks	**43% (N)**

Appendix 1: *Additional information besides the highest consensus in clinical reasoning* **(in bolt letters)** *(continued)*

sign and/or symptom => =>	diagnostic test => =>	intervention => =>	Direct goal of intervention	Evaluation test	Consensus % (J/N)
Reduced control/ endurance, symptoms of motor control impairment	**Motor control test** - Endurance test	**Endurance exercises**	**Increase endurance** - Increase postural control	**Neck endurance test**	**79% (J)**
Postural dysfunction, symptoms of imbalance - Imbalance in posture, muscular imbalance - Pain with sustained activities	**Postural control test** - NPRS	**Endurance exercises**	**Increase postural control** - Increase endurance - Control pain	***Neck endurance test*** - NPRS	**50% (N)**
Weakness, postural dysfunction, coordination difficulties	**Active stability test**	**Stabilization exercises**	**Improve movement control** - Improve active stability - Reduce pain - Decrease the incidence of recurrence	**Active stability tests**	**36% (N)**
Weakness, postural dysfunction, coordination difficulties	**Proprioception test/ CCFT**	**Coordination exercises**	**Improve motor control** - Reduce pain	**Proprioceptive tests/CCFT**	**79% (J)**
Dizziness, Unsteadiness, coordination difficulties, symptoms of motor control impairment - Reasons for dizziness, unsteadiness? Cervical, general proprioceptive, vascular? - Neurological weakness	**Movement pattern assessment** - Tests of motor control, movement coordination - Dizziness Handicap Inventory or Balance Test	**Coordination exercises**	**Improve motor control** - Reduce pain	**Movement pattern assessment** - Dizziness Handicap Inventory or Balance Test	**79% (J)**

5

Completeness of the description of manipulation and mobilisation techniques in randomized controlled trials in neck pain; a review using the TiDieR checklist

Jan Pool
François Maissan
Nanna de Waele
Harriët Wittink
Raymond Ostelo

Musculoskeletal Science and Practice 45 (2020) 102098

Abstract

Study design: A secondary analysis of a systematic review

Background: Manipulations or mobilizations are commonly used interventions in patients with mechanical neck pain. The treatment effects have often been studied in randomized controlled trials (RCT) which are generally considered the gold standard in evaluating the treatment effects, mainly due to its high internal validity. External validity is defined as the extent to which the effects can be generalised to clinical practice. An important prerequisite for this is that interventions used in clinical trials can be replicated in clinical practice. It can be questioned if interventions utilized in randomized controlled trials can be translated into clinical practice.

Objectives: The overall aim of this study is to examine whether the quality of the description of manipulation and mobilization interventions is sufficient for to replication of these interventions in clinical practice.

Methods: A comprehensive literature search was performed. Two independent researchers used the Template for Intervention Description and Replication (TIDieR) which is a 12-item checklist for describing the completeness of the interventions.

Results: Sixty-seven articles were included that used manipulation and/ or mobilization interventions for patients with mechanical neck pain. None of the articles describe the intervention e.g. all the items on the TIDieR list. Considering item 8 (a-f) of the TIDieR checklist only one article described the used techniques completely.

Conclusion: Manipulation or a mobilization interventions are poorly reported in RCTs, which jeopardize the external validity of RCTs, making it difficult for clinicians and researchers to replicate these interventions.

Keywords: Randomized controlled trial; Mobilization; Spinal manipulation:
TiDieR checklist

Introduction

The randomized controlled trial (RCT) is generally considered the gold standard in evaluating the effects of treatment [1]. Internal validity of these studies is of importance as this determines the level of confidence for making treatment recommendations because the conclusions of a specific trial are then valid for the population of interest. Furthermore, studies must also be of sufficient external validity to allow for generalizability and replication of the interventions in clinical practice [2]. External validity has been defined in many different ways in the literature [2,3,4]. Rothwell defined it as the extent to which the results of a trial are relevant to clinical practice, among other things, the extent to which the intervention is likely to be replicated when applied to patients in a particular clinical setting [3].

One of the challenges identified in the reporting of clinical trials is the quality of the description of the intervention [5]. Providing sufficient details about interventions is fundamental in the scientific process and is critical for the development of evidence informed practice [6]. As Hoffman et al stated; “Without a complete published description of interventions, clinicians cannot reliably implement interventions that are shown to be useful, and other researchers cannot replicate or build on research findings”.

The CONSORT statement for RCTs recommends precise specification of trial processes including details of the intervention being studied or components of that intervention [7]. Despite this recommendation, health care providers in daily practice are not provided with a complete description of the intervention in most RCTs. Glasziou et al demonstrated that in back pain trials, only 13 % of the interventions could be replicated [8]. Given the importance of adequate reporting of interventions in clinical trials, the Template for Intervention Description and Replication (TIDieR) was developed by Hoffman et al. [5]. This template was developed to guide the complete reporting of an intervention and is an extension to the CONSORT 2010 statement. The TiDieR checklist was published in 2014 as an official extension of the Consolidated Standards of Reporting Trials (CONSORT) 2010 statement. The CONSORT statement suggests that authors should report on “The interventions with sufficient details to allow replication”.

In this article we consider interventions used in patients with non-specific neck pain. Neck pain is the fourth major cause of disability worldwide. In 2015, more than a third of a billion people worldwide had neck pain of more than 3 months duration [9], which makes neck pain a serious health threat.

The interventions that are used for the treatment of neck pain are exercises, manipulative therapies, mobilization, massage, physical methods, and multidisciplinary biopsychosocial rehabilitation, or a combination of these modalities.The most frequently used physiotherapeutic interventions in patients with non-specific neck pain are manipulations or mobilizations [10].

The TIDieR checklist assesses all the relevant issues related to an intervention, such as for example why the intervention was performed, by whom and where. Item 8 of the TIDieR checklist focuses specificly on the used techniques, such as the 'segmental level', 'frequency','direction', 'intensity', 'dosage'. In this article we focused on both, that is, all the relevant issues related to the intervention, as well as the specific manipulation and mobilization techniques and the replication of these techniques.

A manipulation technique is defined as: A passive, high velocity, low amplitude thrust applied to a joint complex within its anatomical limit with the intent to restore optimal motion, function, and/ or to reduce pain. A mobilization technique is defined as : A technique comprising a continuum of skilled passive movements that are applied at varying speeds and amplitudes to joints, muscles or nerves with the intent to restore optimal motion, function, and/or to reduce pain (www.IFOMPT.org). The description of these techniques can be found in textbooks [11] and videos on the Internet. This article examines whether the quality of the description of manipulation and mobilization interventions as well as the techniques is sufficient for replication in clinical practice.

This has led to the following research questions;

1. *Are interventions which include manipulation and or mobilization techniques, used in clinical trials on patients with non-specific neck pain described complete according to all items on the TIDdier check list ?*
2. *Are the manipulation or mobilisation techniques described in a reproducible manner?*

This review was performed as a secondary analysis alongside a review on the clinical reasoning process in randomized clinical trials with patients with non-specific neck pain [10] .

Methods

A comprehensive literature search was performed in MEDLINE, CINAHL and PEDro from inception to September 2018. We used a sensitive search strategy that we used in a previous review [12]. To collect all potentially eligible RCTs, the search strategy combined two primary pathways. The first combined neck pain with physical therapy and the second concerned the combination neck pain with the subheadings “rehabilitation”, “therapy” and “prevention and control” because these subheadings included most likely also physical therapy. The first and second pathways were combined with the Boolean term “OR”. Subsequently, the outcome was limited for RCTs with the “Cochrane Highly Sensitive Search Strategy” for identifying randomized trials”.

In CINAHL the same strategy was used as in MEDLINE with an adapted Cochrane search strategy. In PEDro the Abstract and Title box was filled with “neck”, the problem box with “pain” and the method box with “clinical trial”.

Study selection

A study was included if it met the following criteria: full-text original article, published in English, adult patients (>18 years) with non-specific neck pain as their main complaint, the intervention consisting of mobilisation or manipulation techniques and randomized controlled trial (RCT) as the study design. Non-specific neck pain was defined as pain (with or without radiation) located in the cervical spine and/or occiput region and/or cervico thoracic junction and muscles originating from the cervical region acting on the head and shoulders. The selection process was performed by two independent researchers (FM,JP). After independently selecting the studies, the differences were discussed until consensus was reached. If no consensus was reached, a third researcher (HW) was consulted and consensus was reached based on discussion between them.

Two reviewers (FM and NW) independently selected the RCTs with manipulation or mobilization interventions. If a manipulation or mobilization intervention was combined with other interventions, only the manipulation or mobilization intervention part was assessed.

Data-extraction

To determine whether the reporting manipulation and/or mobilization intervention performed on patients with non-specific mechanical neck pain was complete we used the 12-item TIDieR checklist to determine the replicability of these interventions [5].

Each item was scored with no (0) or yes (1), except for items 1, 5 and 8 which are scored with a describtion or actual scores. A score of "0" for item 1, 5 or 8 (a-f) means that it is not described. A score of "0" for item 8 means that the manipulations or mobilizations were not used in combination with other interventions. Again, if no consensus was reached, a third researcher (JP) was consulted and consensus was reached based on discussion between them.
To answer research question 1, all items from the Tidier list were used and to answer research question 2, item 8 (a-f).

Risk of bias assessment
The TiDieR checklist was published in 2014 as an additional exploration. Therefore, we compared the articles published before 2015 with articles published after 2015 to get an impression if the description of interventions was improved and to explore publication bias.

Analysis
A descriptive analysis was performed using the software package of IBM SPSS Statistics 23.0 (SPSS Inc., Chicago, Il).

Results

We found 67 articles using manipulation or mobilization techniques as the intervention under study **(see Figure1)** [13-72].
None of the articles described all the items on the TIDieR list **(see Table 1)**. For example, in only 55,2 % of the RCTs a rationale for the intervention was described. For a complete overview of the scored percentages on the TIDieR items we refer to **Table 2.**

In 17% of the studies manipulation of the cervical spine was used and in 22% manipulation was applied to the thoracic spine (22.6 %); in 18.9 % both manipulation and mobilization techniques were used and in 41.5% only mobilization was used.

Several mobilization techniques were used; in 9.4 % specific Maitland mobilization techniques (11), in 1.9 % Snags techniques by Mulligan and in 30.2 % all other modalities of mobilization techniques. In 47.2% of all treatment sessions a combination of modalities was used, for example the addition of exercises.

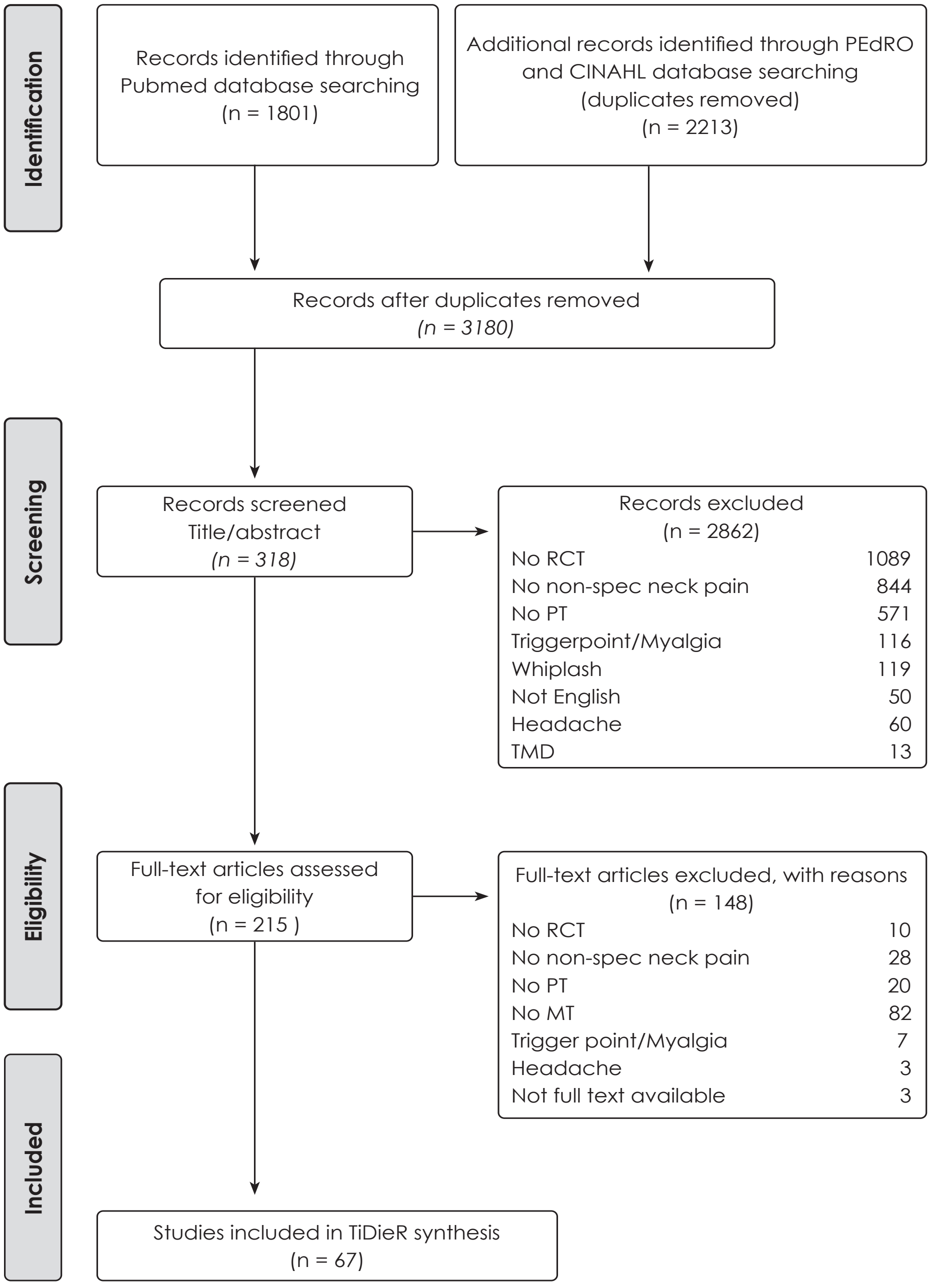


Figure 1: *Flowchart of articles reviewed*

Considering item 8 (a-f) of the TIDieR checklist only one article described the technique completely [73]. In 94.3 % of the articles the number of treatment sessions was described, with a range of 1 to 30 treatment sessions; in 45.3 % of the treatment sessions only 1 technique was the topic of research. The duration of the treatment sessions varied from 1 minute to 45 minutes but in most cases (60.4%) it was not described. In this review the intensity or dose of the techniques was described in 32.1 % of the included articles. In half of these trials, grades of movement were used according to Maitland(11). The vertebral level at which the technique was applied was described in 15.1 %. The majority of authors (51 %) used the results of the physical examination as reference for the level of intervention but did not report what the results were.

After the studies were stratified according to publication date (before or after 2015) the scores of items 2, description of the intervention rationale, 8b frequency and 8f level of the intervention increased slightly **(see Table 3)**.

Discussion

Main findings

None of the articles fully described the manipulation or mobilization interventions used in clinical trials on non-specific neck pain, considering all items of the checklist. Only one article completely described the manipulation or mobilisation technique, considering item 8 of the checklist [13]. The TIDieR checklist intends to check the intervention as a whole. Within the checklist (more specifically, using item 8) we considered the used manipulation and or mobilisation techniques.

We consider the most relevant items in the checklist for replication of these techniques, the rationale (the why), the expertise, the background or level of training of the therapist (the who) and the parameters of the intervention (the what) such as the amount of time, number of sessions, the duration, the intensity and level. Less relevant items are the name of the intervention, "materials" because for these interventions no materials were needed. Tailoring and modifications are expected because each treatment is tailored to the individual patient, although these items were not always reported as such (n = 37,3 %). As Tuttle et al stated [74] that applying parameters of
The second rationale is neuro-physiological, with the aim to influence the patient's pain by applying an input on the neuro-musculo skeletal system. Another issue was the rationale for the selection of a manipulation or mobilization technique. In the context of safety, it seems important whether and, if so, when one prefers manipulation

Table 1: *Score of the TIDieR checklist (n= 67)*

	1	2	3	4	5	6	7	8a	8b	8c	8d	8e	8f	8g	9	10	11	12
	Intervention	rationale	materials	procedure	person	modes	location	sessions	frequency	duration	intensity	Type	level	combination	tailoring	modification	adherance	delivered
Akhter et al., 2014	1	0	0	1	0	1	2	6	2	0	0	1	0	1	1	1	1	0
Ali et al., 2014	2	0	0	1	2	1	1	24	4	0	0	5	0	0	0	1	1	1
Aquino et al., 2009	2	1	0	1	1	1	0	1	0	0	0	4	0	0	1	0	0	0
Beltran-Alacreu et al., 2015	3	0	0	0	1	1	0	8	2	0	0	6	0	0	0	0	0	1
Casanova-Mendez et al., 2014	1	1	0	1	1	1	0	1	0	0	0	2	0	0	0	0	0	0
Celenay et al., 2016	2	1	0	1	1	1	0	12	3	20	0	6	0	1	1	0	0	0
Cleland et al., 2007a	1	1	0	1	1	1	1	2	2	0	0	2	0	1	0	0	0	1
Cleland et al., 2007b	3	0	0	1	1	1	1	1	0	3	0	4	0	1	0	0	0	0
Cleland et al., 2005	1	1	0	1	1	1	1	1	0	0	0	2	0	0	1	0	0	0
de Camargo et al., 2011	1	1	0	1	2	1	0	1	0	0	0	1	1	0	0	0	0	0
Deepa et al., 2014	2	1	0	1	1	1	0	0	0	0	1	6	0	0	1	0	0	0
Dunning et al., 2012	1	1	0	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0
Dziedzic et al., 2005	2	0	0	0	4	1	1	8	0	20	0	6	0	1	1	0	0	0
Evans et al., 2012	1	0	0	0	3	1	1	20	2	20	0	3	0	1	1	1	1	1
Ganesh et al., 2014	2	1	0	0	2	1	2	10	5	10	1	6	0	1	1	1	1	1
Gonzalez-Iglesias et al., 2009b	1	1	0	1	0	1	1	3	1	0	0	2	0	1	0	0	0	0
Gonzalez-Iglesias et al., 2009a	1	0	0	1	1	1	1	3	1	0	0	2	1	1	0	0	0	0

5

Table 1: *Score on the TIDieR checklist (n= 67) (continued)*

	1	2	3	4	5	6	7	8a	8b	8c	8d	8e	8f	8g	9	10	11	12
	Intervention	rationale	materials	procedure	person	modes	location	sessions	frequency	duration	intensity	Type	level	combination	tailoring	modification	adherance	delivered
Griswold et al., 2015	2	0	0	0	4	1	0	3	0	0	1	4	0	1	1	0	0	0
Haas et al., 2003	1	1	0	1	3	1	1	1	0	0	0	1	0	0	1	0	0	0
Hakkinen et al., 2007	2	1	0	1	0	1	2	8	2	10	0	6	0	1	0	0	0	0
Hoving et al., 2002b	2	0	0	1	2	1	1	6	1	45	0	6	0	0	1	0	0	1
Izquierdo Perez et al., 2014	3	0	0	1	1	1	0	4	2	0	0	3	0	0	1	1	1	1
Kanlayanaphotporn et al., 2009	2	0	0	1	2	1	1	1	1	2	1	4	0	0	1	0	0	0
Kanlayanaphotporn et al., 2010	2	0	0	1	2	1	1	1	0	2	1	4	0	0	1	0	0	0
Ko et al., 2010	2	1	0	1	0	1	2	18	6	3	1	6	1	1	0	0	0	0
Krauss et al., 2008	1	1	0	1	2	1	1	1	0	0	0	2	0	0	1	0	0	0
Lau et al., 2011	1	1	0	1	2	1	0	8	2	0	0	2	0	1	1	0	0	1
Leaver et al., 2010	3	0	0	0	1	1	0	4	2	0	0	3	0	1	1	0	0	1
Lee et al., 2013	2	1	0	0	1	1	2	12	3	15	0	6	0	0	1	0	0	0
Lee and Kim, 2016	1	1	0	1	1	1	0	30	3	10	0	2	0	1	1	0	0	1
Lluch et al., 2014	2	1	0	1	1	1	0	1	0	3	0	6	0	1	1	0	0	1
Lopez-Lopez et al., 2014	3	0	0	1	0	1	1	1	0	0	1	3	0	0	1	0	0	0
Madson et al., 2010	2	0	0	1	1	1	1	12	3	30	1	6	0	1	1	0	0	1
Mansilla-Ferragut et al., 2009	1	1	0	1	2	1	0	1	0	0	1	1	0	0	0	0	0	0
Martel et al., 2011	1	0	0	1	3	1	0	15	2	10	1	1	0	1	0	0	0	1
Martinez-Segura et al., 2006	1	0	0	1	2	1	0	1	0	0	0	3	0	0	1	0	0	1
Masaracchio et al., 2013	3	0	0	1	1	1	0	2	0	0	1	3	1	1	1	0	0	1

Pillastrini et al., 2016	2	1	0	1	0	1	0	0	0	30	1	6	0	1	0	0	0	1
Pires et al., 2015	1	1	0	1	2	1	2	1	0	0	0	2	1	0	0	0	0	1
Puentedura et al., 2011	1	1	0	1	2	1	0	0	2	0	0	3	0	1	0	0	0	1
Puntumetakul et al., 2015	1	1	0	1	1	1	0	1	0	0	0	2	0	0	0	0	0	1
Saavedra-Hernandez et al., 2012	1	0	0	1	0	1	1	1	0	0	0	1	0	0	0	0	0	1
Saavedra-Hernandez et al., 2013	1	0	0	1	2	1	1	1	0	0	0	3	0	0	1	0	0	1
Saayman et al., 2011	1	1	0	1	0	1	1	6	2	30	0	1	0	1	1	0	0	1
Salom-Moreno et al., 2014	1	0	0	1	1	1	0	1	0	0	1	2	0	0	0	0	0	1
Schomacher, 2009	2	0	0	1	1	1	0	1	0	4	1	6	0	0	1	0	0	0
Sillevis et al., 2010	1	1	0	1	2	1	0	1	0	0	0	2	1	0	0	0	0	0
Snodgrass et al., 2014	2	0	0	1	1	1	1	1	0	3	1	6	0	0	1	0	0	1
Sterling et al., 2001	2	1	0	1	2	1	2	1	0	0	1	6	1	0	1	0	0	0
Walker et al., 2008	3	0	0	1	1	1	2	6	2	0	0	1	0	1	1	0	0	0
Yang et al., 2015	1	1	0	1	0	1	0	18	3	0	0	3	0	1	0	0	0	0
Ylinen et al., 2007	2	0	0	1	1	1	2	8	2	30	0	6	0	1	0	0	0	1
Baustita et al 2017	1	1	0	1	2	1	1	1	1	0	0	3	3	0	1	0	0	0
Buykturan et al 2018	2	1	0	1	1	1	2	10	5	0	1	5	2	1	1	0	1	1
Cho et al 2017	2	1	0	1	1	1	2	10	2	5	0	6	2	1	0	0	0	1
Duymaz et al 2018	2	1	0	1	0	1	0	10	5	5	0	5	0	1	0	0	0	1
Farooq et al 2018	2	0	0	1	1	1	2	10	2	0	1	4	0	1	1	0	0	1
Galindez et al 2018	1	1	0	1	2	1	1	1	1	0	0	3	0	1	0	0	0	1

5

Table 1: *Score on the TIDieR checklist (n= 67)* *(continued)*

	1	2	3	4	5	6	7	8a	8b	8c	8d	8e	8f	8g	9	10	11	12
	Intervention	rationale	materials	procedure	person	modes	location	sessions	frequency	duration	intensity	Type	level	combination	tailoring	modification	adherance	delivered
Gorrel et al 2016	1	1	0	1	0	1	2	1	1	0	0	1	0	1	1	0	0	1
Griswold et al 2018	3	1	0	1	2	1	1	2	0	0	0	3	0	1	1	0	0	1
Groeneweg et al 2017	2	0	0	1	2	1	1	6	1	30	0	6	0	1	1	0	0	1
Karas et al 2014	1	1	0	1	2	1	1	1	1	0	0	2	0	0	1	1	0	1
Lee et al 2017	2	0	0	1	1	1	0	6	3	60	1	4	0	1	1	0	0	1
Maiers et al 2014	1	0	0	1	3	1	1	20	2	0	0	1	0	1	1	0	1	
NiNiWin et al 2015	1	0	0	0	3	1	1	2	1	0	0	1	1	0	1	0	0	0
Petersen et al 2015	3	1	0	1	2	1	0	1	1	0	0	1	0	1	1	0	0	1

1)1= manipulation 2= mobilization 3= both:
5)0 = not known, 1= PT, 2=MT, 3= chiropractor, 4= other profession
7) 0= not known, 1= private practice, 2= institutional
8a) no of sessions
8b) no of sessions per week
8c) duration of a session
8d) description of , 0 = no, 1 = yes
8e) 1= manip Cx, 2= manip Tx, 3= combination Cx Tx, 4= maitland mob, 5= Snags, 6= diversity
8f) level of manipulation /mobilization, 0= no, 1 = yes 8g) combination with other modality, 0= no, 1= yes

Table 2: *Summary of scores TIDieR checklist (n=67)*

TIDieR items	
1. Description of the name of the intervention	100 %
2. Description of the intervention rationale, theory or goal of the elements essential to the intervention	55,2 %
3. Description of materials used in the intervention	46,0 %
4. Detailed description of procedures used in the intervention	88,0 %
5. Description of the person who provided the intervention	83,6 %
6. Description of the modes of delivery (such as face to face)	100 %
7. Description of the location where the intervention occurred	60,7 %
8. Description of the parameters regarding the intervention	
8a. No of sessions	95,5 %
8b. Frequency	56,7 %
8c. Duration (min)	37,3 %
8d. Intensity or dose	32,8 %
8e. Type of intervention	91,0 %
8f. Level of intervention	20,9 %
8g. Combination of intervention	53,7 %
9. Was the intervention tailored i.e. personalized?	42,0 %
10. Was the intervention modified during the treatment?	18,0 %
11. Was the adherence of the intervention assessed	75 %
12. If so: was the intervention delivered as planned?	55,2 %

Table 3: *Differences in scores TIDieR checklist articles published* before *2015 versus published* after *2015*

TIDieR items	48 <2015	19> 2015
1. Description of the name of the intervention	100 %	100 %
2. Description of the intervention rationale, theory or goal of the elements essential to the intervention	52.8 %	73.7 %
3. Description of materials used in the intervention	na	
4. Detailed description of procedures used in the intervention	86.8 %	89.5 %
5. Description of the person who provided the intervention	100 %	100 %
6. Description of the modes of delivery (such as face to face)	100 %	100 %
7. Description of the location where the intervention occurred	54.7 %	52.6 %
8. Description of the parameters regarding the intervention		
8a. No of sessions	94.3 %	94.7%
8b. Frequency	47.2 %	68.4 %
8c. Duration (min)	39.6 %	36.8 %
8d. Intensity or dose	32.1 %	42.1 %
8e. Type of intervention	88.7 %	100 %
8f. Level of intervention	15.1 %	36.8 %
8g. Combination of intervention	47.2 %	73.7 %
9. Was the intervention tailored i.e. personalized?	58.5 %	11 %
10. Was the intervention modified during the treatment?	30.2 %	21 %
11. Was the adherence of the intervention assessed	7.7 %	5.3 %
12. If so: was the intervention delivered as planned?	48.1 %	68.4%

na = not applicable ; [shaded] *= Improved TIDieR items compared to < 2015* techniques are no longer predetermined but are continuously adjusted due to the feedback of the patients and the adjustment of the therapist to the responses of the individual patient. Furthermore, the perception of the patient as well of the therapist on the performed intervention is of influence on the tailoring of the intervention.

Finally, intervention adherence assessment is a less relevant item because adherence does not provide information about how the intervention is performed.

Although a rationale is not necessary to replicate an intervention, we consider it a relevant tem because there must be a hypothesis present on why the intervention could be effective. The rationale for the use of the techniques was described in only 55.2 % of the included studies. Two main rationales for the use of manipulation and/or mobilization were described.

First, the biomechanical rationale; the therapist identified a hypo mobile segment or articular dysfunction and used a manipulation or mobilization technique to restore mobility. over mobilization [75,76]. From this current review it was impossible to determine why a manipulation or a mobilization was chosen.

Because the checklist was published in 2014, we also compared the articles published before 2015 with articles published after 2015 to get an impression if the description of interventions had improved. The description improved slightly after the publication of the TIDieR checklist. See **Table 3**.

A diagnostic clinical reasoning process for determining the segmental level at which the intervention would be applied to was used in 55% of the trials, however the specific levels were not always reported. Although the specific segmental level is potentially relevant, its relevance can also be questioned because several studies showed that the validity and reliability of determining a segment to be treated is low [17] . In addition, Hegedus et al [78] stated in a recently published review, that it is still unclear whether it is necessary to determine a specific level with cervical mobilization. In this review we found that in 20,9 % the level was described as “high cervical spine”, “mid cervical spine” or “thoracal spine”.

The dose or the intensity of the manipulation or mobilisation technique seems important [79], however, enormous variations in research exist [80]. In this survey the intensity or dose of the techniques was described in 32.1 % of include articles, half of these trials, grades of movement were used according to Maitland [11]. Furthermore, inter-reliability of assessing the grade of movement was poor (ICC= 0.23) and intra-reliability was moderate to good (ICC 0.83-0.94) [80,81]. Further research must demonstrate whether the dosage or intensity matters.

Comparison with existing literature

Overall, the interventions were poorly reported , the used manipulation and or mobilisation techniques somewhat better. One reason for this may be the result of the word limits imposed on authors by journals [82]. Conn et al reported that only 7% of the space in an article was used for description of the intervention in 141 studies in Nursing Research Journals [6]. A possible solution could be attaching an appendix describing the details of the intervention or a design article with the complete description of the intervention [83]. Also, specific register forms can be used to describe every used technique during the treatments sessions [84]. However, it remain problematic to describe an optimal dose, level and frequency. Another option was suggested by Glasziou et al.[8], these authors suggested to video three interventions in advance of conducting the clinical trial with, for example, a mild, moderate or intense intervention which is an option to consider. However, this is a time consuming and potentially costly method, although the use of a smartphone can make it more accessible.

Weaknesses

The results of this study should be interpreted in the light of some limitations. As far as we know, nothing is known about the methodological properties of the TIDieR list. The use of a dichotomous response options on most items on the checklist restrict full information about topics concerning the intervention, a more qualitative description could be more informative.

Furthermore, does the description of intervention or the description of a specific techniques fully resembles what happens in daily practice? This is also related to the heterogeneity of patient's problems, patient's reaction and patient's perception. Also, the beliefs of the patients and of the therapist plays a role in the application of techniques and or the intervention as a whole [85].

Finally, we only included English-language research. There is a chance that this has affected the results, although given the amount of included articles, this chance seems small.

Strengths

A strength of this study is the use of a sensitive search strategies in multiple databases, developed in collaboration with a medical information specialist.

Another strength is the use of the TIDieR checklist as this includes all relevant aspect that should be described.

Implication

What can be expected of a description of the intervention within the methodology of a RCT? A manipulation or mobilization technique is in most cases tailored to the individual patient as the dosage, velocity and segmental level for example. Also, the inconsistency of the applied manual forces during spinal mobilization has to be taken into account [80,86] which makes it very difficult to describe the amount of force and the replication of it, which can be a topic for future research. In our opinion, the TIDieR checklist covers the most important items to give an impression of the completeness of the description of an manipulation or mobilisation intervention, although the specific description of the technique should be considered for addition to the TiDieR checklist.

Conclusion

In conclusion, interventions with manipulation or mobilization techniques are poorly reported in RCTs. Poor reporting and incomplete descriptions of the techniques jeopardize the external validity of RCTs, making it difficult for clinicians and researchers to replicate the techniques. It is also important to investigate which aspects matter with regard to the effectiveness of manipulations and mobilizations.

References

1. Akobeng AK. Understanding randomised controlled trials. Arch Dis Child. 2005;90(8):840–4.
2. Nasser M, van Weel C, van Binsbergen JJ, van de Laar FA. Generalizability of systematic reviews of the effectiveness of health care interventions to primary health care: concepts, methods and future research. Fam Pract. 2012;29 Suppl 1:i94–103.
3. Rothwell PM. Commentary: External validity of results of randomized trials: disentangling a complex concept. Int J Epidemiol. 2010;39(1):94–6.
4. Dekkers OM, Elm E von, Algra A, Romijn JA, Vandenbroucke JP. How to assess the external validity of therapeutic trials: a conceptual approach. Int J Epidemiol. 2010;39(1):89–94.
5. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ.2014;348:g1687. WW
6. Conn VS. Unpacking the black box: countering the problem of inadequate intervention descriptions in research reports. West J Nurs Res. 2012;34(4):427–33.
7. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
8. Glasziou P, Chalmers I, Altman DG, Bastian H, Boutron I, Brice A, et al. Taking healthcare interventions from trial to practice. BMJ. 2010;341:c3852.
9. Hurwitz EL, Randhawa K, Yu H, Côté P, Haldeman S. The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. European Spine Journal. 2018; 27 (6), 796-801
10. Maissan F, Pool J, de Raaij E, Mollema J, Ostelo R, Wittink H. The clinical reasoning process in randomized clinical trials with patients with non-specific neck pain is incomplete: A systematic review. Vol. 35, Musculoskeletal Science and Practice. 2018. p. 8–17.
11. Maitland G, Hengeveld E, Banks K, English K. Maitland's Vertebral Manipulation. 2005.
12. Maissan F, Pool J, Stutterheim E, Wittink H, Ostelo R. Clinical reasoning in unimodal interventions in patients with non-specific neck pain in daily physiotherapy practice, a Delphi study. Musculoskelet Sci Pract. 2018; 14 (9),8-16
13. Akhter S, Khan M, Ali SS, Soomro RR. Role of manual therapy with exercise regime versus exercise regime alone in the management of non-specific chronic neck pain. Pak J Pharm Sci. 2014;27(6):2125–8.
14. Ali A, Shakil-Ur-Rehman S, Sibtain F. The efficacy of Sustained Natural ApophysGlides with and without Isometric Exercise Training in Non-specific Neck Pain. Pakistan J Med Sci. 2014;30(4):872–4.
15. Beltran-Alacreu H, Lopez-de-Uralde-Villanueva I, Fernandez-Carnero J, La Touche R. Manual Therapy, Therapeutic Patient Education, and Therapeutic Exercise, an Effective Multimodal Treatment of Nonspecific Chronic Neck Pain: A Randomized Controlled Trial. Am J

Phys Med & Rehabil /Assoc Acad Physiatr. 2015;94(10 Suppl 1):887–97.

16. Casanova-Mendez A, Oliva-Pascual-Vaca A, Rodriguez-Blanco C, Heredia-Rizo AM, Gogorza-Arroitaonandia K, Almazan-Campos G. Comparative short-term effects of two thoracic spinal manipulation techniques in subjects with chronic mechanical neck pain: a randomized controlled trial. Man Ther. 2014;19(4):331–7.
17. Celenay ST, Akbayrak T, Kaya DO. A Comparison of the Effects of Stabilization Exercises Plus Manual Therapy to Those of Stabilization Exercises Alone in Patients With Nonspecific Mechanical Neck Pain: A Randomized Clinical Trial. J Orthop Sports Phys Ther. 2016;46(2):44–55.
18. Celenay ST, Kaya DO, Akbayrak T. Cervical and scapulothoracic stabilization exercises with and without connective tissue massage for chronic mechanical neck pain: A prospective, randomised controlled trial. Man Ther. 2016;21:144–50.
19. Cleland JA, Glynn P, Whitman JM, Eberhart SL, MacDonald C, Childs JD. Short-term effects of thrust versus nonthrust mobilization/manipulation directed at the thoracic spine in patients with neck pain: a randomized clinical trial. Phys Ther. 2007;87(4):431–40.
20. Cleland JA, Childs JD, McRae M, Palmer JA, Stowell T. Immediate effects of thoracic manipulation in patients with neck pain: a randomized clinical trial. Man Ther. 2005;10(2):127–35.
21. Cleland JA, Childs JD, Fritz JM, Whitman JM, Eberhart SL. 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. 2007;87(1):9–23.
22. de Camargo VM, Alburquerque-Sendin F, Berzin F, Stefanelli VC, de Souza DP, Fernandez-de-las-Penas C. Immediate effects on electromyographic activity and pressure pain thresholds after a cervical manipulation in mechanical neck pain: a randomized controlled trial. J Manipulative Physiol Ther. 2011;34(4):211–20.
23. Deepa A, Dabholkar Y. T, Yardi S. Comparison of the efficacy of Maitland Thoracic Mobilization and Deep Neck Flexor Endurance Training Versus Only Deep Neck Flexor Endurance Training in Patients with Mechanical Neck Pain. Indian J Physiother & Occup Ther. 2014;8(3):77–82.
24. Dunning JR, Cleland JA, Waldrop MA, Arnot CF, Young IA, Turner M, et al. Upper cervical and upper thoracic thrust manipulation versus nonthrust mobilization in patients with mechanical neck pain: a multicenter randomized clinical trial. J Orthop Sports Phys Ther. 2012;42(1):5–18.
25. Dziedzic K, Hill J, Lewis M, Sim J, Daniels J, Hay EM. Effectiveness of manual therapy or pulsed shortwave diathermy in addition to advice and exercise for neck disorders: a pragmatic randomized controlled trial in physical therapy clinics. Arthritis Rheum. 2005;53(2):214–22.
26. Raney NH, Petersen EJ, Smith TA, Cowan JE, Rendeiro DG, Deyle GD, et al. Development of a clinical prediction rule to identify patients with neck pain likely to benefit from cervical traction and exercise. Eur Spine J. 2009;18(3):382–91.
27. Evans R, Bronfort G, Schulz C, Maiers M, Bracha Y, Svendsen K, et al. Supervised exercise

with and without spinal manipulation performs similarly and better than home exercise for chronic neck pain: a randomized controlled trial. Spine (Phila Pa 1976). 2012;37(11):903–14.
28. Ganesh GS, Mohanty P, Pattnaik M, Mishra C. Effectiveness of mobilization therapy and exercises in mechanical neck pain. Physiother Theory Pract. 2014;1–8.
29. Gonzalez-Iglesias J, Fernandez-de-las-Penas C, Cleland JA, Gutierrez-Vega Mdel R. Thoracic spine manipulation for the management of patients with neck pain: a randomized clinical trial. J Orthop Sports Phys Ther. 2009;39(1):20–7.
30. Gonzalez-Iglesias J, Fernandez-de-las-Penas C, Cleland JA, Alburquerque-Sendin F, Palomeque-del-Cerro L, Mendez-Sanchez R. Inclusion of thoracic spine thrust manipulation into an electro therapy/thermal program for the management of patients with acute mechanical neck pain: a randomized clinical trial. Man Ther. 2009;14(3):306–13.
31. Griswold D, Learman K, O'Halloran B, Cleland J. A preliminary study comparing the use of cervical/upper thoracic mobilization and manipulation for individuals with mechanical neck pain. J Man & Manip Ther. 2015;23(2):75–83.
32. Haas M, Groupp E, Panzer D, Partna L, Lumsden S, Aickin M. Efficacy of cervical endplay assessment as an indicator for spinal manipulation. Spine (Phila Pa 1976). 2003;28(11):1091–6; discussion 1096.
33. Hakkinen A, Salo P, Tarvainen U, Wiren K, Ylinen J. Effect of manual therapy and stretching on neck muscle strength and mobility in chronic neck pain. J Rehabil Med. 2007;39(7):575–9.
34. Hoving JL, Koes BW, De Vet HCW, Van der Windt DAWM, Assendelft WJJ, Van Mameren H, et al. Manual therapy, physical therapy, or continued care by a general practitioner for patients with neck pain: A randomized, controlled trial. Ann Intern Med. 2002;136(10).
35. Izquierdo Perez H, Alonso Perez JL, Gil Martinez A, La Touche R, Lerma-Lara S, Commeaux Gonzalez N, et al. Is one better than another?: A randomized clinical trial of manual therapy for patients with chronic neck pain. Man Ther. 2014;19(3):215–21.
36. Kanlayanaphotporn R, Chiradejnant A, Vachalathiti R. Immediate effects of the central posteroanterior mobilization technique on pain and range of motion in patients with mechanical neck pain. Disabil Rehabil. 2010;32(8):622–8.
37. Kanlayanaphotporn R, Chiradejnant A, Vachalathiti R. The immediate effects of mobilization technique on pain and range of motion in patients presenting with unilateral neck pain: a randomized controlled trial. Arch Phys Med Rehabil. 2009;90(2):187–92.
38. Krauss J, Creighton D, Ely JD, Podlewska-Ely J. The immediate effects of upper thoracic translatoric spinal manipulation on cervical pain and range of motion: a randomized clinical trial. J Man Manip Ther 2008;16(2):93–9.
39. Lau HM, Wing Chiu TT, Lam TH. The effectiveness of thoracic manipulation on patients with chronic mechanical neck pain - a randomized controlled trial. Man Ther. 2011;16(2):141–7.
40. Leaver AM, Maher CG, Herbert RD, Latimer J, McAuley JH, Jull G, et al. A randomized controlled trial comparing manipulation with mobilization for recent onset neck pain. Arch Phys

Med Rehabil. 2010;91(9):1313–8.

41. Lee J, Lee Y, Kim H, Lee J. The effects of cervical mobilization combined with thoracic mobilization on forward head posture of neck pain patients. J Phys Ther Sci. 2013;25(1):7–9.
42. Lee KW, Kim WH. Effect of thoracic manipulation and deep craniocervical flexor training on pain, mobility, strength, and disability of the neck of patients with chronic nonspecific neck pain: a randomized clinical trial. J Phys Ther Sci. 2016;28(1):175–80.
43. Lluch E, Schomacher J, Gizzi L, Petzke F, Seegar D, Falla D. Immediate effects of active cranio-cervical flexion exercise versus passive mobilisation of the upper cervical spine on pain and performance on the cranio-cervical flexion test. Man Ther. 2014;19(1):25–31.
44. Lopez-Lopez A, Alonso Perez JL, Gonzalez Gutierez JL, La Touche R, Lerma Lara S, Izquierdo H, et al. Mobilization versus manipulations versus sustain appophyseal natural glide techniques and interaction with psychological factors for patients with chronic neck pain: Randomized control Trial. Eur J Phys Rehabil Med. 2014; 51 (2), 121-132
45. Madson TJ, Cieslak KR, Gay RE. Joint mobilization vs massage for chronic mechanical neck pain: a pilot study to assess recruitment strategies and estimate outcome measure variability. J Manipulative Physiol Ther. 2010;33(9):644–51.
46. Mansilla-Ferragut P, Fernandez-de-Las Penas C, Alburquerque-Sendin F, Cleland JA, Bosca-Gandia JJ. Immediate effects of atlanto-occipital joint manipulation on active mouth opening and pressure pain sensitivity in women with mechanical neck pain. J Manipulative Physiol Ther. 2009;32(2):101–6.
47. Martel J, Dugas C, Dubois JD, Descarreaux M. A randomised controlled trial of preventive spinal manipulation with and without a home exercise program for patients with chronic neck pain. BMC Musculoskelet Disord. 2011;12:41.
48. Martinez-Segura R, Fernandez-de-las-Penas C, Ruiz-Saez M, Lopez-Jimenez C, Rodriguez-Blanco C. Immediate effects on neck pain and active range of motion after a single cervical high-velocity low-amplitude manipulation in subjects presenting with mechanical neck pain: a randomized controlled trial. J Manipulative Physiol Ther. 2006;29(7):511–7.
49. Masaracchio M, Cleland JA, Hellman M, Hagins M. Short-term combined effects of thoracic spine thrust manipulation and cervical spine nonthrust manipulation in individuals with mechanical neck pain: a randomized clinical trial. J Orthop Sports Phys Ther. 2013;43(3):118–27.
50. Pillastrini P, de Lima E Sa Resende F, Banchelli F, Burioli A, Di Ciaccio E, Guccione AA, et al. Effectiveness of Global Postural Re-education in Patients With Chronic Nonspecific Neck Pain: Randomized Controlled Trial. Phys Ther. 2016;96(9):1408–16.
51. Pires PF, Packer AC, Dibai-Filho A V, Rodrigues-Bigaton D. Immediate and Short-Term Effects of Upper Thoracic Manipulation on Myoelectric Activity of Sternocleidomastoid Muscles in Young Women With Chronic Neck Pain: A Randomized Blind Clinical Trial. J Manipulative Physiol Ther. 2015;38(8):555–63.
52. Puentedura EJ, Landers MR, Cleland JA, Mintken PE, Huijbregts P, Fernandez-de-Las-Penas

C. Thoracic spine thrust manipulation versus cervical spine thrust manipulation in patients with acute neck pain: a randomized clinical trial. J Orthop Sports Phys Ther. 2011;41(4):208–20.

53. Puntumetakul R, Suvarnnato T, Werasirirat P, Uthaikhup S, Yamauchi J, Boucaut R. Acute effects of single and multiple level thoracic manipulations on chronic mechanical neck pain: a randomized controlled trial. Neuropsychiatr Dis Treat. 2015;11:137–44.
54. Saavedra-Hernandez M, Castro-Sanchez AM, Arroyo-Morales M, Cleland JA, Lara-Palomo IC, Fernandez-de-Las-Penas C. Short-term effects of kinesio taping versus cervical thrust manipulation in patients with mechanical neck pain: a randomized clinical trial. J Orthop Sports Phys Ther. 2012;42(8):724–30.
55. Saavedra-Hernandez M, Arroyo-Morales M, Cantarero-Villanueva I, Fernandez-Lao C, Castro-Sanchez M. A, Puentedura J. E, et al. Short-term effects of spinal thrust joint manipulation in patients with chronic neck pain: a randomized clinical trial. Clin Rehabil. 2013;27(6):504–12.
56. Saayman L, Hay C, Abrahamse H. Chiropractic manipulative therapy and low-level laser therapy in the management of cervical facet dysfunction: a randomized controlled study. J Manipulative Physiol Ther. 2011;34(3):153–63.
57. Salom-Moreno J, Ortega-Santiago R, Cleland Aland, Palacios-Ceña, M., Truyols-Domínguez, S., Fernández-de-las-Peñas, C. Immediate Changes in Neck Pain Intensity and Widespread Pressure Pain Sensitivity in Patients With Bilateral Chronic Mechanical Neck Pain: A Randomized Controlled Trial of Thoracic Thrust Manipulation vs Nonâ€“Thrust Mobilization. J Manip & Physiol Ther. 2014;37(5):312–9.
58. Schomacher J. The effect of an analgesic mobilization technique when applied at symptomatic or asymptomatic levels of the cervical spine in subjects with neck pain: a randomized controlled trial. J Man & Manip Ther 2009;17(2):101–8.
59. Sillevis R, Cleland J, Hellman M, Beekhuizen K. Immediate effects of a thoracic spine thrust manipulation on the autonomic nervous system: a randomized clinical trial. J Man & Manip Ther 2010;18(4):181–90.
60. Snodgrass J. S, Rivett A. D, Sterling M, Vicenzino B. Dose Optimization for Spinal Treatment Effectiveness: A Randomized Controlled Trial Investigating the Effects of High and Low Mobilization Forces in Patients With Neck Pain. J Orthop & Sport Phys Ther. 2014;44(3):141–52.
61. Sterling M, Jull G, Wright A. Cervical mobilisation: concurrent effects on pain, sympathetic nervous system activity and motor activity. Man Ther. 2001;6(2):72–81.
62. Walker MJ, Boyles RE, Young BA, Strunce JB, Garber MB, Whitman JM, et al. The effectiveness of manual physical therapy and exercise for mechanical neck pain: a randomized clinical trial. Spine 2008;33(22):2371–8.
63. Yang J, Lee B, Kim C. Changes in proprioception and pain in patients with neck pain after upper thoracic manipulation. J Phys Ther Sci. 2015;27(3):795–8.

64. Griswold D, Learman K, Kolber MJ, O'Halloran B, Cleland JA. Pragmatically Applied Cervical and Thoracic Nonthrust Manipulation Versus Thrust Manipulation for Patients With Mechanical Neck Pain: A Multicenter Randomized Clinical Trial. J Orthop Sport Phys Ther. 2018; 48 (3):137-145
65. Bautista-Aguirre F, Oliva-Pascual-Vaca Á, Heredia-Rizo A, Bosca-Gandia J, Ricard F, Rodriguez-Blanco C. Effect of cervical vs. thoracic spinal manipulation on peripheral neural features and grip strength in subjects with chronic mechanical neck pain: a randomized controlled trial. Eur J Phys Rehabil Med. 2017; 53 (3), 333-341
66. Buyukturan O, Buyukturan B, Sas S, Kararti C, Ceylan I. The effect of mulligan mobilization technique in older adults with neck pain: A randomized controlled, double-blind study. Pain Res Manag. 2018; 2856375.
67. Farooq MN, Mohseni-Bandpei MA, Gilani SA, Ashfaq M, Mahmood Q. The effects of neck mobilization in patients with chronic neck pain: A randomized controlled trial. J Bodyw Mov Ther. 2018; 22 (1), 24-31
68. Groeneweg R, van Assen L, Kropman H, Leopold H, Mulder J, Smits-Engelsman BCM, et al. Manual therapy compared with physical therapy in patients with non-specific neck pain: A randomized controlled trial. Chiropr Man Ther. 2017; 25 (12)
69. Gorrell LM, Beath K, Engel RM. Manual and Instrument Applied Cervical Manipulation for Mechanical Neck Pain: A Randomized Controlled Trial. J Manipulative Physiol Ther. 2016; 39 (5), 319-329
70. Maiers M, Bronfort G, Evans R, Hartvigsen J, Svendsen K, Bracha Y, Schulz C, Schulz K, Grimm R. Spinal manipulative therapy and exercise for seniors with chronic neck pain. Spine J. 2014; 14 (9), 1879-1889
71. Petersen SB, Cook C, Donaldson M, Hassen A, Ellis A, Learman K. The effect of manual therapy with augmentative exercises for neck pain: a randomised clinical trial. J Man Manip Ther. 2015; 23 (5),264-275
72. Win NN, Jorgensen AMS, Chen YS, Haneline MT. Effects of upper and lower cervical spinal manipulative therapy on blood pressure and heart rate variability in volunteers and patients with neck pain: A randomized controlled, cross-over, preliminary study. J Chiropr Med. 2015; 14 (1), 1-9
73. Ko T, Jeong U, Lee K. Effects of the inclusion thoracic mobilization into cranio-cervical flexor exercise in patients with chronic neck pain. J Phys Ther Sci. 2010;22(1):87–91.
74. Tuttle N, Hazle C. An empirical, pragmatic approach applying reflection in interaction approach to manual therapy treatments. Physiother Theory Pract. 2019; 6, 1-12
75. Kranenburg HA, Schmitt MA, Puentedura EJ, Luijckx GJ, van der Schans CP. Adverse events associated with the use of cervical spine manipulation or mobilization and patient characteristics: A systematic review. Musculoskeletal Science and Practice. 2017. 30
76. Haynes MJ, Vincent K, Fischhoff C, Bremner AP, Lanlo O, Hankey GJ. Assessing the risk of

stroke from neck manipulation: a systematic review. Int J Clin Pract. 2012;66(10):940–7.

77. Harlick JC, Milosavljevic S, Milburn PD. Palpation identification of spinous processes in the lumbar spine. Man Ther. 2007;
78. Slaven EJ, Goode AP, Coronado RA, Poole C, Hegedus EJ. The relative effectiveness of segment specific level and non-specific level spinal joint mobilization on pain and range of motion: results of a systematic review and meta-analysis. J Man Manip Ther. 2013; 21 (1), 7-17
79. Gross A, Miller J, D'Sylva J, Burnie SJ, Goldsmith CH, Graham N, et al. Manipulation or mobilisation for neck pain: A Cochrane Review. Manual Therapy. 2010. 15 (4) 315-333
80. Snodgrass SJ, Rivett DA, Robertson VJ. Manual Forces Applied During Posterior-to-Anterior Spinal Mobilization: A Review of the Evidence. Journal of Manipulative and Physiological Therapeutics. 2006; 29 (4), 316-329
81. Snodgrass SJ, Rivett DA, Robertson VJ, Stojanovski E. Cervical spine mobilisation forces applied by physiotherapy students. Physiotherapy. 2010; 96 (2) 120-129
82. Yamato TP, Maher CG, Saragiotto BT, Hoffmann TC, Moseley AM. How completely are physiotherapy interventions described in reports of randomised trials? Physiotherapy. 2016;102(2):121–6.
83. Pool JJM, Ostelo RWJG, Köke AJ, Bouter LM, de Vet HCW. Comparison of the effectiveness of a behavioural graded activity program and manual therapy in patients with sub-acute neck pain: Design of a randomized clinical trial. Man Ther. 2006;11(4).
84. Holtorp J, Molenaar N, Plaatsman GP, Veen L, Pool JJM. What is the most frequent used intervention by non-specific neckpain in Dutch manual therapy practises. In: IFOMPT Congress proceedings 2008 Rotterdam. 2008.
85. Bialosky JE, Beneciuk JM, Bishop MD, Coronado RA, Penza CW, Simon CB, et al. Unraveling the mechanisms of manual therapy: Modeling an approach. Journal of Orthopaedic and Sports Physical Therapy. 2018. 48 (1), 8-18
86. Snodgrass SJ, Rivett DA, Robertson VJ, Stojanovski E. Forces Applied to the Cervical Spine During Posteroanterior Mobilization. J Manipulative Physiol Ther. 2009; 32 (1) 72-83

6

The diagnostic accuracy of self-report and physical tests for measuring limitations in the range of motion of the neck.

François Maissan
Jan Pool
Edwin de Raaij
Marloes Thoomes-de Graaf
Raymond Ostelo
Harriët Wittink

Physiotherapy Theory and Practice.
Under review.

Abstract

Objective: To determine the diagnostic accuracy of a self-report test, eight physical tests, and combinations of both, for the measurement of limitations in cervical range of motion (ROM) as index tests, compared to the Cervical Range of Movement (CROM) device as reference test.

Method: Subjects with non-specific neck pain were included. A self-report test and eight physical examination tests were investigated separately, and combinations of the self-report test with the best physical test for a particular movement direction were also investigated. Diagnostic accuracy was determined by calculating sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios (LR+, LR-) and their 95% confidence intervals.

Results: In total, 128 subjects were included. In general, combining the self-report test with the best physical tests had the best diagnostic accuracy. The LR+ for the best combination of tests ranged from 2.96 (right rotation) to 1.39 (flexion). The LR- ranged from 0.61 (flexion) to 0.19 (left lateral flexion).

Conclusion: The LRs of the combination of the self-report test and the best physical tests were small but sometimes clinically important. Unfortunately, they have not demonstrated to be useful for all directions of movement. Therefore, we advise the use of the CROM device to determine a restricted ROM of the neck.

Key words:
Diagnostic test, neck pain, range of motion

Introduction

Neck pain is a common condition that causes substantial disability. Globally, the point prevalence of neck pain is 4.9% (95% confidence interval (CI) 4.6 to 5.3) [1]. Worldwide, disability-adjusted life years increased from 23.9 million (95% CI 16.5 to 33.1) in 1990 to 33.6 million (95% CI 23.5 to 46.5) in 2010 [1]. In 2015, more than 333 million people worldwide had neck pain for more than three months' duration [2].

Limitations in cervical Range of Motion (ROM) have been associated with the development of neck pain [3] and people with non-specific neck pain often have significant limitations in cervical ROM, compared to those without [4]. Cervical ROM is often measured to document baseline status and treatment effect, and to readjust treatment plans as needed in physiotherapy [5].

There are many methods of evaluating cervical ROM, ranging from more traditional subjective methods (self-report, physical examination) to more objective measurement devices using more or less advanced technology [6]. In international clinical guidelines, no specific recommendations are made about the assessment of ROM [7,8]. To objectively assess cervical ROM, reliable and valid tests or instruments are needed with small measurement errors. To date, no study has assessed which of the commonly used clinical measurement methods yields precise estimates of restrictions of movement of the cervical spine.

Therefore, our main research question was: What is the diagnostic accuracy of commonly used methods compared with an objective measurement device to assess restrictions in ROM of the neck in subjects with non-specific neck pain.
This main research question is divided into three sub-questions:

- *What is the diagnostic accuracy of a self-report of limitations in cervical ROM by the patient as part of history-taking?*
- *What is the diagnostic accuracy of tests to assess restrictions in cervical ROM as part of physical examination?*
- *What is the diagnostic accuracy of the combination of self-report and physical examination tests in assessing restrictions in cervical ROM?*

Methods

The reporting of this study of diagnostic accuracy follows the STARD statement [9]. This study was approved by the Institutional Review Board of HU University of Applied Sciences Utrecht (reference number 78_000_2018).

Participants

Consecutive patients who entered physiotherapy consultations between December 2018 and June 2019 and met the inclusion criteria were recruited. Physiotherapists were allowed to evaluate patient according to their discretion. Based on history-taking and, if necessary, additional physical examination tests, eligible subjects were recruited by their physiotherapists. The inclusion criteria were: ≥18 years of age; non-specific neck pain (acute and chronic); and good understanding of the Dutch language. Non-specific neck pain was defined as pain located in the cervical spine and/or occiput region and/or cervico-thoracic junction and/or muscles originating from the cervical region acting on the head and shoulders. The exclusion criterion was: underlying pathology (such as trauma (fractures), infection, inflammatory disorders, neurological pathology or systemic disease) [10]. To further rule out cervical radiculopathy, the upper limb tension test (ULTT) had to be negative [11]. Eligible subjects were so informed and invited to participate in the study. On agreement, subjects gave written informed consent prior to data collection.

Data collection

Reference test

The reference test was the Cervical Range of Motion (CROM) device, a valid measurement device of the ROM of the cervical spine [12]. A systematic review rated the concurrent validity positively if the correlation coefficient was above 0.65. Construct validity of the CROM was determined by comparing CROM with radiographics, and an optoelectronic system. The Pearson correlations with radiographics were for flexion 0.97, extension 0.98, lateral flexion left 0.82 and right 0.84. The Pearson correlations with an optoelectronic system were for flexion 0.98, extension 0.99, rotation right 0.89, rotation left 0.94, lateral flexion right 0.91 and left 0.89. Furthermore, the CROM device has high inter- and intra-rater reliability in patients with non-specific neck pain [13]. Finally, the standard error of measurement for flexion was 2.8°, extension 4°, rotation right 2.4°, rotation left 2.3°, lateral flexion right 2.5° and left 2.5° [14]. This makes the CROM one of the most valid and reliable measurement devices for determining the ROM of the cervical spine in daily physiotherapy practice.

The CROM measurement procedure was as follows: subjects were positioned sitting on a chair with a backrest and a seat height of 45 cm, with both feet were resting on the ground. Subjects were instructed on how to perform the movements of interest (e.g. without compensation from the thoracic spine; remaining in contact with the chair) and were asked to slowly move the neck to what they felt to be the end of their range of motion. Participants were asked to perform these movements once prior to the test as a warm-up and to ensure they understood the instructions. The CROM device was calibrated to a “zero” starting position prior to the test. Participants performed flexion, extension, rotation right and left, lateral flexion right and left once each (in a 2-dimensional movement plane) in random order (using a computerized randomization program). End range positions were held for 5 seconds, and the range of motion was recorded in whole degrees for each movement (see link below for an exact demonstration).

Index tests

An index test is a diagnostic test that is evaluated against a reference standard test. One of these, recommended by the Dutch guidelines on neck pain, concerns self-reported restriction in cervical ROM, with perceived limitations being inquired about during history-taking. In the Netherlands, physiotherapists quantify patients’ experienced functional restrictions using a Patient-Specific Functional Scale [7]. Subjects were asked about their perceived cervical ROM individually for each direction, namely flexion, extension, rotation (left/right) and lateral flexion (left/right). First, the patient was asked to perform the movement as far as they could, with the physiotherapist demonstrating if not performed correctly. After performing the movement correctly at least once, subjects were asked to answer the following question in respect of each movement direction: “To what extent do you feel restricted in moving your neck?”, using a 0-10 numerical rating scale (NRS) (0 “no restriction” to 10 “fully restricted”). For our main analyses, we dichotomized the results of this self-report test to “not limited” (0) or “limited” (1 to 10).

There is no consensus on which physical tests can best be used to determine limitations in cervical ROM. Through a Delphi-survey among experts, eight groups of physical examination tests have been suggested for assessing the presence of an articular dysfunction [15] which can be used for this purpose. These tests are also commonly used by the participating physiotherapists in their diagnostic practice.

Despite already performing these tests regularly, the participating physiotherapists practised them for a total of five hours. The aim of this was to reach consensus on the protocol, the performance of the CROM test, and the performance of the physical examination tests.

The following eight tests [15] were investigated:

1. Six unidirectional movement tests performed in sitting position: active flexion, extension, left/right rotation and left/right lateral flexion. The result is positive when unilateral compression pain and/or stretch pain occurs during active movement.
2. Four tests in sitting position performed to determine restriction in ROM and the end-feel (the sensation felt by the physiotherapist at the end of each movement) during passive combined movement testing: three-dimensional (3D) flexion/extension with rotation and lateral flexion in the same left and right direction. The result is positive when an aberrant end-feel is sensed by the physiotherapist. An aberrant hard end-feel is defined as "*a firm, abrupt end of motion…when passively moving the joint to its end range*" [16]. The result is also positive when a decreased ROM is detected by the physiotherapist.
3. Hold-relax test, performed in a sitting position, for only the most limited 3D movement as rated by the physiotherapist. A hold-relax test for all four 3D movements was considered too stressful for the patient. The result is positive if relaxation of relevant myofascial structures does **not** result in an increased passive ROM.
4. Traction test performed in supine position with neck in neutral position. The result is positive if traction reduces pain or other symptoms in the neck region.
5. Palpation test for muscle tension performed in supine position with neck in neutral position. The result is positive if increased muscle tension is identified by the physiotherapist in the dorsal neck region.
6. Myofascial pain provocation by palpation test performed in supine position with neck in neutral position. The result is positive when pain is provoked by palpating for pressure soreness in the dorsal neck region.
7. End-feel test in neutral position performed in supine position. The physiotherapist places one hand under the neck and observes the end-feel for each segment with pressure against the spinal process. The result is positive when a restriction of intervertebral movement at an impaired segment (experienced by the physiotherapist as resistance or a barrier to further motion) is identified.
8. Unilateral Posterior-Anterior (UPA) provocation test performed in prone position with neck in neutral position. The result is positive when pain or other symptom provocation occurs in the entire neck region.

The physical examination tests were performed in several directions: 1) the same as in the reference test; 2) in 3D directions; or 3) in neutral position. If tests were performed in the same direction as the reference test, then this direction was used to determine diagnostic accuracy. A 3D movement has to be resolved into three 2D movements to enable evaluation of the 3D test. We argue, based on the coupling behaviour of the cervical spine [17], that a limited 3D movement will manifest itself as a limitation in at least one of the three 2D planes that make up the 3D movement. For example: the 3D reference test: flexion - left rotation - left lateral flexion was compared with each individual 2D plane on the index test, that is, flexion, left rotation and left lateral flexion. The neutral position physical tests were compared with all anatomical planes as assessed by the index test because they are not specific to one direction of motion.

The reference test (during physical examination) was performed minutes after the self-report index test (during history-taking) and before the physical examination tests. The tests were performed in the order described, except for the end-feel test in neutral position. This was performed before the myofascial provocation test to preclude any influence of the latter on its outcome. The UPA was performed when no more pain was experienced from the myofascial provocation test. The physiotherapists were not aware of the pooled normative values prior to the reference test.

Online data collection system

An online data collection system named "lime survey" was used (https://community.limesurvey.org/licence-trademark/). Using "lime survey" guarantees non-traceable personal data from the subjects, in compliance with European Privacy laws. If there was insufficient time within the physiotherapy process, subjects could complete the baseline characteristics themselves in an online digital form of the digital database "lime survey".

Data collection process

Five physiotherapists (PT) collected data in six primary care physiotherapy practices in the Netherlands. Three were in their final year of an advanced physiotherapy MSc programme in Orthopaedic Manual Therapy). These three had a mean work experience of 3 years. The two others were also manual therapists (MT) (MSc) (work experience: 13 years PT/10 years MT and 30 years PT/23 years MT).

Data collection took place during the physical therapeutic diagnostic usual care process of history-taking and physical examination.

Baseline characteristics were: gender, age, education, neck pain during the last 24 hours measured with a numeric rating scale (NRS), and the Neck Disability Index (NDI). The NRS was used to capture the participant's level of pain intensity [18] over the last 24 hours using a scale ranging from 0 (no pain) to 10 (worst pain imaginable). This NRS is more efficient in an online data collection system than a visual analogue scale (VAS), with patients seeming to find the VAS more difficult to understand [19].

The NDI is a valid questionnaire to measure disability, with a total score ranging from 0 (no disability) to 50 (maximum disability) [20].

The outcomes of the index and reference tests were recorded during the diagnostic process by the physiotherapist. For transparency purposes, we filmed the tests. An exact demonstration of the index and reference tests can be seen via the following link: https://www.youtube.com/watch?v=qIvk0j1eZoI&feature=youtu.be.

Sample size

The sample size calculation is based on Buderer's formula for sensitivity and specificity of diagnostic health studies [21]. For a sensitivity and specificity of 0.7, a prevalence of 0.3, a precision of 0.15 and a 95% confidence level, a sample of 121 subjects was needed.

Cut-off point

A recent systematic review presented pooled normative data, stratified by age category [22]. When the ROM is less than the norm value, it could be considered limited. However, this limited ROM may also be the result of the normal anatomical variation and therefore may not reflect a true limitation in ROM. Thus, there is an overlap between the distribution of ROM measurements in people with and without non-specific neck pain. The cervical ROM was considered limited if the ROM on the reference test was less than the pooled normative value minus one standard deviation per age category.

Statistical analyses

Descriptive statistics were used for baseline characteristics of the subjects and cross-tabs. Statistical analyses were performed with SPSS Version 25 (IBM Corporation, Armonk, NY).

As recommended by the STARD statement [23], diagnostic accuracy was determined by calculating sensitivity, specificity, positive and negative predictive values (PPV, NPV), positive and negative likelihood ratios (LR +, LR-) and their 95% confidence intervals (CI) and the "area under the curve" (AUC) [24]. A LR+ between 1 and 2 is considered small, 2-5 small but sometimes important, 5-10 moderate and >10 large [25]. An LR- between 0.5 and 1 is considered small and rarely important, 0.5-0.2 small but sometimes important, 0.1-0.2 moderate and < 0.1 large (24). An AUC of ≤ 0.75 was considered as not clinically useful [26].

First, the diagnostic accuracy of the self-report test for each movement direction was determined. Thereafter, the diagnostic accuracy of each of the eight physical examination tests was determined separately for each direction of movement. Finally, the diagnostic accuracy of the self-report test was combined with the best test of the nine physical tests for each direction of movement.

Determining the diagnostic accuracy when combining tests, in this case a self-report test and a physical examination test, can be done in two ways, namely using either a serial or parallel testing strategy [27]. With serial testing, the second test is performed only if the result of the first test is positive. With parallel testing, both tests are performed and the results are subsequently combined. The consequence of serial testing is a higher specificity at the cost of lower sensitivity, and the opposite with parallel testing. Clinically, it means that with serial testing one can better demonstrate an activity limitation or can exclude this with parallel testing [28]. Therefore, we analysed the diagnostic accuracy of the combination of tests according to both strategies.

Sensitivity analysis

For our main analyses, we dichotomized the results of the self-report test into 'not limited' (0) and 'limited' (1 to 10). However, we hypothesized that patients might have some difficulties distinguishing between an NRS of 0, 1 or 2. Therefore, we performed a sensitivity analysis categorizing 1) NRS scores of "0 and 1", and 2) the NRS "0, 1 and 2", as 'not limited' **(Appendix 1)**.

Results

A total of 131 subjects were invited to participate in the study and 128 were enrolled. Three subjects refused to participate for privacy reasons. Female subjects numbered 83 (65%) and males 45 (35%). Ages ranged from 20 to 81 years. Other characteristics are described in **Table 1**.

Table 1: *Characteristics of enrolled patients*

		n=128	Mean	SD
Gender	Female	83 (64.8%)		
	Male	45 (35.2%)		
Age			50.8	14.9
Education	Primary education	3 (2.3%)		
	Secondary education	27 (21.1%)		
	Specialized VET programs	64 (50.0%)		
	Higher professional education	29 (22.7%)		
	Research-oriented education	5 (3.9%)		
Neck pain	NPRS		5.0	2.1
Disability	NDI		11.5	7.1

SD = standard deviation; VET = Vocational education and training; NDI = Neck disability index; NPRS = Numeric pain rating scale

The prevalence of limited cervical ROM at a cut-off point of 1σ, assessed with the CROM, was 24% for flexion, 40% for extension, 45% for left rotation, 52% for right rotation, 27% for left lateral flexion and 32% for right lateral flexion.

For the self-reported test **(Appendix 1)**, sensitivity ranged from 0.42 (95% CI 0.25-0.61) for flexion to 0.83 (95% CI 0.72-0.91) for right rotation, and specificity from 0.41 (95% CI 0.30-0.54) for left rotation to 0.62 (95% CI 0.50-0.77) for extension. LR+ ranged from 0.99 (95% CI 0.62-1.60) for flexion to 1.7 (95% CI 1.20-2.41) for extension. LR- ranged from 1.00 (95% CI 0.73-1.38) for flexion to 0.34 (95% CI 0.19-0.60) for right rotation. PPV ranged from 24% (95% CI 14-38%) for flexion to 64% (95% CI 53-74%) for right rotation, and NPV ranged from 72% (95% CI 60-82%) for extension to 87% (95% CI 73-95%) for left lateral flexion. The AUC ranged from 0.50 for flexion to 0.72 for left rotation.

When conducting the sensitivity analysis for the diagnostic accuracy of the self-report test, categorizing the sum of scores of "0 and 1" and "0, 1 and 2" **(Appendix 1)** as "not limited", we found a minimal improvement in accuracy. Overall, diagnostic accuracy decreased slightly for sensitivity, LR- and NPV, as increasing numbers of scores (0, 1 and 0, 1, 2) were considered as “not limited”. Conversely, diagnostic accuracy improved slightly for specificity, LR+ and PPV, as increasing numbers of scores (0, 1 and 0, 1, 2) were considered as “not limited” **(Appendix 1)**. It can be concluded from this sensitivity analysis that including ratings of 1or 2 in our definition of ‘limitation of the cervical spine’ has hardly any influence on the outcomes.

For the physical examination tests, the highest sensitivity values ranged from 0.71 for extension (95% CI 0.57-0.82 for end-feel 3D extension left rotation-lateral flexion and UPA) to 0.80 for right lateral flexion (95% CI 0.64-0.90 UPA) **(Appendix 2)**. The highest specificity values ranged from 0.86 for flexion (95% CI 0.77-0.92 muscle tension palpation) and left rotation (95% CI 0.75-0.93 traction test) to 0.89 for extension (95% CI 0.80-0.00 muscle tension palpation). The highest LR+ values ranged from 1.43 for left lateral flexion (95% CI 0.67-3.06 traction test) to 1.73 for right rotation (95% CI 0.79-3.79 traction test). The LR- values ranged from 0.81 for left lateral flexion (95% CI 0.46-1.41 end-feel 3D extension left rotation /lateral flexion) to 0.51 for left rotation (95% CI 0.25-0.68 UPA). The PPV ranged from 32% for flexion (95% CI 21-45% active flexion movement with unilateral compression and/or stretch pain) to 65% for right rotation (95% CI 43-83% traction test). The NPV ranged from 59% for right rotation (95% CI 42-73% UPA) to 83% for flexion (67-92% UPA).

For an overview of all outcomes regarding the diagnostic accuracy of all physical examination tests, see **Appendix 2**.

For the best combination of the self-report test and the physical test for any direction expressed in a LR+, serial testing is required. The best combination of tests for any direction of movement were the self-report test plus:

- Flexion: Active flexion movement test (1.39; 95% CI 0.67-2.88)
- Extension: Muscle tension (palpation) test (2.52; 95% CI 0.63-10.07)
- Left rotation: Traction test (2.66; 95% CI 0.98-7.20)
- Right rotation: Traction test (2.96; 95% CI 1.02-8.59)
- Left lateral flexion: Traction test (2.07; 95% CI 0.83-6.95)
- Right lateral flexion: Muscle tension (palpation) test (1.77; 95% CI 0.57-5.47)

For the best combination of the self-report test and the best physical test per direction expressed in a LR-, parallel testing is required. The best combination of tests per direction of movement were the self-report test plus:

- Flexion: Active flexion movement test (0.61; 95% CI 0.28-1.31)
- Extension: End-feel and 3D extension left rotation-lateral flexion test (0.42; 95% CI 0.17-1.06)
- Left rotation: UPA test (0.20; 95% CI 0.05-0.86)
- Right rotation: UPA test (0.23; 95% 0.07-0.77)
- Left lateral flexion: UPA test (0.19; 95% CI 0.03-1.39)
- Right lateral flexion: UPA test (0.33; 95% CI 0.08-1.38)

Table 2 further summarizes the diagnostic accuracy of the combination of the self-report test and the best physical test for any direction of movement, both for serial and parallel testing.

Table 2 *Diagnostic accuracy self-report test + physical tests*

Physical tests	Serial		Parallel	
	Sensitivity (95%CI)	Specitivity (95%CI)	Sensitivity (95%CI)	Specitivity (95%CI)
Flexion				
UPA	0.32 (0.17-0.51)	0.73 (0.63-0.82)	**0.87 (0.70-0.96)**	0.20 (0.13-0.30)
Muscle tension (palpation)	0.06 (0.00-0.11)	**0.94 (0.87-0.98)**	0.51 (0.33-0.70)	0.50 (0.39-0.60)
Extension				
End feel 3D extension left rotation-lateral flexion / UPA	0.46 (0.33-0.62)	0.76 (0.66-0.85)	**0.90 (0.79-0.97)**	0.23 (0.15-0.34)
Muscle tension (palpation)	0.10 (0.03-0.21)	**0.96 (0.89-0.99)**	0.71 (0.56-0.83)	0.55 (0.43-0.66)
Left rotation				
UPA	0.64 (0.50-0.76)	0.65 (0.53-0.77)	**0.96 (0.88-0.99)**	0.17 (0.09-0.28)
Traction test	0.19 (0.10-0.31)	**0.92 (0.84-0.98)**	0.86 (0.75-0.94)	0.35 (0.23-0.47)
Right rotation				
UPA	0.61 (0.49-0.73)	0.69 (0.56-0.80)	**0.96 (0.87-0.99)**	0.19 (0.11-0.32)
Muscle tension (palpation)	0.12 (0.05-0.22)	**0.93 (0.84-0.98)**	0.85 (0.74-0.92)	0.43 (0.30-0.56)
Left lateral flexion				
UPA	0.61 (0.42-0.76)	0.62 (0.52-0.72)	**0.96 (0.85-0.99)**	0.14 (0.08-0.24)
Muscle tension (palpation)	0.12 (0.03-0.27)	**0.93 (0.85-0.97)**	0.85 (0.70-0.95)	0.37 (0.26-0.46)
Right lateral flexion				
UPA	0.54 (0.37-0.69)	0.65 (0.55-0.75)	**0.94 (0.83-0.99)**	0.16 (0.08-0.24)
Muscle tension (palpation)	0.12 (0.04-0.26)	**0.94 (0.86-0.97)**	0.73 (0.57-0.86)	0.38 (0.28-0.49)
	Serial		**Parallel**	
	LR+ (95%CI)	LR- (95%CI)	LR+ (95%CI)	LR- (95%CI)
Flexion				
Active flexion movement with unilateral compression and/or stretch pain	**1.39 (0.67-2.88)**	0.91 (0.73-1.14)	1.19 (0.95-1.48)	**0.61 (0.28-1.31)**
UPA	1.20 (0.66-2.21)	0.93 (0.71-1.21)	1.10 (0.93-1.30)	0.63 (0.23-1.69)
Extension				
Muscle tension (palpation)	**2.52 (0.63-10.07)**	0.94 (0.85-1.04)	1.55 (1.15-2.10)	0.54 (0.34-0.86)
End feel 3d extension left rotation-lateral flexion	2.01 (1.22-3.31)	0.69 (0.52-0.92)	1.18 (1.01-1.37)	**0.42 (0.17-1.06)**

Table 2 *Diagnostic accuracy self-report test + physical tests* *(continued)*

Left rotation				
Traction test	**2.66 (0.98-7.20)**	0.87 (0.76-1.00)	1.31 (1.08-1.60)	0.40 (0.20-0.83)
UPA	1.86 (1.28-2.71)	0.55 (0.38-0.81)	1.17 (1.04-1.31)	**0.20 (0.05-0.86)**
Right rotation				
Traction test	**2.96 (1.02- 8.59)**	0.86 (0.75-0.99)	1.51 (1.19-1.91)	0.32 (0.16-0.62)
UPA	1.96 (1.29-2.99)	0.56 (0.40-0.80)	1.19 (1.04-1.36)	**0.23 (0.07-0.77)**
Left lateral flexion				
Traction test	***2.07 (0.83-0.95)***	0.89 (0.74-1.06)	1.37 (1.13-1.66)	0.32 (0.12-0.84)
UPA	1.59 (1.09-2.32)	0.64 (0.41-0.99)	1.14 (1.03-1.39)	**0.19 (0.03-1.39)**
Right lateral flexion				
Muscle tension (palpation)	**1.77 (0.57-5.47)**	0.94 (0.83-1.07)	1.18 (0.92-1.52)	0.70 (0.40-1.24)
UPA	1.56 (1.04-2.34)	0.71 (0.49-1.02)	1.12 (1.00-1.25)	**0.33 (0.08-1.38)**
	Serial		**Parallel**	
	PPV (95%CI)	NPV (95%CI)	PPV (95%CI)	NPV (95%CI)
Flexion				
Active flexion movement with unilateral compression and/or stretch pain	**31% (18%-48%)**	78% (74%-81%)	27% (23%-32%)	**84% (71%92%)**
UPA	28% (17%-41%)	77% (72%-82%)	26% (23%-29%)	84% (65%-93%)
Extension				
Muscle tension (palpation)	**63% (30%-87%)**	62% (59%-64%)	51% (43%-58%)	74% (64-82%)
End feel 3d extension left rotation-lateral flexion	57% (45%-69%)	69% (62%-74%)	44% (40%-48%)	**78% (59%-90%)**
Left rotation				
Traction test	**69% (45%-86%)**	58% (55%-62%)	52% (47%-57%)	75% (60%-86%)
UPA	55% (46%-64%)	73% (6%-80%)	44% (41%-47%)	**88% (64%-58%)**
Right rotation				
Traction test	**76% (53%-90%)**	52% (48%-55%)	62% (56%-67%)	75% (60%-85%)
UPA	68% (58%-76%)	62% (56%-73%)	56% (53%-60%)	**80% (55%-93%)**
Left lateral flexion				
Traction test	**43% (24%-65%)**	75% (72%-78%)	34% (30%-38%)	89% (76%-96%)
UPA	37% (29%-46%)	81% (73%-87%)	30% (28%-32%)	**93% (66%-99%)**
Right lateral flexion				
Muscle tension (palpation)	**45% (21%-72%)**	69% (66%-72%)	36% (30%-42%)	75% (63%-84%)
UPA	42% (33%-52%)	75% (68%-81%)	35% (32%-37%)	**87% (61%-96%)**

The diagnostic parameters as presented for each physical test are to be interpreted as the combination of the self-report test with the physical tests (either serial or parallel)
CI = confidence interval, LR = likelihood ratio, NPV = Negative predictive value, PPV = Positive predictive value
UPA = Unilateral Posterior-Anterior Provocation test,
3D = 3 dimensional.
The highest values are shown in bold numbers.

Discussion

Main results

We consider the LR+ of all individual tests to be small, and therefore rarely important. The LR- for the tests for flexion, extension and right lateral flexion were small (rarely important) and for the tests for left rotation, right rotation and left lateral flexion small but sometimes important. The AUC for the self-report test is considered not clinically useful. Based on the LRs and AUCs, the diagnostic accuracies of the self-report test and the physical examination tests separately were small and thus rarely clinical important.

A combination of the self-report test and the best physical examination test for any direction of movement had the highest diagnostic accuracy. As some LR+s were between 2 and 5 (serial testing) and some LR-s were between 0.5 and 0.2 (parallel testing), these may be clinically important. For example, the muscle tension test for extension and traction test for left and right rotation in combination with the self-report test are potentially clinically important, based on the LR+. The UPA test for left rotation and left lateral flexion in combination with the self-report test is potentially clinically important, based on the LR-.

Discussion of findings

There are many physical tests to choose from when assessing ROM. To increase clinical relevance, we wanted to base our choice on scientific research. The Delphi study by de Witte et al. [15] identified physical tests for an articular dysfunction and we assumed that there would be a relationship between articular dysfunction and limited ROM. Therefore, articular dysfunction tests should be able to signal limited ROM. In retrospect, based on our results, this assumption turned out to be incorrect.
The cluster of tests we examined included a non-specific test to determine ROM, namely the traction test. This may adversely affected diagnostic accuracy of this test. Strikingly, this non-specific test worked best in combination with self-report to identify a cervical ROM restriction.

Although patients with neck pain have significantly limited ROM, compared to those without [4], the underlying cause of this limitation may well be something other than articular dysfunction. For example, possible causes of restricted movement could be myofascial in origin [29], exacerbated by psychological factors [30]. This could explain why the physical examination tests for articular dysfunction that we investigated did not accurately diagnose limitations in ROM.

Because there are multiple underlying causes for limited ROM, it could be argued that an investigation such as ours might work better in a homogeneous population with one cause for limited ROM. However, it is not clear how such a population could be identified. One could argue that using treatment-based classification systems might achieve this but we are of the opinion that using such systems would not help to select more homogenous populations since their methodological quality is predominantly low and they do not identify specific causes for limited ROMs [31].

Pain intensity might also be thought to be an underlying cause of a restricted ROM. However, this does not seem to be the case since the correlation between pain intensity and perceived restriction or between pain intensity and a measured restriction in ROM, using the CROM, was small ($r < 0.30$) for all movement directions of the cervical spine.

Strength and limitations

To determine the existence of a true limitation in cervical ROM, it is necessary to establish an optimal cut-off point for a restriction, as opposed to normal variation [23,32]. A limitation of our study is that there are no well-defined cut-off points for classifying a true restriction in cervical ROM. On the other hand, we know that ROM changes with age so the cut-off points have been stratified by age [33]. We used the one standard deviation (1σ) cut-off point as we considered this the most appropriate, but this usage can be debated. With cut-off points of 2σ or 3σ for a limitation in cervical ROM, the prevalence decreased dramatically, (for 2σ to below 31% and for 3σ to below 9%) **(see Appendix 1)**. As, according to Stenneberg et al. [4], significant differences were found in active ROM between subjects with non-specific neck pain and those without, these cut-off points no longer seem to represent clinical practice. In addition, this decrease in prevalence results in an increase in false negatives (decrease in specificity see **Appendix 1, 2**). This means that a cut-off point of 2 or 3σ leads to an increase in people who are incorrectly classified as not restricted and would therefore not be treated. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group indicates that, for relatively safe treatments, high sensitivity is preferable to high specificity [34]. The last argument for choosing 1σ was that the accuracy hardly increased with higher cut-offs because of the small differences in LR between 1, 2 and 3σ **(see Appendices 1, 2)**. For the sake of interest, all three cut-off point calculations are reported in **Appendices 1 and 2**. We believe, for this study, that using the 1σ cut-off point is the most appropriate but further research will be needed to determine whether this is sufficiently accurate [34].

A second limitation was that the measurements of the index and reference tests were not technically independent so that the results may have been overestimated [9,23]. For our study, this means that, if the outcomes of this research were an overestimation, the actual outcomes are even worse. This procedure was unavoidable because data collection took place during the routine physiotherapy process of usual care. A first measure aimed at reducing information bias was to randomize the order of directions in which CROM measurements were performed. As a consequence, these CROM measurements were always taken in a different order than in the self-report test or physical examination tests. In addition, the correlation between the physical therapist's interpretation of the outcome of a CROM measurement and the objective limitation of ROM is moderate. This moderate correlation reduces the chance of overestimation as a result of information bias [35]. Therefore, we made sure that the physiotherapists were not aware of the cut-off points because without knowing those, the outcome of CROM measurements are difficult or impossible to interpret. The interpretation of the CROM measures (limited yes / no according to the cut-off point) was performed by an independent researcher to further reduce information bias.

Finally, the hold-relax test was assessed in the most limited of the four 3D directions, as determined by the physiotherapists. However, this direction was not recorded and therefore we were unable to link the hold-relax test to the relevant anatomical planes. As a result, the diagnostic accuracy for this test could not be determined.
A strength of this study was the practice-oriented setting where real-time results were obtained during the normal diagnostic process of physiotherapists, meaning that these results reflect daily practice

We followed the STARD statement recommendations concerning a rationale for test positivity cut-offs of the reference standard and the use of the recommended estimates of diagnostic accuracy (such as sensitivity, specificity, predictive values, or Area Under the Curve values) and their precision (such as 95% CIs) [23].

External validity
There are only three studies on diagnostic accuracy of International Classification of Functioning (ICF) related physiotherapeutic variables (36-38). One study describes the accuracy of the manual diagnosis for cervical zygapophysial joint pain syndromes (37) and a one study describes the diagnostic accuracy of the cervical flexion–rotation test (36) and one study describes the diagnostic accuracy of joint position error (38). This lack of diagnostic accuracy research confirms Verhagen's statement: "Unfortunately, little is known regarding the diagnostic value of general physical examination for patients with neck pain" [39].

Implications

First, it should be stressed that measuring ROM is not the same as determining a limitation in ROM. To determine this, a cut-off point must be used as otherwise we simply know what the ROM is, not whether it is limited.

The traction and UPA tests, in combination with the self-report test, have some clinical relevance but have not been shown useful for all movement directions. Nevertheless, they can be used in daily practice to get an impression of whether cervical ROM is restricted or not, especially if no CROM is available. However, if the results of this study are considered to overestimate the true outcomes, the use of these tests should be advised with caution. Therefore, if a CROM device is available, its use to determine a restriction in ROM, using a cut-off point of 1σ, is preferable.

For instance, the CROM measurement could be performed first to determine whether the movement is restricted or not, using the cut-off point. Then a physical test could be performed to determine the possible underlying cause of the restriction in ROM. Unfortunately, insufficient research has been carried out to determine which tests test valid underlying causes.

This is a first study of the diagnostic accuracy of self-report and physical examination tests, or a combination of these, to demonstrate or rule out a restriction of movement of the neck. As it is impossible to examine every physical examination test of the cervical spine in one study, we have looked at those recommended by experts. Other physical examination tests might possibly perform better, perhaps with a more appropriate subgroup of subjects. Therefore, further research into the diagnostic accuracy of physical examination tests is needed.

Conclusion

A first insight has been obtained into the diagnostic accuracy of determining a restricted ROM of the neck during the entire diagnostic physiotherapy process in a usual care setting, consisting of history-taking and physical examination tests.

The overall diagnostic accuracy of the self-report test or the best physical examination tests to assess the restriction of ROM in subjects with non-specific neck pain seems limited. We consider the LRs of the combination of the self-report test and the best physical tests to be small but sometimes clinically important. Therefore, we advise the use the CROM device to determine a restricted ROM of the neck.

Acknowledgements
We want to thank Melissa Baltussen, Ilona Boers and Tamara van Meeteren for their help in collecting the data.
We want to thank Les Hearn for scientific proofreading and English editing (les_hearn@yahoo.co.uk)

References

1. Hoy D, March L, Woolf A, Blyth F, Brooks P, Smith E, Vos T, Barendregt J, Blore J, Murray C, Burstein, R., Buchbinder R. The global burden of neck pain: Estimates from the global burden of disease 2010 study. Annals of the Rheumatic Diseases 2014; 73: 1309-1315.
2. Hurwitz EL, Randhawa K, Yu H, Cote P, Haldeman S. The global spine care initiative: A summary of the global burden of low back and neck pain studies. European Spine Journal 2018; 27: 796-801.
3. Lee H, Nicholson LL, Adams RD. Cervical range of motion associations with subclinical neck pain. Spine 2004; 29: 33-40.
4. tenneberg MS, Rood M, de Bie R, Schmitt MA, Cattrysse E, Scholten-Peeters GG. To what degree does active cervical range of motion differ between patients with neck pain, patients with whiplash, and those without neck pain? A systematic review and meta-analysis. Archives of Physical Medicine and Rehabilitation 2017; 98: 1407-1434.
5. Pool JJM, Maissan F, Waele de N, Wittink H, Ostelo RWJG. Completeness of the description of manipulation and mobilisation techniques in randomized controlled trials in neck pain; a review using the TiDieR checklist. Musculoskeletal Science Practice 2020; 45: 102098.
6. Nordin M, Carragee EJ, Hogg-Johnson S, Weiner SS, Hurwitz EL, Peloso PM, Guzman J, van der Velde G, Carroll J, Holm LW, Cote, P., Cassidy, D., Haldeman, S. Assessment of neck pain and its associated disorders: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Spine 2008; 33: S101-22.
7. Bier JD, Scholten-Peeters WGM, Staal JB, Pool J, van Tulder MW, Beekman E, Knoop J, Meerhoff G, Verhagen AP. Clinical practice guideline for physical therapy assessment and treatment in patients with nonspecific neck pain. Physical Therapy 2018; 98: 162-171.
8. Blanpied PR, Gross AR, Elliott JM, Devaney LL, Clewley D, Walton DM, Sparks C, Robertson EK. Neck pain: Revision. The Journal of Orthopaedic and Sports Physical Therapy 2017; 47: A1-A83.
9. ossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HCW, Lijmer JG. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. Clinical Chemistry 2003; 49: 7-18.
10. Hogg-Johnson, S., van der Velde, G., Carroll, L. J., Holm, L. W., Cassidy, J. D., Guzman, J., Cote, P., Haldeman, S., Ammendolia, C., Carragee, E., Hurwitz, E., Nordin, M., Peloso, P. 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 2008; 33: S39-51.
11. Thoomes EJ, van Geest S, van der Windt DA, Falla D, Verhagen AP, Koes BW, Thoomes-de Graaf M, Kuijper B, Scholten-Peeters WGM, Vleggeert-Lankamp CL. Value of physical tests

6

in diagnosing cervical radiculopathy: A systematic review. The Spine Journal 2018; 18: 179-189.

12. de Koning CH, van den Heuvel SP, Staal JB, Smits-Engelsman BC, Hendriks EJ. Clinimetric evaluation of active range of motion measures in patients with non-specific neck pain: A systematic review. European Spine Journal 2008; 17: 905-921.
13. Rondoni A, Rossettini G, Ristori D, Gallo F, Strobe M, Giaretta F, Battistin A, Testa M. Intrarater and inter-rater reliability of active cervical range of motion in patients with nonspecific neck pain measured with technological and common use devices: A systematic review with meta-regression. Journal of Manipulative and Physiological Therapeutics 2017; 40: 597-608.
14. Fletcher JP, Bandy WD. Intrarater reliability of CROM measurement of cervical spine active range of motion in persons with and without neck pain. The Journal of Orthopaedic and Sports Physical Therapy 2008; 38: 640-645.
15. Dewitte V, Peersman W, Danneels L, Bouche K, Roets A, Cagnie B. Subjective and clinical assessment criteria suggestive for five clinical patterns discernible in nonspecific neck pain patients. A Delphi-survey of clinical experts. Manual Therapy 2016; 26: 87-96.
16. Manning DM, Dedrick GS, Sizer PS, Brismée JM. Reliability of a seated three-dimensional passive intervertebral motion test for mobility, end-feel, and pain provocation in patients with cervicalgia. The Journal of Manual & Manipulative Therapy 2012; 20: 135-141.
17. Cook C, Hegedus E, Showalter C, Sizer PS. Coupling behavior of the cervical spine: A systematic review of the literature. Journal of Manipulative and Physiological Therapeutics 2006; 29: 570-575.
18. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. Archives of Physical Medicine and Rehabilitation 2008; 89: 69-74.
19. Gries K, Berry P, Harrington M, Crescioni M, Patel M, Rudell K, Safikhani S, Pease S, Vernon M. Literature review to assemble the evidence for response scales used in patient-reported outcome measures. Journal of Patient-Reported Outcomes 2018; 2: 41.
20. terling M, Rebbeck T. The Neck Disability Index (NDI). The Australian Journal of Physiotherapy 2005; 51: 271.
21. Zaidi SMH, Waseem HF, Ansari FA, Irfan M, Fahim S. Sample size estimation of diagnostic test studies in health sciences. Paper presented at: 14th International Conference on Statistical Sciences, Karachi, Pakistan. Proceedings 2016; 29: 239-246.
22. Thoomes-de Graaf M, Thoomes E, Fernández-de-Las-Peñas C, Plaza-Manzano G, Cleland JA. Normative values of cervical range of motion for both children and adults: A systematic review. Musculoskeletal Science and Practice 2020; 49: 102182.
23. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HCW, Kressel, H.Y., Rifai, N., Golub, R.M., Altman, D.G., Hooft, L., Korevaar, D.A., Cohen, J.F. STARD An updated list of essential items for reporting diagnostic accuracy

studies. British Medical Journal 2015; 351: h5527.

24. Chu K. An introduction to sensitivity, specificity, predictive values and likelihood ratios. Emergency Medicine 1999; 11: 175-181.
25. Jaeschke, R., Guyatt, G. H., Sackett, D. L. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. The Journal of the American Medical Association 1994; 271: 703-707.
26. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. Canadian Journal of Emergency Medicine 2006; 8: 19-20.
27. Lewis G, Sheringham J, Kalim K, Crayford T. Chapter 2C Diagnosis and Screening. In Mastering Public Health: A postgraduate guide to examinations and revalidation, 2008; pp. 185-206. London: Royal Society of Medicine Press Ltd.
28. van Stralen KJ, Stel VS, Reitsma JB, Dekker FW, Zoccali C, Jager KJ. Diagnostic methods I: Sensitivity, specificity, and other measures of accuracy. Kidney International 2009; 75: 1257-1263.
29. Haritha, P., Shanthi, C., Mashavi, K. Efficacy of post isometric relaxation versus static streching in subjects with chronic non specific neck pain. Int J Physiother 2015; 2: 1097-1102. doi:10.15621/ijphy/2015/v2i6/80774
30. Cresswell C, Galantino ML, Myezwa H. The prevalence of fear avoidance and pain catastrophising amongst patients with chronic neck pain. The South African Journal of Physiotherapy 2020; 76: 1326.
31. Maissan F, Pool J, de Raaij E, Wittink H, Ostelo R. Treatment based classification systems for patients with non-specific neck pain. A systematic review. Musculoskeletal Science and Practice 2020; 47: 102133.
32. Gogtay NJ, Thatte. UM Statistical evaluation of diagnostic tests (Part 1): Sensitivity, specificity, positive and negative predictive values. The Journal of the Association of Physicians of India 2017; 65: 80-84.
33. Pan F, Arshad R, Zander T, Reitmaier S, Schroll A, Schmidt H. The effect of age and sex on the cervical range of motion - A systematic review and meta-analysis. Journal of Biomechanics 2018; 75: 13-27.
34. Brozek JL, Akl EA, Jaeschke R, Lang DM., Bossuyt P, Glasziou P, Helfand M, Ueffing E, Alonso-Coello P, Meerpohl J, Phillips, B., Horvath, A.R., Bousquet, J., Guyatt, G.H., Schunemann, H.J. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. the GRADE approach to grading quality of evidence about diagnostic tests and strategies. Allergy 2009; 64: 1109-1116.
35. Thoomes-de Graaf M, Thoomes E, Falla D, Fernández-de-Las-Peñas C, Maissan F, Cleland JA. Does the patient and clinician perception of restricted range of cervical movement agree with the objective quantification of movement in people with neck pain? And do cli-

nicians agree in their interpretation? Musculoskeletal Science and Practice 2020; 50: 102226.
36. Hall, T. M., Briffa, K., Hopper, D., & Robinson, K. Comparative analysis and diagnostic accuracy of the cervical flexion-rotation test. The Journal of Headache and Pain 2010; 11: 391-397. Jull G, Bogduk N, Marsland A. The accuracy of manual diagnosis for cervical zygapophysial joint pain syndromes. The Medical Journal of Australia 1988; 148: 233-236.
37. Treleaven J, Jull G, Sterling M. Dizziness and unsteadiness following whiplash injury: Characteristic features and relationship with cervical joint position error. Journal of Rehabilitation Medicine 2003; 35: 36-43.
38. Verhagen AP 2021. Physiotherapy management of neck pain. Journal of Physiotherapy, 67: 5-11.

Appendix 1: *Diagnostic values self reported test*

	prevalence	sensitivity	95% CI	specificity	95% CI	PPV	95% CI	NPV	95% CI	LR+	95% CI	LR-	95% CI	ROC
flexion														
Self reported not limited = 0	24%	0.42	0.25-0.61	0.58	0.47-0.68	0.24	0.14-0.38	0.76	0.64-0.85	0.99	0.62-1.60	1.00	0.73-1.38	0.50
Not limited = 0+1	24%	0.39	0.22-0.58	0.65	0.55-0.74	0.26	0.15-0.41	0.77	0.66-0.85	1.10	0.66-1.86	0.94	0.71-1.26	
Not limited = 0+1+2	24%	0.29	0.15-0.48	0.72	0.62-0.81	0.25	0.13-0.42	0.76	0.66-0.84	1.04	0.55-1.97	0.98	0.78-1.24	
extension														
0	40%	0.65	0.50-0.77	0.62	0.50-0.77	0.53	0.40-0.66	0.72	0.60-0.82	1.70	1.20-2.41	0.57	0.39-0.84	0.67
1	40%	0.59	0.44-0.72	0.70	0.58-0.80	0.57	0.42-0.70	0.72	0.60-0.81	1.94	1.29-2.93	0.59	0.42-0.83	
2	40%	0.55	0.41-0.69	0.76	0.65-0.85	0.61	0.45-0.75	0.72	0.60-0.81	2.32	1.44-3.72	0.59	0.43-0.81	
left rotation														
0	45%	0.81	0.68-0.89	0.41	0.30-0.54	0.53	0.43-0.64	0.73	0.56-0.85	1.38	1.10-1.75	0.46	0.26-0.81	0.72
1	45%	0.81	0.68-0.90	0.50	0.38-0.62	0.57	0.46-0.68	0.76	0.61-0.87	1.62	1.24-2.11	0.38	0.22-0.66	
2	45%	0.72	0.56-0.83	0.64	0.52-0.75	0.63	0.50-0.74	0.74	0.61-0.84	2.03	1.43-2.88	0.43	0.28-0.66	
right rotation														
0	52%	0.83	0.72-0.91	0.49	0.36-0.62	0.64	0.53-0.74	0.73	0.57-0.85	1.64	1.25-2.16	0.34	0.19-0.60	0.71
1	52%	0.76	0.63-0.85	0.56	0.43-0.68	0.65	0.53-0.75	0.68	0.53-0.80	1.71	1.25-2.33	0.44	0.28-0.68	
2	52%	0.68	0.55-0.79	0.67	0.54-0.78	0.69	0.56-0.80	0.66	0.53-0.77	2.08	1.40-3.09	0.47	0.33-0.69	
left lateroflexion														
0	27%	0.83	0.66-0.93	0.42	0.32-0.53	0.35	0.25-0.46	0.87	0.73-0.95	1.42	1.13-1.80	0.41	0.19-0.87	0.66
1	27%	0.71	0.53-0.85	0.45	0.34-0.55	0.32	0.23-0.44	0.80	0.67-0.90	1.28	0.97-1.68	0.65	0.37-1.12	
2	27%	0.69	0.51-0.83	0.57	0.46-0.67	0.38	0.26-0.51	0.83	0.71-0.91	1.59	1.15-2.20	0.55	033-0.91	
right lateroflexion														
0	32%	0.68	0.52-0.81	0.43	0.32-0.54	0.36	0.26-0.48	0.74	0.59-0.85	1.19	0.90-1.57	0.76	0.46-1.20	0.57
1	32%	0.61	0.45-0.75	0.46	0.35-0.57	0.35	0.24-0.50	0.71	0.58-0.82	1.13	0.83-1.54	0.85	0.56-1.28	
2	32%	0.54	0.38-0.69	0.59	0.48-0.69	0.38	0.26-0.52	0.73	0.61-0.82	1.30	0.89-1.89	0.79	0.56-1.12	

Appendix 2: *Diagnostic values physical examination tests*

	prevalence	sensitivity	95% CI	specificity	95% CI	PPV	95% CI	NPV	95% CI	LR+	95% CI	LR-	95% CI
flexion													
Active flexion movement with unilateral compression and/or strech pain	2%	1.00	0.20-1.00	0.52	0.43-0.61	0.03	0.01-0.12	1.00	0.93-1.00	2.07	1.72-2.47		
End feel 3d flexion left rotation/lateroflexion (passive)	2%	1.00	0.20-1.00	0.36	0.28-0.45	0.02	0.01-0.09	1.00	0.90-1.00	1.56	1.37-1.78		
End feel 3d flexion right rotation/lateroflexion (passive)	2%	1.00	0.20-1.00	0.37	0.28-0.46	0.03	0.00-0.10	1.00	0.90-1.00	1.58	1.38-1.80		
Restricted ROM 3d flexion left rotation/lateroflexion (passive)	2%	1.00	0.20-1.00	0.36	0.27-0.45	0.02	0.00-0.09	1.00	0.90-1.00	1.56	1.37-1.77		
Restricted ROM 3d flexion right rotation/lateroflexion(passive)	2%	1.00	0.20-1.00	0.38	0.30-0.47	0.03	0.00-0.10	1.00	0.91-1.00	1.61	1.41-1.85		
Traction	2%	1.00	0.20-1.00	0.81	0.73-0.88	0.08	0.01-0.28	1.00	0.95-1.00	5.35	3.70-7.73		
Muscle tension (palpation)	2%	1.00	0.20-1.00	0.87	0.79-0.92	0.11	0.02-0.35	1.00	0.96-1.00	7.41	4.76-11.5		
Pain provocation by palpation	2%	1.00	0.20-1.00	0.29	0.21-0.38	0.02	0.00-0.09	1.00	0.88-1.00	1.40	1.26-1.57		
Aberrant segmental end feel left	2%	1.00	0.20-1.00	0.54	0.44-0.63	0.03	0.01-0.13	1.00	0.93-1.00	2.16	1.78-2.61		
Aberant segmental endfeel right	2%	0.50	0.03-0.97	0.50	0.41-0.59	0.02	0.00-0.10	0.98	0.90-1.00	1.01	0.25-4.08	0.99	0.24-4.03
Unilateral Posterio-Anterior Provocation test (UPA)	2%	1.00	0.20-1.00	0.33	0.25-0.42	0.02	0.00-0.09	1.00	0.89-1.00	1.49	1.32-1.68		
extension													
Active extension movement with unilateral compression and/or strech pain	4%	0.60	0.17-0.93	0.40	0.31-0.49	0.04	0.01-0.12	0.96	0.85-0.99	1.00	0.48-2.07	1.00	0.34-3.01
End feel 3d extension left rotation/lateroflexion (passive)	4%	0.60	0.17-0.93	0.33	0.25-0.42	0.04	0.01-0.11	0.95	0.89-0.99	0.90	0.44-1.86	1.20	0.40-3.62
End feel 3d extension right rotation/lateroflexion (passive)	4%	0.80	0.30-0.99	0.37	0.29-0.47	0.05	0.02-0.13	0.98	0.87-1.00	1.28	0.81-2.02	0.54	0.09-3.17
Restricted ROM 3d extension left rotation/lateroflexion (passive)	4%	0.60	0.17-0.93	0.33	0.25-0.42	0.04	0.01-0.11	0.95	0.83-0.99	0.90	0.44-1.86	1.20	0.40-3.62
Restricted ROM 3d extension right rotation/lateroflexion (passive)	4%	0.80	0.30-0.99	0.37	0.29-0.46	0.05	0.02-0.13	0.98	0.87-1.00	1.27	0.80-2.01	0.54	0.09-3.22
Traction	4%	0.20	0.01-0.70	0.82	0.73-0.88	0.04	0.00-0.24	0.96	0.90-0.99	1.09	0.18-6.56	0.98	0.63-1.52
Muscle tension (palpation)	4%	0.00	0.00-0.54	0.86	0.79-0.92	0.00	0.00-0.23	0.96	0.89-0.98			1.16	1.16-1.16

Pain provocation by palpation	4%	0.80	0.30-0.99	0.29	0.21-0.38	0.04	0.01-0.12	0.97	0.84-1.00	1.12	0.71-1.76	0.70	0.12-4.19
Aberrant segmental end feel left	4%	0.40	0.07-0.83	0.53	0.43-0.62	0.03	0.01-0.13	0.97	0.86-0.99	0.84	0.28-2.50	1.14	0.55-2.37
Aberant segmental endfeel right	4%	0.40	0.73-0.83	0.50	0.41-0.59	0.03	0.01-0.12	0.95	0.86-0.99	0.80	0.27-2.38	1.20	0.58-2.49
Unilateral Posterio-Anterior Provocation test (UPA)	4%	0.60	0.17-0.93	0.32	0.24-0.41	0.03	0.01-0.11	0.95	0.82-0.99	0.88	0.43-1.82	1.25	0.41-3.79
left rotation Active extension movement with unilateral compression and/or strech pain	7%	0.78	0.40-0.96	0.35	0.26-0.44	0.08	0.04-0.17	0.95	0.83-0.99	1.19	0.82-1.72	0.65	0.18-2.26
End feel 3d flexion left rotation/lateroflexion (passive)	7%	0.89	0.51-0.99	0.37	0.29-0.47	0.10	0.05-0.19	0.98	0.87-1.00	1.42	1.08-1.86	0.30	0.05-1.94
End feel 3d extension left rotation/lateroflexion (passive)	7%	0.67	0.58-0.74	0.34	0.25-0.43	0.07	0.03-0.15	0.93	0.80-0.98	1.00	0.62-1.62	0.99	0.38-2.58
Restricted ROM 3d flexion left rotation/lateroflexion (passive)	7%	0.89	0.51-0.99	0.37	0.28-0.46	0.10	0.05-0.19	0.98	0.87-1.00	1.41	1.08-1.85	0.30	0.05-1.96
Restricted ROM 3d extension left rotation/lateroflexion (passive)	7%	0.67	0.31-0.91	0.34	0.25-0.43	0.07	0.03-0.15	0.93	0.80-0.98	1.00	0.62-1.62	0.99	0.38-2.58
Traction	7%	0.11	0.01-0.49	0.81	0.72-0.87	0.04	0.00-0.24	0.92	0.85-0.96	0.59	0.09-3.86	1.10	0.87-1.39
Muscle tension (palpation)	7%	0.00	0.00-0.37	0.86	0.78-0.91	0.00	0.00-0.23	0.92	0.85-0.96	0.00	0.00-NA	1.17	1.16-1.16
Pain provocation by palpation	7%	0.78	0.40-0.96	0.29	0.21-0.38	0.08	0.03-0.16	0.94	0.80-0.99	1.09	0.76-1.58	0.77	0.22-2.73
Aberrant segmental end feel left	7%	0.50	0.18-0.83	0.53	0.44-0.62	0.07	0.02-0.17	0.94	0.84-0.98	1.06	0.52-2.18	0.94	0.46-1.91
Aberant segmental endfeel right	7%	0.44	0.15-0.77	0.50	0.40-0.59	0.06	0.02-0.16	0.92	0.82-0.97	0.89	0.41-1.89	1.11	0.61-2.03
Unilateral Posterio-Anterior Provocation test (UPA)	7%	0.67	0.31-0.91	0.32	0.24-0.42	0.07	0.03-0.15	0.93	0.79-0.98	0.98	0.61-1.59	1.04	0.40-2.70
right rotation Active extension movement with unilateral compression and/or strech pain	9%	1.00	0.70-1.00	0.36	0.28-0.46	0.14	0.08-0.24	1.00	0.90-1.00	1.57	1.37-1.80		
End feel 3d extension left rotation/lateroflexion (passive)	9%	0.75	0.43-0.93	0.39	0.30-0.48	0.11	0.06-0.21	0.94	0.82-0.98	1.23	0.86-1.75	0.64	0.24-1.76
End feel 3d extension right rotation/lateroflexion (passive)	9%	0.83	0.51-0.97	0.39	0.30-0.48	0.12	0.06-0.22	0.96	0.84-0.99	1.36	1.02-1.82	0.43	0.12-1.56
Restricted ROM 3d flexion right rotation/lateroflexion (passive)	9%	0.75	0.43-0.93	0.39	0.30-0.48	0.11	0.06-0.21	0.94	0.82-0.98	1.23	0.86-1.75	0.64	0.24-1.76
Restricted ROM 3d extension right rotation/lateroflexion (passive)	9%	0.83	0.51-0.97	0.38	0.30-0.48	0.12	0.06-0.22	0.96	0.84-0.99	1.35	1.01-1.81	0.44	0.12-1.59
Traction	9%	0.27	0.07-0.61	0.83	0.74-0.88	0.13	0.03-0.35	0.92	0.85-0.96	1.56	0.55-4.42	0.88	0.61-1.27
Muscle tension (palpation)	9%	0.08	0.00-0.40	0.86	0.78-0.92	0.06	0.00-0.31	0.90	0.83-0.95	0.60	0.83-0.95	1.06	0.89-1.26
Pain provocation by palpation	9%	0.83	0.51-0.97	0.30	0.22-0.39	0.11	0.06-0.20	0.94	0.80-0.94	1.18	0.89-1.56	0.56	0.15-2.08
Aberrant segmental end feel left	9%	0.5	0.22-0.78	0.53	0.44-0.63	0.10	0.04-0.22	0.91	0.81-0.96	1.07	0.59-1.94	0.94	0.53-1.68

6

Appendix 2: *Diagnostic values physical examination tests (continued)*

Aberant segmental endfeel right	9%	0.58	0.28-0.84	0.51	0.42-0.61	0.11	0.05-0.22	0.92	0.82-0.97	1.20	0.72-2.00	0.81	0.41-1.61
Unilateral Posterio-Anterior Provocation test (UPA)	9%	0.75	0.43-0.93	0.33	0.25-0.43	0.11	0.05-0.19	0.93	0.79-0.98	1.12	0.79-1.59	0.76	0.27-2.09
Left lateroflexion Active extension movement with unilateral compression and/or strech pain	2%	1.00	0.31-1.00	0.31	0.23-0.40	0.03	0.01-0.10	1.00	0.89-1.00	1.45	1.29-1.64		
End feel 3d flexion left rotation/lateroflexion (passive)	2%	1.00	0.31-1.00	0.36	0.28-0.46	0.04	0.01-0.11	1.00	0.90-1.00	1.57	1.37-1.79		
End feel 3d extension left rotation/lateroflexion (passive)	2%	1.00	0.31-1.00	0.34	0.26-0.43	0.04	0.01-0.11	1.00	0.89-1.00	1.52	1.34-1.73		
Restricted ROM 3d flexion left rotation/lateroflexion (passive)	2%	1.00	0.31-1.00	0.36	0.28-00.45	0.04	0.01-0.11	1.00	0.90-1.00	1.56	1.37-1.78		
Restricted ROM 3d extension left rotation/lateroflexion (passive)	2%	1.00	0.31-1.00	0.34	0.26-0.43	0.04	0.01-0.11	1.00	0.89-1.00	1.52	1.34-1.73		
Traction	2%	0.00	0.00-0.69	0.81	0.73-0.87	0.00	0.00-0.18	0.97	0.91-0.99			1.23	1.23-1.24
Muscle tension (palpation)	2%	0.00	0.00-0.69	0.86	0.79-0.92	0.00	0.00-0.23	0.97	0.92-0.99			1.16	1.16-1.16
Pain provocation by palpation	2%	0.67	0.13-0.98	0.28	0.21-0.37	0.02	0.00-0.09	0.97	0.84-1.00	0.93	0.41-2.08	0.97	0.84-1.00
Aberrant segmental end feel left	2%	0.67	0.13-0.98	0.53	0.44-0.62	0.03	0.01-0.12	0.99	0.91-1.00	1.43	0.63-3.25	0.63	0.13-3.14
Aberant segmental endfeel right	2%	0.33	0.02-0.87	0.50	0.41-0.59	0.02	0.00-0.10	0.97	0.88-0.99	0.67	0.13-3.34	1.33	0.59-3.02
Unilateral Posterio-Anterior Provocation test (UPA)	2%	1.00	0.31-1.00	0.33	0.250.42	0.04	0.01-0.11	1.00	0.89-1.00	1.49	1.32-1.69		
right lateroflexion Active extension movement with unilateral compression and/or stretch pain	2%	1.00	0.20-1.00	0.31	0.23-0.40	0.02	0.00-0.09	1.00	0.89-1.00	1.45	1.29-1.63		
End feel 3d flexion right rotation/lateroflexion (passive)	2%	1.00	0.20-1.00	0.38	0.30-0.47	0.03	0.00-0.10	1.00	0.94-1.00	1.62	1.41-1.85		
End feel 3d extension right rotation/lateroflexion (passive)	2%	1.00	0.20-1.00	0.37	0.29-0.46	0.03	0.00-0.10	1.00	0.91-1.00	1.59	1.39-2.83		
Restricted ROM 3d flexion right rotation/lateroflexion (passive)	2%	1.00	0.20-1.00	0.38	0.30-0.47	0.03	0.00-0.10	1.00	0.94-1.00	1.62	1.41-1.85		
Restricted ROM 3d extension right rotation/lateroflexion (passive)	2%	1.00	0.20-1.00	0.37	0.29-0.46	0.03	0.00-0.10	1.00	0.90-1.00	1.58	1.38-1.81		
Traction	2%	0.00	0.00-0.80	0.81	0.73-0.88	0.00	0.00-0.18	0.98	0.92-1.00			1.23	1.23-1.24
Muscle tension (palpation)	2%	0.00	0.00-0.80	0.87	0.79-0.92	0.00	0.00-0.23	0.98	0.93-1.00			1.16	1.15-1.16
Pain provocation by palpation	2%	1.00	0.20-1.00	0.29	0.21-0.38	0.02	0.00-0.09	1.00	0.88-1.00	1.40	1.26-1.57		
Aberrant segmental end feel left	2%	1.00	0.20-1.00	0.54	0.45-0.63	0.03	0.01-0.13	1.00	0.93-1.00	2.16	1.78-2.61		
Aberant segmental endfeel right	2%	0.50	0.03-0.97	0.50	0.41-0.59	0.02	0.00-0.10	0.98	0.91-1.00	1.01	0.25-4.08	0.99	0.24-4.03
Unilateral Posterio-Anterior Provocation test (UPA)	2%	1.00	0.20-1.00	0.33	0.25-0.42	0.02	0.00-0.09	1.00	0.89-1.00	1.49	1.32-1.68		

7

An exploratory, practice-oriented pilot study into matched treatments in patients with non-specific neck pain

François Maissan
Jan Pool
Edwin de Raaij
Marloes Thoomes-de Graaf
Paul Westers
Glenn Kroon
Harriët Wittink
Raymond Ostelo

International Journal of Physiotherapy.
Accepted.

Abstract

Objectives: The objective of this study was 1) to establish the measurement error of the Sensamove Cervical Trainer accelerometer (SCT); 2) to describe the applied treatments for patients with non-specific neck pain with an identified restriction in cervical Range of Motion (ROM) in primary care physiotherapy clinics; 3) to explore if the cervical ROM, pain, (perceived) disability and motor control improved after one manual therapy treatment.

Methods: The standard error of measurement (SEM) and the smallest detectable difference (SDD) were calculated, based on a test-retest study. Second, an explorative, longitudinal study design (follow-up one week) was performed. Inclusion criterion: patients with non-specific neck pain with an identified restriction in cervical ROM. Measurements: pre- (T0) and post treatment (T1) and one week post-treatment (T2). Outcomes: ROM, motor control movement task, Numerical Pain Rating Scale (NPRS) and Patient Specific Function Scale (PSFS).

Results: The SEM varied from 1.62° (lateral flexion right) to 3.46° (extension). The SDD varied from 4.49° (lateral flexion right) to 9.58° (extension). Four physiotherapists included 24 patients and used eight different treatments. The T0-T2 improvement in cervical ROM ranged from 2.95° (SD 6.09) (right lateral flexion) to 11.00° (SD11.87) (left rotation). The movement task was performed 3.96 (SD 4.24) seconds faster. The NPRS decreased 3.08 (SD 1.82) points and PSFS improved 7.71 (SD 5.34) points.

Conclusion: The measurement error has been established. Moreover, this study illustrates that matched treatments, as applied in daily practice, have the potential to induce short-term improvements.

Keywords: neck pain, range of motion, physical therapy modality

Introduction

Non-specific neck pain is a major concern in the adult Western world population and the 12-month prevalence ranges between 30% to 50% [1]. Often a specific diagnosis cannot be made, and neck pain is labelled non-specific, because of the multifactorial aetiology [1].

Physiotherapy interventions for non-specific neck pain have repeatedly been investigated but the results are inconclusive [2-5]. A potential explanation is that treatments are according to a one-size fits all principle, and therefore, recently, "physiotherapeutic validity" has emerged as an important topic. This is defined as a match between the identified impairments (e.g. restricted Range of Motion) and/or activity limitations (e.g. looking backward while driving a car) and specific treatments aiming to improve these impairments and/or activity limitations, with matching outcome measures (i.e. relevant outcome measures linked to the aim of the treatment) [6-8]. This match is important as the clinical reasoning process is a pre-requisite for choosing the most optimal treatment [9]. A recent review assessed the "physiotherapeutic validity" of Randomized Controlled Trials (RCTs) for patients with non-specific neck pain. Only 9% of the 122 included studies had adequate "physiotherapeutic validity" [6].

It is generally believed that the most investigated interventions, mobilisations and manipulations, can improve Range of Motion (ROM) in patients with non-specific neck pain if there is a valid indication for those interventions [7,10]. The reported effects of mobilisations and/or manipulations are small, but they are reported to be more effective when combined with exercise therapy [11,12]. It is unknown which treatment parameters (e.g. the segmental level) of the mobilizations or the manipulations give the best result [13]. Additionally, it has been argued that other interventions also have the potential to improve ROM; e.g exercise therapy [12], hold-relax techniques [14] and pain education [15]. This suggest that restricted ROM may be associated with a variety of factors (e.g. joint, muscle or psychological factors). It therefore remains unclear which interventions or combination of interventions have the greatest potential to improve ROM.

Little is known to which extent changes in other variables, such as pain and/or disability, occur when ROM improves in patients with non-specific neck pain with a restriction of ROM of the neck. To date only one study, which included non-specific neck pain patients with a restriction of ROM of the neck, investigated whether improved ROM was associated with decreased pain intensity [16]. This randomized controlled trial (RCT) compared mobilizations with a motionless manual contact placebo treatment. Mobilization significantly increased ROM compared to the placebo treatment (MD lateral flexion 5.2°(95% CI: 1.84-8.56); MD rotation 4.8° (95% CI: 0.32-9.28) and the difference in pain decrease ranged from 29 to 47% in favour of the intervention group [16].

An improvement of ROM could, in addition to changes in pain and disability, also induce an improvement in motor control, defined as the way in which the nervous system controls posture and movement to perform a specific mo¬tor task, and includes consideration of all the associated motor, sensory, and integrative processes [17]. One study investigated if, in addition to changes in ROM, a simple rotation task of the cervical column also changed, after one treatment with spinal manipulation (SM)[18]. Right rotation varied statistically significantly from 74.75° (SD 7.63°) pre-SM to 78.50° (7.23°) post-SM. No other ROM directions or conditions yielded significant differences. The results of the rotation task showed that the precision of the execution of the rotation task also improved. So, preliminary results seem to suggest that motor control improves after an improved ROM of the neck.

To strengthen physiotherapeutic validity in scientific research, effects of physiotherapy should be investigated in more practice-oriented studies [19]. Practice-oriented does not only mean a physiotherapeutic practice setting, but also that the physiotherapist is free to act in accordance with their normal daily clinical practice.
A physiotherapy treatment generally combines multiple interventions, for example mobilizations with exercise therapy, and is therefore multimodal in nature [20]. In daily practice, physiotherapists choose their own treatment based on their clinical reasoning process. This individualisation of treatment based on the patient's specific needs is called matched care [21]. Despite the fact that physiotherapists attempt matching care to the patient’s needs, there is little knowledge about which treatments are applied in daily practice and which treatments have the most potential to improve ROM in patients with non-specific neck pain with a limited ROM. For the measurement of the cervical ROM we used the Sensamove Cervical Trainer accelerometer (SCT). Before the results of this instrument can be clinically interpreted, insight into the reliability and measurement error of the SCT is necessary [22].

Therefore, the first aim was to establish the measurement error of the SCT. The second aim was to describe the treatments applied by physiotherapists in daily practice, for patients with non-specific neck pain with an identified restriction in cervical ROM. The third aim was to explore if the cervical ROM, pain, (perceived) disability and motor control improved after one matched treatment.

Methods

First, a test-retest design was used to calculate the reproducibility of the SCT. The Medical Ethic Center in Rotterdam approved this part of the study (MEC-2018-129). Second, an explorative prospective, longitudinal pilot study with a follow-up of one week was executed. This study was approved (reference number 96_000_2019) by the Institutional Review Board (department of health studies) of HU University of Applied Sciences Utrecht. Participation was voluntary and written informed consent was obtained. Patients were included from February to May 2019.

Participants

For the test-retest design patients had to meet the following inclusion criteria:

- age >18 years and non-specific neck pain, defined as pain (with or without radiation) located in the cervical spine and/or occiput region and/or cervico thoracic junction and muscles originating from the cervical region acting on the head and shoulders, without underlying pathology [23].
- Proficient in Dutch language.
- to rule out cervical radiculopathy, the upper limb tension test had to be negative [24].

For the pilot study consecutive participants were recruited from three primary care physiotherapy clinics between February 2020 and October 2020. Before participating, participants signed informed consent. The inclusion criteria were identical to the test-retest study, with two additional criteria for the pilot study were:

- a confirmed movement restriction in left and/or right rotation as measured with the SCT. To reduce participant burden, the other directions were measured only if a restriction was found in the left and/or right rotation direction. The ROM was considered restricted if the ROM was less than the pooled norm value minus one standard deviation per age category [25].
- The mandatory primary treatment target of the first treatment was improvement of ROM of the neck.

Physiotherapists

For the test-retest study the measurements were performed by one physiotherapist with five years of work experience. Repeated measurements by this rater on the same day were used to calculate intra-rater reliability.

For the pilot study a convenience sample of four physiotherapists, mean age 39.75 (SD 13.2) was invited to collected data in three primary care physiotherapy practices in the Netherlands (two physiotherapists worked in the same practice). One physiotherapist was in his final year of a 3-year master Orthopaedic Manual Therapy programme. The other physiotherapists were registered manual therapists (MSc). The average work experience as a physiotherapist was 16.25 (SD 13.48) years and as a manual therapist 11.00 (SD 12.06) years.

The physiotherapists were invited because they owned a Sensamove Cervical Trainer (SCT). This is not part of the standard equipment in Dutch physiotherapy practices. The SCT 3D sensor (www.sensamove.com/en/) is a 9 degrees of freedom sensor which combines signals from a 3-axis accelerometer, a 3-axis gyroscope and a 3-axis magnetometer and then translates the 9 separate data points into an orientation vector in x, y and z coordinates and an angle of rotation around the direction of the vector. It is positioned with an adjustable strap and aligned centrally on the forehead just above the bridge of the nose **(Figure 1)**.
The advantage of the SCT is that both the measurement of the ROM and the computerized motor task are combined in one measurement instrument, which reduces patient burden.

Figure 1: *Sensamove Cervical Trainer (SCT)*

Study protocol

Test-retest design
The strap with the accelerometer was attached to the participant's head. The SCT was calibrated before the ROM measurements, after which all movement directions (flexion, extension, rotation left/right, and lateral flexion left/right) were measured. After the measurements, the strap was removed and there was a 5-minute break. Then the strap was reattached for the second round. Recall bias was not an issue as the results of both SCT measurements were displayed in the digital output, not visible to the participant.

Pilot study
This study took place during the usual daily practice of physiotherapists. Eligible participants were asked to participate in the study by their treating physiotherapists. In order to not interfere with their daily practice, the same physiotherapists determined both the inclusion of participants and carried out the treatment.
The aim was to include a minimum of 12 participants to explore treatment effects in

this pilot study [26]. Anticipating loss to follow up we aimed to include 24 participants. During a two-hour session the study protocol was discussed with all participating physiotherapists in order to achieve that the physiotherapists used the SCT in a similar fashion. As the physiotherapists used the SCT in their daily work already, no further training was necessary.

Three measurements were done: baseline (= pre-treatment T0), immediately after treatment (T1) and after a week, before the continuation of further physiotherapy treatments (T2). Patient characteristics were measured at baseline (T0). Highest pain-intensity in the past 24 hours (Numeric Pain Rating Scale (NPRS) [27]) and perceived disability (Patient Specific Function scale (PSFS)[28]) were measured at T0 and T2. ROM of the neck was measured with an the SCT (figure 1) and the movement task performed at T0, T1 and T2. The movement task was practiced twice and was measured the third time.

After the baseline measurements (T0) treatment aimed at improving ROM was performed. The choice for the specific treatment was left to the discretion of the physiotherapists, based on the findings of their history taking and physical examination. Immediately after the treatment, the physiotherapist registered the various components of the treatment online. ROM and the computerized movement task were measured again immediately after treatment (T1). After one week (T2) it was assessed to what extent the ROM and movement task had changed, relative to T1, and to what extent the pain and experienced performance of the neck had changed, relative to T0, plus the seven-point General Perceived Effect (GPE) [29]. Only participants who were measured at all three times (T0, T1, T2) were included in this study.

The data was entered in an online database: Lime Survey (https://community.limesurvey.org/licence-trademark/), which guarantees untraceable personal data in compliance with European Privacy laws.

Outcome measures

Primary outcome

Active ROM of the neck, measured with the SCT. If a movement restriction was identified in active left and/or right rotation, the other directions of movement (flexion, extension and left/right lateral flexion) were also measured.

Secondary outcomes

A tracking task is considered an outcome measure for motor control 30 for which the SCT Neuro Muscular Control (NMC) PRO test was used at level 3 (pan view) (NMC PRO TEST - YouTube). This is a computerized movement task which can be set in such a way that the activity can also be carried out with a movement restriction. This test is focused on controlled movement; the participant starts with the cursor (red dot) on a yellow dot at one side of a predetermined pattern (3D). Once the cursor is inside the yellow dot, the yellow dot starts to move and the participant has to follow the predetermined pattern (by staying inside the yellow dot) by moving the head. If the cursor deviates from the pattern, the yellow dot stops until the participant relocates the cursor inside the yellow dot. The more often the red dot deviates from the yellow dot, the slower the activity proceeds. The test result is the time needed to complete the entire pattern. Psychometric properties of the NMC PRO test are unknown.

The 11-point NPRS captures the participant's level of pain intensity (0 = no pain; 10 (worst pain imaginable) [27] of their current pain over the last 24 hours The Smallest Detectable Difference (SDD) has been reported to be 2.1, whereas the Minimal Clinical Important difference (MCID) was shown to be 1.3, in patients with mechanical neck pain [27].

A modified PSFS was used to measure experienced disability by scoring the general activity limitations [31]. The scale was reversed ranging from 0 "unable to perform" (instead of able) to 10 "able to perform the activity" (instead of unable). The participant reports three activities that are limited and an average rating for all three activities is calculated. The original PSFS has excellent test–retest reliability (ICC 0.92) and a standard error of measure (SEM) of 0.43 for patients with neck pain [32]. The modified PSFS has an ICC of 0.95 (CI 0.92-0.97) (unpublished result). The calculated SDD of the PSFS for participants with neck dysfunction is 1.19 points [33]. The modified PSFS, preferred by Dutch participants, is valid in terms of content validity and construct validity for patients with neck pain [31].

A 7-point General Perceived Effect (GPE) was used to measure perceived recovery, ranging from 1 (fully recovered) to 7 (worse than ever). Intraclass correlation coefficient values of 0.90-0.99 indicate excellent reproducibility [29].

Data analysis

For the test-retest study the following patient's characteristics were described; gender, age, duration of complaints, neck pain intensity (NPRS) and experienced disability with the neck disability index (NDI). To determine a clinically relevant difference,

SEM and the SDD were calculated [36].
For the pilot study the raw quantitative data was transferred from Lime Survey to SPSS. Descriptive statistics were used for baseline characteristics: age, gender, duration of neck pain, neck pain intensity (NPRS), experienced disability (PSFS), ROM, and the NMC PRO test.

The changes in cervical ROM, movement task, pain, and experienced activity limitation, are presented in means and standard deviation (SD) (significance level 5%). Since the NPRS and PSFS were measured twice (T0-T2) the paired samples T-Test was used. For the cervical ROM and movement task a repeated measures Anova was used including (T0, T1 and T2). All analyses were performed with SPSS Version 25 (IBM Corporation, Armonk, NY). Data available on request from the authors.

Results

Test-retest study

Of the 33 consecutive participants who met the inclusion criteria, 31 participated including 15 men, with a mean age of 52.6 (SD 18.8) years, mean duration of complaints 69.2 (SD 96.5) weeks, mean NPRS score 4.9 (SD 1.8), and mean NDI score 23.4 (SD 12). The SEM varied from 1.62 degrees (lateral flexion right) to 3.46 degrees (extension). The SDD varied from 4.49 (lateral flexion right) to 9.58 (extension). **Table 1** presents the results for all directions.

Pilot study

Twenty-four participants were included (mean age 48 (SD 18.99) years). Nine patients had acute (0-6 weeks), 2 sub-acute (6-12 weeks) and 13 chronic neck pain (>12 weeks) [20]. **Table 2** presents all characteristics.

Eight different treatments were applied by the 4 physiotherapists **(Table 3)**. Each treatment led to an improvement in cervical ROM, especially rotation **(Table 4)**. This was as expected as patients were specifically included based on a ROM restriction of the left and/or right rotation. None of the treatments seemed superior at improving ROM. The ROM per direction of the neck (independent from the different treatments) improvement between T0 and T2 ranged from 2.95° (SD 6.09) for right lateral flexion to 11.00° (SD11.87) for left rotation **(Table 4)**. The differences between T0-T2 were all statistically significant (<0.05) except for flexion and right lateral flexion. The not statistically significant difference between T1 and T2 ranged from -1.33° (4.44) for flexion to

0.51° (6.72) for extension. The differences between T1-T2 were therefore an indication of maintenance of improvement after one week.
A result was considered to be clinically relevant if the average difference T0-T2 exceeded the measurement error (SDD) or exceeded the MCIC. The NPRS exceeded both the SDD and MCID and, the PSFS exceeded the SDD. Left and right rotation showed a clinically relevant improvement (SDD) in ROM (T0-T2).
The motor control task improved statistically significant between T0-T2 (3.96 (SD 4.24) seconds; $p<0.05$) and there was a non-significant reduction of 1.21 (SD 6.78) seconds between T1-T2. Pain decreased statistical significantly (<0.05) on average by 3.08 (SD 1.82) points on the NPRS and the activity limitations experienced by the patient improved significantly (<0.05) with 7.71 (SD 5.34) points on the PSFS. Two participants experienced a full recovery, 8 much improvement, 13 somewhat improvement and 1 patient experienced no improvement, reported on the GPE. No patient reported deterioration.

Table 1: *Measurment error*

Direction	SEM	SDD
Flexion	3.42	9.48
Extension	3.46	9.59
Left rotation	2.99	8.29
Right rotation	2.21	6.13
Left lateral flexion	2.21	6.13
Right lateral flexion	1.62	4.49

SDD = Smallest Detectable Difference, SEM = Standard Error of Measurement

Table 2: *Characteristics of enrolled participants*

n = 24	n (%)	Mean (SD)
Sex (Female)	14 (58%)	
Age (Years)		48.42 (18.99)
Acute NP (Weeks)	9 (38%)	4.11 (1.05)
Sub-acute NP (Weeks)	2 (8%)	10.50 (2.12)
Chronic NP (Weeks)	13 (54%)	126.31 (165.17)
Neck pain (NPRS)		6.71 (0.91)
Activity (PSFS)		18.25 (4.35)
DoT (Minutes)		21.63 (4.01)

DoT = duration of treatment, NP = neck pain, n = number of patients, NPRS = Numeric pain rating scale; SD = standard deviation, PSFS = patient specific function scale,
Acute neck pain 0-6 weeks – Sub acute neck pain 6-12 weeks – Chronic neck pain >12 weeks

Table 3: *Multimodal treatments*

n	Multimodal treatments
1	advice, mobilization, manipulation, triggerpoint treatment
1	advice, mobilization, manipulation, strength exercises
1	mobilization, manipulation, triggerpoint treatment
1	mobilization, manipulation, hold relax techniques
1	manipulation, triggerpoint treatment
2	advice, mobilization, motor control exercises
4	advice, mobilization, manipulation
5	mobilization, manipulation
8	advice, mobilization

n = number of patients treated

Table 4: *Range of movement per direction with differences and the computerized movement task*

Direction	n	T0	T1	T2	T0-T1	T0-T2
Flexion	24	46.52 (11.52)	50.84 (11.57)	49.51 (11.36)	4.33 (7.19)*	2.99 (6.60)
Extension	24	51.52 (15.75)	60.12 (14.28)	60.63 (15.87)	8.60 (10.32)*	9.11 (12.12)*
Left rotation	24	51.16 (12.91)	62.81 (14.80)	62.16 (15.64)	11.65(12.35)*	11.00 (11.87)*
Right rotation	24	54.47 (17.40)	63.22 (16.09)	63.19 (16.01)	8.75 6.60)*	8.72 (10.92)*
Left lateral flexion	24	31.25 (12.12)	35.82 (11.34)	34.97 (11.17)	4.57 (6.18)*	3.71 (5.94)*
Right lateral flexion	24	31.66 (13.19)	35.32 (10.97)	34.61 (11.99)	3.66 (7.34)	2.95 (6.09)
Computerized movement task	24	33.96 (13.18)	28.79 (10.02)	30.00 (13.18)	5.17 (7.43)*	3.96 (4.24)*

*Mean (Standard Deviation); n = number of subjects; T0 = pre intervention, T1 = post intervention, T2 = after 1 week; * = P< 0.05*

7

Discussion

Main results

Because there is now insight into the degree of measurement error of the SCT in the measurement of cervical ROM, the results of the ROM changes can now be interpreted clinically.

Eight treatments, all multimodal, were applied. This underpins the assumption that physiotherapy treatment generally consists of more than one intervention. The most frequently applied interventions were mobilisations and manipulations.

Participants had a clinically relevant improvement on all PROMs after only one treatment session. Left and right rotation showed a clinically relevant improvement (SDD) in cervical ROM (T0-T2), the other directions did not. This is in line with expectations since the participants were included based on a ROM rotation limitation. It can therefore be expected that the rotational limitation at baseline is greatest and therefore has the greatest chance of a clinically relevant improvement.

Discussion of findings

A range of interventions were used, however, mobilizations and manipulations were most frequently applied, in line with our expectations as it was an inclusion criterion that the first treatment should be primarily aimed at improving cervical ROM. Furthermore, three of the physiotherapists were manual therapists and one was completing manual therapy training. Therefore, the results found apply primarily to the manual therapeutic care process in patients with non-specific neck pain and restricted active rotation of the neck.

Were the one-week short-term effects on ROM found in this study with an assumed match between the identified limitation and therapy better than results in the published literature? One study[37] investigating short-term effects of manipulations on ROM after one treatment reported a mean improvement for flexion of 1.47° (our study 2.99°), rotation left 0.76° (our study 11.00°), rotation right 1.00° (our study 8.72°), left lateral flexion 1.94° (our study 3.71°) and right lateral flexion 0.65° (our study 3.96°). The extension declined by 1.94° (our study 9.11° improvement). At baseline these participants had a rotation greater than the cut-off point in our study. This means that the study population included a large proportion of patients who had no restriction in ROM, leaving no obvious room for improvement. Three RCTs included participants who had a normal ROM at baseline [38-40]. In these RCTs there was also little or no effect on ROM, even after more treatments. Only one study which did include participants with a restriction of the ROM of the neck found similar results (described in the introduction) as in this study on cervical ROM [16]. It seems important to specifically identify the specific restrictions one aims to improve with the specific treatment to achieve good results. Also the motor control task improved. What remains unclear is whether the motor control improved due to an improvement in ROM or an improvement in pain or both. This is important to understand better the impairments associated with non-specific neck pain [41,42].

Strength and limitations

In our study physiotherapists matched their treatment with their diagnostic process. However, we have no data on the outcomes of the diagnostic process and can therefore not confirm there was actually a match. Further research is needed as to why the physiotherapists applied the interventions performed, but the observation that such good short-term results were achieved by having physiotherapists match the diagnosis and treatment as they are used to in daily practice is an interesting finding.

It was difficult to include participants in this study, partly due to the Covid 19. However, the design of this study also turned out to be too time consuming within daily practice. Compensation for the extra time could potentially speed up the inclusion of participants but no financial remuneration was possible within this study.

The design doesn't allow causal inferences regarding the effectiveness of the applied interventions. In addition, interventions were used in combination, which makes it impossible to make statements about the individual components.
The psychometric properties of the NMC PRO test was unknown, which makes interpretation about clinical relevance difficult, the SDD of the SCT accelerometer was determined so that it could be determined whether the effect on the ROM was clinically relevant or not.

A strength of this study was that, in line with clinical practice, the inclusion criteria were the presence of restricted neck mobility and the primary goal of the first treatment should be to restore ROM. Therefore it can be argued this study has external validity as it is a practice-oriented study, facilitating the translation into daily practice [43,44].

Implications

The effects in this study differ considerably from the effects described in the literature. An important difference between our study and the literature concerns patient selection. Based on this observation, an important implication for researchers is to consider the selection of patients more carefully, in order to improve the match between patient characteristics and the specific treatment goals and interventions.

Although this design prevents causal inferences, the applied treatments seem worthwhile to consider in daily practice in patients with an identified restricted ROM of the neck. Which treatment is best to apply remains unclear because it has not been objectified why a therapist chose the treatment used, as this was outside the scope of this pilot study. The manual therapists applied a combination of different interventions in the same treatment, and no treatment consisted of just one intervention. This seems to confirm the assumption that physiotherapy treatment is predominantly multimodal, even if treatment is primarily aimed at one impairment (here a limited ROM).

Finally, the short time frame between the treatment and the results (T0-T1) seems to indicate that the results may be due to the treatment. In addition, the results remained fairly stable after a week.

Further research

The fact that multiple combinations of interventions are used suggests that different clinical reasoning processes are followed. Further research is needed to understand on what basis manual therapists choose different interventions. Further research is also needed to gain a better understanding of the causal relationship between the applied treatments and their effects. Another intriguing question is if improvements in physical functions lead to an improvement in objective physical activities that include the neck region? Maybe the most important question is: if the participant can perform activities better, will they also improve in their everyday movement behaviour? A last, more general issue is that future research should focus on external validity while also retaining the required internal validity, as flawed research results should not be applied at all, let alone generalized [44].

In conclusion:

The SEM and SDD of the SCT has been established. This study suggests that if manual therapists use their clinical reasoning process, in line with their routine daily practice short-term and clinically significant improvements can be achieved in patients with non-specific neck pain with a restriction of cervical ROM. Therefore, we cautiously conclude that matching the treatment to the identified impairment as performed in daily practice has potential to improve patient outcome.

Acknowledgements

We would like to thank Guido van de Wetering, Koen Kuunders and Erik Thoomes for their help with data collection.

The data that support the findings of this study are available from the corresponding author, [FM], upon reasonable request.

References

1. Haldeman S, Carroll L, Cassidy JD, Schubert J, Nygren A. Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Spine. 2008; 33: S5-7.
2. Gross A, Forget M, St George K, et al. Patient education for neck pain. Cochrane Database Syst Rev. 2012;3:CD005106.
3. Graham N, Gross A, Goldsmith CH, et al. Mechanical traction for neck pain with or without radiculopathy. Cochrane Database Syst Rev. 2008;(3):CD006408.
4. Kay TM, Gross A, Goldsmith CH, et al. Exercises for mechanical neck disorders. Cochrane Database Syst Rev. 2012;8:CD004250.
5. Monticone M., Ambrosini E., Cedraschi C., et al. Cognitive-behavioral treatment for subacute and chronic neck pain: A cochrane review. Spine. 2015;40(19):1495-1504.
6. Maissan F, Pool JJM, Raaij de E, Mollema J, Ostelo RWJG, Wittink H. The clinical reasoning process in randomised clinical trials with patients with non-specific neck pain is incomplete: A systematic review. Musculoskelet Sci Pract. 2018;35:8-17.
7. Maissan F , Pool JJM, Stutterheim E, Wittink H, Ostello R. Clinical reasoning in unimodal interventions in patients with non-specific neck pain in daily physiotherapy practice, a delphi study. Musculoskelet Sci Pract. 2018;37:8-16.
8. Maissan F, Pool J, de Raaij E, Wittink H, Ostelo R. Treatment based classification systems for patients with non-specific neck pain. A systematic review. Musculoskelet Sci Pract. 2020;47:102133.
9. Hendriks HJM, Oostendorp R.A.B, Bernards ATM, van Ravensberg CD, Heerkens YF, Nelson RM. The diagnostic process and indication for physiotherapy: A prerequisite for treatment and outcome evaluation. Phys Ther Rev. 2000;5:29-47.
10. Bialosky JE, Bishop MD, Price DD, Robinson ME, George SZ. The mechanisms of manual therapy in the treatment of musculoskeletal pain: A comprehensive model. Man Ther. 2009;14(5):531-538.
11. Gross AR, Hoving JL, Haines TA, et al. A cochrane review of manipulation and mobilization for mechanical neck disorders. Spine. 2004;29(14):1541-1548.
12. Hidalgo B, Hall T, Bossert J, Dugeny A, Cagnie B, Pitance L. The efficacy of manual therapy and exercise for treating non-specific neck pain: A systematic review. J Back Musculoskelet Rehabil. 2017;30(6):1149-1169.
13. Pool JJM, Maissan F, Waele de N, Wittink H, Ostelo RWJG. Completeness of the description of manipulation and mobilisation techniques in randomized controlled trials in neck pain; a review using the TiDieR checklist. Musculoskelet Sci Pract 2020. 45.
14. Haritha P, Shanthi C, Mashavi K. Efficacy of post isometric relaxation versus static streching in subjects with chronic non specific neck pain. Int J Physiother. 2015;2(6):1097-1102.

15. Cresswell C, Galantino ML, Myezwa H. The prevalence of fear avoidance and pain catastrophising amongst patients with chronic neck pain. S Afr J Physiother. 2020;76(1):1326.
16. Lascurain-Aguirrebeña I, Newham DJ, Casado-Zumeta X, Lertxundi A, Critchley DJ. Immediate effects of cervical mobilisations on global perceived effect, movement associated pain and neck kinematics in patients with non-specific neck pain. A double blind placebo randomised controlled trial. Musculoskelet Sci Pract. 2018;38:83-90.
17. van Dieën JH, Reeves NP, Kawchuk G, van Dillen LR, Hodges PW. Motor control changes in low back pain: Divergence in presentations and mechanisms. J Orthop Sports Phys Ther. 2019;49(6):370-379.
18. Passmore SR, Burke JR, Good C, Lyons JL, Dunn AS. Spinal manipulation impacts cervical spine movement and fitts' task performance: A single-blind randomized before-after trial. J Manipulative Physiol Ther. 2010;33(3):189-192.
19. Tsakitzidis G, Remmen R, Dankaerts W, van Royen P. Non-specific neck pain and evidence-based practice. ESJ. 2013;9(3):1-19.
20. Blanpied PR, Gross AR, Elliott JM, et al. Neck pain: Revision 2017. J Orthop Sports Phys Ther. 2017;47(7):A1-A83.
21. Linton SJ, Nicholas M, Shaw W. Why wait to address high-risk cases of acute low back pain? A comparison of stepped, stratified, and matched care. Pain. 2018;159(12):2437-2441.
22. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol. 2006;59(10):1033-1039.
23. 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. Spine. 2008;33:S39-51.
24. Thoomes EJ, van Geest S, van der Windt DA, et al. Value of physical tests in diagnosing cervical radiculopathy: A systematic review. Spine J. 2018;18(1):179-189.
25. Thoomes-de Graaf M, Thoomes E, Fernández-de-Las-Peñas C, Plaza-Manzano G, Cleland JA. Normative values of cervical range of motion for both children and adults: A systematic review. Musculoskelet Sci Pract. 2020;49:102182. doi: S2468-7812(20)30094-1.
26. Moore CG, Carter RE, Nietert PJ, Stewart PW. Recommendations for planning pilot studies in clinical and translational research. Clin Transl Sci. 2011;4(5):332-337.
27. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the neck disability index and numeric pain rating scale in patients with mechanical neck pain. Arch Phys Med Rehabil. 2008;89(1):69-74.
28. Stratford PW, Gill C, Westaway MD, Binkley JM. Assessing disability and change on individual patients: A report of a patient specific measure. Physiother Can. 1995;47:258-262.
29. Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ. Global perceived effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol.

2010;63(7):760-766.e1.

30. Willigenburg NW, Kingma I, Hoozemans MJ, van Dieën JH. Precision control of trunk movement in low back pain patients. Hum Mov Sci. 2013;32(1):228-239.
31. Thoomes-de Graaf M, Fernández-De-Las-Peñas C, Cleland JA. The content and construct validity of the modified patient specific functional scale (PSFS 2.0) in individuals with neck pain. J Man Manip Ther. 2020;28(1):49-59.
32. Westaway MD, Stratford PW, Binkley JM. The patient-specific functional scale: Validation of its use in persons with neck dysfunction. J Orthop Sports Phys Ther. 1998;27(5):331-338.
33. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. Qual Life Res. 2001;10(7):571-578.
34. Portney LG, Watkins MP. Foundations of clinical research. second edition ed. New Jersey: Prentice Hall Health; 2000.
35. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155-163.
36. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine - A practical guide.. New York: Cambridge University Press; 2011.
37. Hanney WJ, Puentedura EJ, Kolber MJ, Liu X, Pabian PS, Cheatham SW. The immediate effects of manual stretching and cervicothoracic junction manipulation on cervical range of motion and upper trapezius pressure pain thresholds. J Back Musculoskelet Rehabil. 2017;30(5):1005-1013.
38. Hakkinen A, Salo P, Tarvainen U, Wiren K, Ylinen J. Effect of manual therapy and stretching on neck muscle strength and mobility in chronic neck pain. J Rehabil Med. 2007;39(7):575-579.
39. Kanlayanaphotporn R, Chiradejnant A, Vachalathiti R. The immediate effects of mobilization technique on pain and range of motion in patients presenting with unilateral neck pain: A randomized controlled trial. Arch Phys Med Rehabil. 2009;90(2):187-192.
40. Kanlayanaphotporn R, Chiradejnant A, Vachalathiti R. Immediate effects of the central posteroanterior mobilization technique on pain and range of motion in patients with mechanical neck pain. Disabil Rehabil. 2010;32(8):622-628.
41. Falla D, Gizzi L, Parsa H, Dieterich A, Petzke F. People with chronic neck pain walk with a stiffer spine. J Orthop Sports Phys Ther. 2017;47(4):268-277.
42. Salehi R, Rasouli O, Saadat M, Mehravar M, Negahban H, Shaterzadeh MJ. Cervical movement kinematic analysis in patients with chronic neck pain: A comparative study with healthy subjects. Musculoskelet Sci Pract. 2021.jun.
43. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials. 2017;18(1):122-017-1870-2.
44. Knottnerus JA, Tugwell P. Research methods must find ways of accommodating clinical reality, not ignoring it: The need for pragmatic trials. J Clin Epidemiol. 2017;88:1-3.

8

General discussion

General discussion

There is no substantial scientific evidence to support the contention that physiotherapy is effective for people with non-specific neck pain [1]. However, both physiotherapists and patients feel that physiotherapy is effective [2]. Could this discrepancy be due to conformation bias or is there a mismatch between the scientific literature and clinical expertise?

The general aim of this dissertation was to gain insight into the physiotherapeutic validity of physiotherapy research in subjects with non-specific neck pain. We have defined physiotherapeutic validity as the match between the diagnostic process and intervention process. To achieve the general aim, 3 specific aims were formulated for which 2, 1 and 3 studies were conducted, respectively [3-8]. In this chapter, the three specific aims and the main findings for each of the studies are summarized and discussed. Clinical implications and recommendations for future research are formulated.

The first specific aim

The first specific aim of this dissertation was to systematically explore the literature in order to assess whether the intervention matches the diagnostic process in Randomized controlled trials (RCTs) [3] and in Treatment Based Classification Systems (TBCSs) [4] in patients with non-specific neck pain (NSNP).

Main findings

In the first systematic review only a very small proportion, 11 out of 122 RCTs (9%), had a matched intervention to the diagnostic process and matching intervention related outcome measures, thereby determining what needs to be treated and whether the goal of the intervention was reached [9-19]. Since the diagnostic process is essential for selecting appropriate treatment, our systematic review highlights that the absence of some form of diagnostic process in most RCTS, is the most important omission in RCTs with patients with non-specific neck pain. A treatment-based classification system (TBCS) can be expected to match diagnostics with interventions, as they are developed specifically for this match. Our second systematic review, however, showed this not to be the case. The conclusion based on these 2 reviews is that many RCTs and TBCs are insufficiently designed or applied from the perspective of physiotherapeutic validity[3,4].

8

An exception was the StarTBack screening tool [20]. The StarTBack has been examined in a sufficiently physiotherapeutically valid manner [4]. Matched treatments were suggested from the StarTBack, but those proposals are so general that there is no real match between diagnosis and treatment. The StarTBack is not a treatment based classification system that leads to a specific intervention, it does classify patients according to a risk profile, ranging from low to high, of developing chronic back pain, with general treatment suggestions per profile.
The fact that the StarTBack does not lead to specific treatments is confirmed in various studies that aimed to arrive at specific interventions using the StarTBack [21,22]. For example, a study on interventions in patients with a high-risk profile describes that physiotherapists must first learn to identify potential targets for appropriate interventions [21]. This means that after the risk profile has been established (this is where the StarTBack stops), additional diagnostics on risk factors should be performed to establish treatment goals. Another example is a targeted treatment protocol named "the StarTBack trial study protocol" [22]. The name of the protocol suggests that interventions follow directly from the StarTBack outcome. The medium risk group package of care, however, targets physical characteristics which should be diagnosed first. Even for the high-risk group, the physiotherapist still needs to further identify biopsychosocial risk factors before treatment can take place. It is not only characteristic of the StarTBack that a specific interpretation of the treatment of those risk profiles still has to be sought. An RCT on stratified care to prevent chronic low back pain in high-risk patients also shows that the implementation of specific interventions is still a challenge [23].
The StarTBack only provides a level of risk assessment and should be applied as such.

Clinical implications
In the introduction of this thesis we described that, theoretically, it should be easy to integrate the Evidence Based Medicine model into physiotherapy interventions. However, in order to be able to use this model, scientific research must be designed and conducted in a more clinically relevant manner in order to have sufficient physiotherapeutic validity to be able to translate the results of research into daily practice.

In chapter 2, the systematic review of RCTs, we found that only eleven RCTs out of the 122 RCTs included were designed in a physiotherapeutically valid manner. The findings from these eleven studies could be translated into daily practice. I will next describe the studies with positive treatment effects in more detail. Two studies [9,14] investigated interventions to increase endurance of the cervical muscles, namely, postural exercises versus no treatment and craniocervical flexion training versus high load strength training. They used the Cranial Cervical Flexion Test (CCFT) to include participants with a reduced cervical muscle endurance (unable to control more than the second stage of the test) and to evaluate the effect of the intervention. The effect sizes (ES) for the improvement of the muscle endurance was 0.97 for postural exercises and for craniocervical flexion training the ES ranged from 0.91 to 1.15. Effect sizes > 0.8 are considered large [24].Two studies investigated interventions to increase the joint position sense (JPS) of the cervical muscles [10,13]. They used the JPS test to include participants with reduced joint position sense (joint position sense was considered to be impaired, if the deviation of the head was greater than 3° in at least 2 of 8 repositioning tasks) and to evaluate the effect of the intervention. One study [10] investigated the effect of balance training versus no training of the body (ES 0.79) and the other [13] craniocervical flexion training and proprioceptive training of the neck (ES -0.03 for the extension movement to 0.54 for right rotation). In this last study, both interventions appeared to improve the JPS. The PEDro score of these four studies ranged from five to seven. The risk of bias can also be visualized using the Cochrane collaboration tool [25]. What is most striking in both the endurance intervention and the JPS studies is the lack of blinding of the therapist and the assessor. In addition, no between-group comparison was described. Beer[9] did not describe long-term effects in the endurance interventions. Beinert[10] in the JPS intervention studies had no random allocation. It can be concluded that overall there was high RoB.

Cochrane collaboration tool. Endurance training

	Selection bias		**Performance bias**	**Detection bias**	**Attrition bias**	**Reporting bias**	**Other bias**
	Random sequence generation	Allocation concealment	Blinding of participants or personal	Blinding of outcome assessment	Incomplete outcome data	Selective	
Jull et al; 2009	+	+	?	–	+	?	
Beer et al; 2012	+	+	?	–	+	–	Small sample size

Cochrane collaboration tool. JPS training

	Selection bias		**Performance bias**	**Detection bias**	**Attrition bias**	**Reporting bias**	**Other bias**
	Random sequence generation	Allocation concealment	Blinding of participants or personal	Blinding of outcome assessment	Incomplete outcome data	Selective	
Jull et al; 2009	+	+	?	–	+	?	
Beer et al; 2012	–	+	?	–	+	?	

+ = low risk of bias, - = high risk of bias, ? = unclear risk of bias

The conclusion is that the quality of evidence is moderate. Further research may have an important impact on the confidence in the effect estimate. Therefore, if applicable to the patient, we cannot make strong recommendation for these interventions to be used in daily physiotherapy practice yet. However, based on the within[13] or between[9,10,14] group improvements, both the endurance as the JPS training could possibly contribute to the recovery in patients with non-specific neck pain with an decreased endurance or JPS.

Recommendations for future research

If we want to investigate interventions, we must be able to diagnose the problem for which the intervention is intended. All eleven physiotherapeutically valid RCTs had diagnostic cut-off points that they used for the inclusion the population under study. However, none of these cut-off points have been tested for their diagnostic accuracy. There is, therefore, still more work to be done to obtain more valid diagnostic tests for physiotherapists as for now, there are not enough diagnostic tests for the cervical spine with a known diagnostic accuracy [26]. This lack of high quality diagnostic tests ties in with one of the limitations of matched care, namely that accurate matching of the results of the diagnostic process is under development [27].

The second aim

The second aim of this dissertation was to examine expert opinion regarding matching interventions to the results of the diagnostic process in patients with non-specific neck pain. With this aim, we conducted a Delphi study [5].

Main findings

None of the experts considered physiotherapy treatment indicated in patients with non-specific neck pain without a complaint of activity limitation or participation restriction and without any positive signs and/or symptoms or positive diagnostic tests. However, six out of fourteen (43%) experts named one possible treatment in the absence of all these signs and symptoms, namely pain education. However, as pain mechanisms are complex and can be divided into different classifications [28-30], pain education should be specific to the classification of the underlying pain mechanism. It follows that for pain education a diagnostic process is needed as well, to ensure that the education fits in well with the underlying pain mechanism and the degree of understanding by the patient.

We asked the experts which diagnostic tests they used and which they considered to be diagnostically valid. The experts named exactly those tests as valid which were

also used in the physiotherapeutically valid RCTs (CCFT, JPS) [9,10,13]. In addition, the Cervical Range Of Movement (CROM) measurement tool was cited as a valid measurement instrument [5].
Finally, the degree of consensus among experts on linear clinical reasoning processes was investigated. A physiotherapeutic linear (unimodal) clinical reasoning process consists of three sequential phases: the diagnostic, the therapeutic and the evaluative phase. Sequential linear clinical reasoning is defined as the transition ('a line') from signs and symptoms to diagnostic tests, from diagnostic tests to an intervention with matching treatment goal and evaluation based on outcome measurements related to the matched goals. Eighteen lines of sequential linear clinical reasoning leading to the following interventions were examined for the degree of consensus: massage, strength exercises, traction, dry needling, relaxation therapy, mobilization, endurance exercises, stabilization exercises and coordination exercises. An intervention could have multiple lines of sequential linear clinical reasoning. For example experts matched the intervention "mobilisation" with either a ROM test, a joint mobility assessment or an end feel test. This then leads to three lines of sequential linear clinical reasoning. Only six out of 18 lines of sequential linear clinical reasoning reached more than 50% consensus. Three out of these six lines of sequential linear clinical reasoning were consistent, as mentioned earlier, with physiotherapeutically valid RCTs performed.

Clinical implications

The results of this study are consistent with the results of the "the clinical reasoning process in randomized clinical trials study" [3]. That is, experts apply interventions that match the results of the diagnostic process and that are applied in the RCTs with high physiotherapeutical validity.
Endurance training [9,14] and JPS training [10,13] could be effective and can with restraint therefore, if indicated, be used in daily physiotherapy practice, also according to the experts. The experts probably base this on the results of the studies. However, as described earlier, the quality of the evidence is moderate.
This could be hypothesised that this is a first indication that physiotherapeutically valid studies with good results might have been translated into daily practice by physiotherapists. Whether this translation can be attributed to the good effects of the interventions or to the valid design of the studies or both needs to be further investigated.

Recommendations for future research

It appears to become more likely for experts to reach consensus when positive scientific evidence is available. The low level of consensus on consecutive linear clinical reasoning regarding unimodal interventions seems to confirm that there is little positive scientific evidence for the use of most physiotherapy interventions for the cervical spine. This seems to confirm that almost all physiotherapy interventions for non-specific neck pain have yet to be investigated in a physiotherapeutically valid way. Positive scientific evidence can then also be more easily integrated into daily practice [31,32].

The third aim

The third aim of this thesis was to investigate from the perspective of physiotherapeutic validity, the most commonly used physiotherapy intervention, namely manipulations or mobilizations, and their indication in patients with non-specific neck pain. More specifically, the aim was to investigate the diagnostic accuracy of tests used to diagnose a ROM limitation of the neck. In addition, to investigate the effect of mobilizations/manipulations in a population with non-specific neck pain with this ROM limitation as a matching indication for these interventions.

Main findings

Manipulation and mobilization techniques were poorly reported in the investigated RCTs [6]. It therefore remains unclear how the techniques were performed and what the treatment parameters (for example the duration of the intervention or the level at which the intervention was applied) were. It was clear, however, that the aim of these interventions often was to improve the ROM of the neck. However, other interventions, such as exercise therapy, were also investigated for their effects on the ROM of the neck.

As argued earlier, research into the effect of interventions should start with valid diagnostics.

The CROM was mentioned by the experts as a valid measurement instrument [5]. The literature is in line with this [33,34]. Although experts consider the CROM a valid measurement instrument for ROM, it is not recommended in the Dutch neck pain guideline [35]. Dutch physiotherapists assess cervical ROM based on their diagnostic process of history taking and physical examination tests. The first challenge was to determine how reliable and valid this diagnostic process is at assessing limitations in cervical ROM relative to using the CROM. First, we determined a substantiated cut-off point for diagnosing ROM limitation. We then assessed whether questions of perceived movement limitation per direction of movement in combination with the best physical examination test per the same direction of movement were diagnostically accurate. This turned out not to be the case. Therefore, we decided to use the CROM to determine a ROM limitation in our exploratory intervention study.

The next challenge was to investigate commonly used manipulation and mobilization techniques for the treatment of decreased/ limited ROM.
In an exploratory intervention study we wanted to gain insight into what manual therapists do in patients with an limited ROM of the neck where the manual therapist had a primary treatment goal of improving this ROM. Various (combinations of) interventions were applied, such as (as expected) mobilizations and manipulations, but also exercise therapy or giving information. The ROM per direction of the neck (independent from the multimodal treatments) improved between pre and post treatment, ranging from 3.66 for right lateral flexion to 11.65 degrees for left rotation after only 1 treatment by manual therapists and that progress remained stable in any case after 1 week. Not only did ROM improve, but also a computerized tracking task and perceived pain and disability improved.
Although these results might be perceived as promising, the observational study design does not lend itself to making causal statements. That was also not the main aim of the study.
Finally, there were also methodological limitations that could have been avoided. In retrospect, the internal validity would have increased if an independent physiotherapist (and not the physiotherapist who also treated the patient) would have included the patients. However, consideration should be given to offering the treating physiotherapist the option of excluding the patient for treatment as yet. In daily practice, the treating physiotherapist is ultimately responsible for determining whether there is an indication or not.
It would have been informative if we could have investigated what considerations play a role for a therapist in choosing a particular intervention. In-depth interviews were beyond the scope and possibilities of this study, as the primary aim was to determine whether interventions as given in daily practice improve the patient's symptoms. Future research should focus on the decision making of therapists in order to determine which considerations play a role in the choice of therapy.

Clinical implications

Both the diagnostic study and the exploratory intervention study were practice-oriented. We conducted a study with a practice-oriented design to determine the best diagnostic test for establishing a limitation of the ROM of the neck. By practice-oriented we do not only mean a physiotherapeutic practice setting, but that the physiotherapists were free to act in accordance with his/her normal physiotherapy activities. In the diagnostic study, we asked the physiotherapist to perform the CROM measurements in addition to the regular diagnostic process, based on the practice-oriented design. This was then in accordance with daily practice. We used this test as a starting point to include the right target/ diagnostic group matching the interventions in the exploratory intervention study. We let the therapists match their treatments as they do in their daily work. This approach therefore can be labelled as a matched care approach [27]. We allowed the therapists to work as they always do, which made the research practice oriented. This way we tried to improve the translatability of the entire study into daily practice. This facilitates working according to the principles of evidence-based medicine (EBM) defined as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients" [36]. This means that the physiotherapist should integrate research evidence into clinical decision-making whenever possible. Integration can only take place if the physical therapist recognizes his daily practice or if the clinical line of reasoning is understandable in the research (physiotherapeutically valid) and the research is also of sufficient quality (low RoB).

Recommendations for future research

A first step of our diagnostic examination was the determination of a ROM limitation. However, future research should focus on which constructs or mechanisms could potentially contribute to neck movement limitation. Insight into these underlying constructs and mechanisms could in turn lead to an improvement of our physiotherapeutic diagnostics. This in turn facilitates the tailoring of interventions to the results of the diagnostic process.

In our exploratory intervention study the primary goal of treatment was to improve one impairment namely, a ROM limitation of the neck. This suggests the use of one intervention. However, each treatment consisted of two or more interventions. Was this necessary? Or did the participating physiotherapists apply an eclectic approach with a range of interventions that increases the likelihood that the effective intervention will also be included?

It is possible several constructs underly a limmitation of movement, for example a myogenic or an arthrogenic construct. In addition to physical constructs, psychological constructs such as illness perceptions or fear of movement may also lead to a limitation of the mobility of the neck.

A final note

This dissertation arose from the fundamental question that physiotherapists should be able to answer: why do physiotherapists do what they do? We were the first to show that RCTs that include patients with non-specific neck pain hardly use diagnostic tests to guide treatment. It would be interesting to investigate how this is done in other regions of the body.

The various studies align well with The New Agenda for Neck Pain Research [37]. For example the second item on the Agenda for Primary Care Research on Neck Pain is "Translating research evidence on neck pain management into the clinical setting". We believe that practice-oriented research will contribute to this translation. Point nine on this research agenda is: "Identifying clinical features that can be used to direct treatment decisions and identifying which treatment lead to better outcomes for specific individuals with neck pain. This item includes identifying distinct subpopulations with regard to the diagnosis of and prognosis for patients with neck pain". We think that the necessary diagnostic accuracy research will have to be carried out to identify these clinical features. We cannot emphasize enough that before starting to investigate interventions, diagnostics must first be developed and/or demonstrated that the diagnostics are sufficiently reliable and valid [38].

We have made a first cautious step in investigating appropriate treatments (the 'what?' question) for a ROM limitation (the 'why?' question) of the neck. We hope that this tentative initial design of the exploratory study, but also the other studies, will somewhat convince and enthuse researchers to conduct more physiotherapeutically valid and possibly also more practice-oriented scientific research into physiotherapy.

References

1. Damgaard P, Bartels EM, Ris I, Christensen R, Juul-Kristensen B. Evidence of physiotherapy interventions for patients with chronic neck pain: A systematic review of randomised controlled trials. ISRN Pain. 2013;2013:567175.
2. Hush JM, Cameron K, Mackey M. Patient satisfaction with musculoskeletal physical therapy care: A systematic review. Phys Ther. 2011;91(1):25-36.
3. Maissan F, Pool JJM, Raaij de E, Mollema J, Ostelo RWJG, Wittink H. The clinical reasoning process in randomised clinical trials with patients with non-specific neck pain is incomplete: A systematic review. Musculoskeletal Science and Practice. 2018;35:8-17.
4. Maissan F, Pool J, de Raaij E, Wittink H, Ostelo R. Treatment based classification systems for patients with non-specific neck pain. A systematic review. Musculoskelet Sci Pract. 2020;47:102133.
5. Maissan F , Pool JJM, Stutterheim E, Wittink H, Ostello R. Clinical reasoning in unimodal interventions in patients with non-specific neck pain in daily physiotherapy practice, a delphi study. 2018;37:8-16.
6. Pool JJM, Maissan F, Waele de N, Wittink H, Ostelo RWJG. Completeness of the description of manipulation and mobilisation techniques in randomized controlled trials in neck pain; a review using the TiDieR checklist. Musculoskelet.Sci.Pract. 2020.
7. Maissan F, Pool J, de Raaij E, Thoomes-de Graaf M, Ostelo R, Wittink H. The diagnostic accuracy of self-report and physical tests for measuring limitations in the range of motion of the neck. Physiotherapy science and practise.
8. Maissan F, Pool J, de Raaij E, et al. An exploratory, practice-oriented study into matched multimodal treatments in patients with non-specific neck pain. International Journal of Physiotherapy.
9. Beer A, Treleaven J, Jull G. Can a functional postural exercise improve performance in the cranio-cervical flexion test?--a preliminary study. Man Ther. 2012;17(3):219-224.
10. Beinert K, Taube W. The effect of balance training on cervical sensorimotor function and neck pain. J Mot Behav. 2013;45(3):271-278.
11. Cleland JA, Mintken PE, Carpenter K, et al. 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. 2010;90(9):1239-1250.
12. Dawood RS, Kattabei OM, Nasef SA, battarjee KA, Abdelraouf OR. Effectiveness of kinesio taping versus cervical traction on mechanical neck dysfunction. International Journal of Therapies and Rehabilitation Research. 2013;2(2).
13. Jull G, Falla D, Treleaven J, Hodges P, Vicenzino B. Retraining cervical joint position sense: The effect of two exercise regimes. J Orthop Res. 2007;25(3):404-412.

14. Jull GA, Falla D, Vicenzino B, Hodges PW. The effect of therapeutic exercise on activation of the deep cervical flexor muscles in people with chronic neck pain. Man Ther. 2009;14(6):696-701.
15. Kim JY, Kwag KI. Clinical effects of deep cervical flexor muscle activation in patients with chronic neck pain. J Phys Ther Sci. 2016;28(1):269-273.
16. Lee J, Lee Y, Kim H, Lee J. The effects of cervical mobilization combined with thoracic mobilization on forward head posture of neck pain patients. Journal of Physical Therapy Science. 2013;25(1):7-9.
17. Lee KW, Kim WH. Effect of thoracic manipulation and deep craniocervical flexor training on pain, mobility, strength, and disability of the neck of patients with chronic nonspecific neck pain: A randomized clinical trial. J Phys Ther Sci. 2016;28(1):175-180.
18. Mansilla-Ferragut P, Fernandez-de-Las Penas C, Alburquerque-Sendin F, Cleland JA, Bosca-Gandia JJ. Immediate effects of atlanto-occipital joint manipulation on active mouth opening and pressure pain sensitivity in women with mechanical neck pain. J Manipulative Physiol Ther. 2009;32(2):101-106.
19. O'Leary S, Jull G, Kim M, Vicenzino B. Specificity in retraining craniocervical flexor muscle performance. J Orthop Sports Phys Ther. 2007;37(1):3-9.
20. Bier JD, Ostelo RWJG, Koes BW, Verhagen AP. Validity and reproducibility of the modified STarT back tool (dutch version) for patients with neck pain in primary care. Musculoskelet Sci Pract. 2017;31:22-29.
21. Main CJ, Sowden G, Hill JC, Watson PJ, Hay EM. Integrating physical and psychological approaches to treatment in low back pain: The development and content of the STarT back trial's 'high-risk' intervention (StarT back; ISRCTN 37113406). Physiotherapy. 2012;98(2):110-116.
22. Hay EM, Dunn KM, Hill JC, et al. A randomised clinical trial of subgrouping and targeted treatment for low back pain compared with best current care. the STarT back trial study protocol. BMC Musculoskelet Disord. 2008;9:58-2474-9-58.
23. Delitto A, Patterson CG, Stevans JM, et al. Stratified care to prevent chronic low back pain in high-risk patients: The TARGET trial. A multi-site pragmatic cluster randomized trial. EClinicalMedicine. 2021;34:100795.
24. Hojat M, Xu G. A visitor's guide to effect sizes: Statistical significance versus practical (clinical) importance of research findings. Adv Health Sci Educ Theory Pract. 2004;9(3):241-249.
25. Higgins JP, Altman DG, Gøtzsche PC, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
26. Verhagen AP. Physiotherapy management of neck pain. J Physiother. 2021;67(1):5-11. doi: S1836-9553(20)30141-7 [pii].
27. Linton SJ, Nicholas M, Shaw W. Why wait to address high-risk cases of acute low back pain? A comparison of stepped, stratified, and matched care. Pain. 2018;159(12):2437-2441.

28. Dallel R, Voisin D. Towards a pain treatment based on the identification of the pain-generating mechanisms? Eur Neurol. 2001;45(2):126-132.
29. Hüllemann P, Keller T, Kabelitz M, et al. Clinical manifestation of acute, subacute, and chronic low back pain in different age groups: Low back pain in 35,446 patients. Pain Pract. 2018;18(8):1011-1023.
30. Smart KM, Blake C, Staines A, Doody C. The discriminative validity of "nociceptive," "peripheral neuropathic," and "central sensitization" as mechanisms-based classifications of musculoskeletal pain. Clin J Pain. 2011;27(8):655-663.
31. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials. 2017;18(1):122-017-1870-2.
32. Knottnerus JA, Tugwell P. Research methods must find ways of accommodating clinical reality, not ignoring it: The need for pragmatic trials. J Clin Epidemiol. 2017;88:1-3.
33. de Koning CH, van den Heuvel SP, Staal JB, Smits-Engelsman BC, Hendriks EJ. Clinimetric evaluation of active range of motion measures in patients with non-specific neck pain: A systematic review. Eur Spine J. 2008;17(7):905-921.
34. Rondoni A, Rossettini G, Ristori D, et al. Intrarater and inter-rater reliability of active cervical range of motion in patients with nonspecific neck pain measured with technological and common use devices: A systematic review with meta-regression. J Manipulative Physiol Ther. 2017;40(8):597-608.
35. Bier JD, Scholten-Peeters WGM, Staal JB, et al. Clinical practice guideline for physical therapy assessment and treatment in patients with nonspecific neck pain. Phys Ther. 2018;98(3):162-171.
36. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't. 1996. Clin Orthop Relat Res. 2007;455:3-5.
37. Silva PV, Costa LOP, Maher CG, Kamper SJ, Costa LDCM. The new agenda for neck pain research: A modified delphi study. J Orthop Sports Phys Ther. 2019;49(9):666-674. doi: 10.2519/jospt.2019.8704 [doi].
38. Hendriks HJM, Oostendorp R.A.B., Bernards A.T.M., van Ravensberg C.D., Heerkens Y.F., Nelson R.M. The diagnostic proces and indication for physiotherapy: A prerequisite for treatment and outcome evaluation. PHYS THER REV. 2000;5:29-47.

Appendix

Summary

Neck pain is the fourth major cause of disability worldwide. Additionally, neck pain poses an important socio-economic burden on society because pain, stiffness or loss of mobility associated with neck pain often results in utilization of diagnostic assessments and treatments.

The physiotherapist is expected to conduct an examination and subsequently provide an appropriate treatment. In order to maximize the effectiveness of physiotherapy interventions , the physiotherapist must base treatment on the best available evidence (evidence based practice) to give the most effective treatment possible.

Physiotherapists are qualified and professionally required to:

- undertake a comprehensive examination/assessment of the patient/client or needs of a client group
- evaluate the findings from the examination/assessment to make clinical judgments regarding patients/clients
- formulate a diagnosis, prognosis within their expertise and determine when patients/clients need to be referred to another professional
- implement a physiotherapist intervention/treatment programme
- determine the outcomes of any interventions/treatments
- make recommendations for self-management

The physiotherapist should make use of relevant questions during history taking. After history taking, the physiotherapist chooses valid and reliable tests. Information about relevant questions and the validity and or reliability of the tests can be obtained from scientific research. The choice of the intervention also depends on the outcome from scientific research. This whole process is called evidence based clinical reasoning.

Ultimately, the point is that, based upon a thorough clinical reasoning process, the physiotherapist can explain why he / she wants to perform a certain intervention that is appropriate for the specific clinical signs and symptoms of a particular patient. This way of matching patient and care is called “matched care”. Matched care is the new innovative approach in treatment and prevention. In this dissertation we refer to this match between diagnostics, interventions and outcome measures as “physiotherapeutic validity”. In order to be able to translate scientific research into daily physiotherapy practice, it is desirable that scientific research is physiotherapeutically

valid. Physiotherapeutic validity can be equalled to "external validity". External validity is the validity of applying the conclusions of a scientific study outside the context of that study.

The general aim of this dissertation is to gain insight into the physiotherapeutic validity of physiotherapy research in subjects with non-specific neck pain.

Chapter 1 describes the background of the research and the research questions and gives an overview of the studies performed. The studies are divided into 3 parts. Chapters 2 and 3 investigated the match between diagnostics and treatment in scientific research. Chapter 4 examined which diagnostics experts match with predetermined interventions. Chapters 5 and 6 investigated the replicability of scientifically investigated interventions and the validity of diagnostic tools pertaining to Range of Movement (ROM), respectively in people with non-specific neck pain. Finally, chapter 7 describes an exploratory, practice-oriented study into matched treatments in patients with non-specific neck pain.

Chapter 2 presents the results of a systematic review (SR) of the completeness of the clinical reasoning process within the methodology of the RCT in patients with non-specific neck pain. Peer-reviewed literature was systematically searched in the MEDLINE, CINAHL and PEDro databases. A study was included if it met the following criteria: a full text original article published in English, adult patients (> 18 years old) with non-specific neck pain, monodisciplinary physical therapy intervention, and randomized controlled trial. The "Risk of Bias" was assessed using the PEDro checklist. Clinical reasoning was assessed with a HOAC II based clinical reasoning rating scale consisting of 6 items.
We rated a clinical reasoning process as complete when:

1. The problem experienced by the patient was described,
2. A cause was 'diagnosed' or 'argued',
3. The main goal of the intervention was related to the 'cause' (as identified in step 2),
4. The intervention corresponded to the main goal,
5. The intervention-related outcome measure corresponded to the physiotherapist's main goal of the intervention
6. The problem-related outcome measure corresponded to the problem experienced by the patient.

Examples of intervention-related outcome measures are Range of motion (ROM) if the intervention is aimed at improving ROM or muscle strength if the intervention is aimed at improving muscle strength. This often concerns performance-based outcome measures. However, patient reported outcome measures (PROMs) can also be used. In contrast, problem-related outcomes are almost always PROMs.

It remains unclear if risk of bias of a study is associated with the extent to which this study used (and described) a clinical reasoning process. Therefore, it was assessed whether there was an association between the degree of "Risk of Bias" and the completeness of the clinical reasoning process.

For the SR analysis 122 studies were included. Eleven of the 122 studies had an inclusion criterion that diagnosed a possible cause of the complaint (step 2) and objectified the effect with an appropriate intervention-related outcome measure (step 5). Eight of the 122 studies also objectified the change in the patient's complaint with an appropriate outcome measure (step 6). In the majority of studies (70%) the described clinical reasoning process was incomplete. A very small proportion (6%) had a 'diagnosed cause'. There was scarcely any association between the degree of risk of bias and the completeness of the clinical reasoning process, indicating that better methodological quality does not necessarily imply a better description of clinical reasoning process.

This study was a first step to provide insight into the completeness of the physiotherapeutic clinical reasoning process within RCTs in patients with non-specific neck pain.

Chapter 3 presents the results of a SR in which we sought to identify published classification systems with a targeted treatment approach (treatment-based classification systems (TBCSs)) for patients with non-specific neck pain. We also aimed to assess the quality and effectiveness of these systems. Literature was systematically searched in the MEDLINE, CINAHL, EMBASE and PEDro databases. Studies were included if they reported on classification systems on which treatment was based. The system must include physiotherapy treatments in patients with non-specific neck pain. We used the framework developed by Buchbinder to describe the characteristics of a classification system. This framework consists of seven items: (1) purpose of the study; (2) method of development; (3) patient population and setting; (4) specific patient exclusions, (5) categories that describe the specific subgroup; (6) criteria used to assign patients to the subgroup and finally (7) treatments corresponding to the categories. To critically appraise the quality of the classification systems, we used the scoring system also developed by Buchbinder that uses seven criteria: target, content validity, face validity, feasibility, construct validity, (diagnostic) reliability and generalisability. The treatment effect of the systems was assessed from published effect studies of these TBCSs.

Thirteen TBCSs were identified. The overall quality of the thirteen TBCSs varied from low to moderate. We found two RCTs, both with a low risk of bias, that compared the effectiveness of two systems, the Cleland classification system and the McKenzie system, to alternative treatments. The results showed that treatments based on both classification systems were not superior to alternative treatments.

In conclusion, existing treatment-based classification systems are of moderate quality at best. Moreover, these systems were not more effective than alternative treatments. Therefore, we do not recommend the use of these systems in daily physiotherapy practice.

Chapter 4 describes a Delphi study of the clinical reasoning process of physiotherapy experts in unimodal interventions in patients with non-specific neck pain. This study had three goals. First, we aimed explore the expert opinions on the indication for physiotherapy when, other than neck pain, there are no positive signs and symptoms, no positive diagnostic tests or complaints of limitations in functioning or restrictions in participation. Second, we focused on the experts' use of measurement tools and when they are used to support and objectify the clinical reasoning process. Finally, we wanted to reach consensus among experts on the use of unimodal interventions in patients with non-specific neck pain, i.e. their sequential linear clinical reasoning. We considered over 50% consensus in responses as the consensus cut-off point.

Fifteen international experts took part in this study. The Delphi consisted of three rounds. According to all experts, pain alone was not considered to be an indication for physiotherapy. Patient reported outcome measures were mainly used for evaluative purposes and physical tests for diagnostic and evaluative purposes. Eighteen different variants of sequential linear clinical reasoning were examined in our Delphi study. Only 6 of the 18 variants of sequential linear clinical reasoning reached a consensus of more than 50%.

Insight has been obtained about the indication for physiotherapy and when and which measuring instruments are used. There was little consensus on sequential clinical reasoning.

Chapter 5 describes a review that examined the completeness of the description of manipulation and mobilization interventions in randomized controlled trials of subjects with non-specific neck pain. The aim of this study was to investigate whether the quality of the description of manipulation and mobilization interventions is sufficient to replicate these interventions in clinical practice.

Literature was systematically searched in the MEDLINE, CINAHL and PEDro databases. A study was included if it met the following criteria: a full text original article published in English, adult subjects (> 18 years old) with non-specific neck pain, monodisciplinary physiotherapy intervention, and randomized controlled trial. To assess whether the mobilizations / manipulations were fully described, we used the so-called TIDieR checklist. This checklist consists of 12 items.

For the analyses 67 articles were included. Only one article described the technique 'sufficiently' as to be reproducible. However, since the technique is only part of the intervention none of the articles described all the items on the TIDieR list, thus making treatment irreplicable.

In conclusion, mobilization or manipulation interventions are poorly reported in RCTs, compromising the external validity of RCTs, making it difficult for clinicians and researchers to replicate these interventions.

Chapter 6 investigated the diagnostic physiotherapeutic process regarding limited ROM of the neck. First, we investigated the diagnostic accuracy of a self-reported test by the subjects as part of history taking, for determining a movement restriction of the cervical spine (= index test). The subjects were asked to answer the following question: "To what extent do you feel restricted in moving your neck?" scored by a 0-10 rating on a numeric rating scale (NRS) (0 means "no restriction" and 10 means "fully restricted"). For our main analyses we dichotomized this into not limited (0) and limited (1). Second, we investigated the diagnostic accuracy of nine groups of articular dysfunction tests for determining a movement restriction of the cervical spine (= index test), as part of the physical examination.

The reference test was the Cervical Range of Motion device (CROM). The CROM is a valid measurement device for measuring the Active Range of Motion (AROM) of the cervical spine. A recent systematic review presented pooled normative Active Range of Motion (AROM) values for each direction, stratified by age category. For the current study, we classified a movement as limited if the AROM was less than the pooled mean norm value minus 1 standard deviation. The reference test (during physical examination) was performed a few minutes after the index test (during history taking).

Five physiotherapists collected the data in daily practice during the regular physiotherapy process. Subjects were included when they were > 18 years old, had non-specific neck pain (acute and chronic) and understood the Dutch language sufficiently to complete the PROM. Eligible subjects were informed and then asked to participate in the study. If the subject agreed to participate in the study, the subject signed an informed consent prior to data collection. The total number of participating subjects was 128. Diagnostic accuracy of the self-reported test was low to moderate based on the area under the curve (AUC), positive predicted value, negative predicted value, Likelihood Ratio + (LR) and LR - . The diagnostic accuracy remained low when combining the self-reported test and the best physical test per movement direction.

It can be concluded that the overall diagnostic accuracy of physical examination is limited (compared to the CROM measurement). Therefore, a measurement device should be used in daily physical therapy practice to assess if a movement direction is restricted.

Chapter 7 describes an exploratory, practice-oriented study into matched treatments in patients with non-specific neck pain. The objective of this study was 1) to establish the measurement error of the used accelerometer; 2) to determine which different treatments are used for patients with non-specific neck pain with an identified restriction in Range of Motion (ROM) in primary care physiotherapy clinics; 3) to explore if the cervical ROM, pain, (perceived) disability and motor control improved after one treatment.

To be able to interpret the effects on cervical ROM clinically, insight into the measurement error of the ROM, measured with the Sensamove Cervical Trainer accelerometer (SCT), is necessary. That is why first a reproducibility study was performed on the SCT. The measurement error is calculated as the standard error of measurement (SEM) and the Smallest detectable difference (SDD).

For the pilot study participants were recruited from three primary care physiotherapy clinics. The most important inclusion criterion was an identified restriction in cervical Range of Motion (ROM) (measured with an accelerometer). Four manual therapists performed the treatments. Since there is no evidence from the literature which specific treatments lead to an increased ROM, the choice for the specific treatments matched to the individual patient was left to the discretion of the manual therapists based on their clinical reasoning process. This is in line with how manual therapists act in daily practice and therefore we label this study as a practice-oriented study.

Measurements took place pre and post treatment and after a week. Outcome measures were: cervical ROM (flexion/extension, left/right rotation and left /right lateral flexion), a motor control task (both measured with an accelerometer), pain with a Numerical Pain Rating Score (NPRS) and perceived disability with the Patient Specific Function Scale (PSFS). After one week a general perceived effect (GPE) was also measured.

The SCT is a reliable accelerometer for measuring neck ROM, with a small measurement error. Eight different treatments were carried out by the 4 physiotherapists. The NPRS, the PSFS and left and right rotation showed a clinically relevant improvements (exceeded the measurement error). Twenty-three out of twenty- four participants experienced improvement measured with the GPE.

Chapter 8 comprises the general discussion. The general discussion presents an overview of this dissertation and discusses the strengths and limitations of the studies and possible implications of the results and recommendations for future research.

We have made a first cautious step in investigating appropriate treatments (the 'what?' question) for a ROM limitation (the 'why?' question) of the neck. We hope that this tentative initial design of the exploratory study, but also the other studies, will somewhat convince and enthuse researchers to conduct more physiotherapeutically valid, and possibly also more practice-oriented, scientific research into physiotherapy.

Samenvatting

Nekpijn is wereldwijd de op drie na belangrijkste oorzaak van beperkingen in het functioneren of beperkingen in participatie. Bovendien vormt nekpijn een belangrijke sociaal-economische last voor de samenleving, omdat nekpijn gerelateerde pijn, stijfheid of verlies van mobiliteit vaak leidt tot het gebruik van diagnostische beoordelingen en behandelingen.

De fysiotherapeut wordt geacht een onderzoek te doen en vervolgens een passende behandeling te geven. Om de effectiviteit van fysiotherapeutische interventies te optimaliseren, moet de fysiotherapeut de behandeling baseren op het best beschikbare bewijs (evidence based practice).

Fysiotherapeuten zijn gekwalificeerd voor /in staat tot:

- het uitvoeren van een uitgebreid onderzoek van de patiënt/cliënt of behoeften van een cliëntgroep
- het evalueren van de bevindingen van het onderzoek/de beoordeling om klinische beslissingen te nemen over patiënten/cliënten
- het formuleren van een diagnose en prognose binnen hun expertise en het bepalen wanneer patiënten/cliënten doorverwezen moeten worden naar een andere professional
- het implementeren van een fysiotherapeutische interventie of behandelprogramma
- het bepalen van de gewenste uitkomsten van eventuele interventies/behandelingen
- het doen van aanbevelingen voor zelfmanagement

A

De fysiotherapeut dient bij de anamnese gebruik te maken van relevante vragen. Na de anamnese kiest de fysiotherapeut valide en betrouwbare testen. Informatie over relevante vragen en de validiteit en of betrouwbaarheid van de tests kan worden verkregen uit wetenschappelijk onderzoek. De keuze van de interventie hangt ook af van de uitkomst van wetenschappelijk onderzoek.
Dit hele proces wordt evidence-based klinisch redeneren genoemd.

Uiteindelijk gaat het erom dat de fysiotherapeut op grond van een gedegen klinisch redeneerproces kan uitleggen waarom hij/zij een bepaalde interventie wil uitvoeren die past bij de specifieke klinische klachten en symptomen van een bepaalde patiënt. Deze manier van matchen van patiënt en zorg wordt "matched care" genoemd. Matched care is de nieuwe innovatieve benadering in behandeling en preventie. In dit proefschrift noemen we deze match tussen diagnostiek, interventies en uitkomstmaten "fysiotherapeutische validiteit". Om wetenschappelijk onderzoek te kunnen vertalen naar de dagelijkse praktijk van de fysiotherapie, is het wenselijk dat wetenschappelijk onderzoek fysiotherapeutisch valide is. Fysiotherapeutische validiteit kan gelijkgesteld worden aan "externe validiteit". Externe validiteit is de validiteit van het toepassen van de conclusies van een wetenschappelijk onderzoek buiten de context van dat onderzoek.

Het algemene doel van dit proefschrift is om inzicht te krijgen in de fysiotherapeutische validiteit van fysiotherapeutisch onderzoek bij personen met aspecifieke nekpijn.

In **Hoofdstuk 1** worden de achtergrond van het onderzoek en de onderzoeksvragen beschreven evenals een overzicht van de uitgevoerde onderzoeken.
De onderzoeken zijn opgedeeld in 3 delen. In Hoofdstukken 2 en 3 onderzochten we de match tussen diagnostiek en behandeling in wetenschappelijk onderzoek. Hoofdstuk 4 onderzocht welke diagnostische uitkomsten experts matchen met vooraf bepaalde interventies. Hoofdstuk 5 en 6 onderzochten respectievelijk de repliceerbaarheid van wetenschappelijk onderzochte interventies en de validiteit van diagnostische middelen met betrekking tot het beoordelen van de bewegingsuitslag (Range of Motion = ROM) bij mensen met a-specifieke nekpijn. Ten slotte beschrijft hoofdstuk 7 een verkennend, praktijkgericht onderzoek naar gematchte behandelingen bij patiënten met aspecifieke nekpijn.

Hoofdstuk 2 presenteert de resultaten van een systematische review (SR) over de volledigheid van het klinisch redeneerproces binnen de methodologie van het RCT bij patiënten met a-specifieke nekpijn. Peer-reviewed literatuur werd systematisch doorzocht in de MEDLINE-, CINAHL- en PEDro-databases. Een studie werd opgenomen als deze aan de volgende criteria voldeed: een origineel artikel met volledige tekst gepubliceerd in het Engels, volwassen patiënten (> 18 jaar oud) met a-specifieke nekpijn, monodisciplinaire fysiotherapie-interventie en gerandomiseerde gecontroleerde studie. De "Risk of Bias" is beoordeeld aan de hand van de PEDro checklist. Klinisch redeneren werd beoordeeld met een op de HOAC II gebaseerde beoordelingsschaal voor klinisch redeneren, bestaande uit 6 items.

We beoordeelden een klinisch redeneerproces als voltooid wanneer:
1. het door de patiënt ervaren probleem werd beschreven;
2. er een oorzaak werd 'gediagnosticeerd' of 'beredeneerd';
3. het belangrijkste doel van de interventie gerelateerd was aan de 'oorzaak' (zoals geïdentificeerd in stap 2);
4. de interventie overeen kwam met het hoofddoel;
5. de interventie-gerelateerde uitkomstmaat overeen kwam met het hoofddoel van de interventie van de fysiotherapeut;
6. de probleem-gerelateerde uitkomstmaat overeen kwam met het door de patiënt ervaren probleem.

Voorbeelden van interventie-gerelateerde uitkomstmaten (punt 5) zijn ROM als de interventie gericht is op het verbeteren van de ROM of spierkracht als de interventie is gericht op het verbeteren van de spierkracht. Vaak gaat het dan om prestatiegerichte uitkomstmaten. Patiënt-gerapporteerde uitkomstmaten (PROM's) kunnen ook worden gebruikt. Daarentegen zijn probleem-gerelateerde uitkomsten (punt 6) bijna altijd PROM's.

Het blijft onduidelijk of het risico op vertekening van een onderzoek samenhangt met de mate waarin dit onderzoek een klinisch redeneerproces gebruikte (en beschreef). Daarom werd beoordeeld of er een verband was tussen de mate van "Risk of Bias" en de volledigheid van het klinisch redeneerproces.

Voor de SR-analyse werden 122 studies geïncludeerd. Elf van de 122 studies hadden als inclusiecriterium dat een mogelijke oorzaak van de klacht vastgesteld moest zijn (stap 2) en het effect geobjectiveerd moest zijn met een passende interventiegerelateerde uitkomstmaat (stap 5). In 8 van de 122 onderzoeken werd ook de verandering in de klacht van de patiënt geobjectiveerd met een passende uitkomstmaat (stap 6). In de meeste onderzoeken (70%) was het beschreven klinische redeneerproces onvolledig. Een zeer klein deel (6%) had een 'gediagnosticeerde oorzaak'. Er was nauwelijks een verband tussen de mate van risico op bias en de volledigheid van het klinisch redeneerproces, wat aangeeft dat een betere methodologische kwaliteit niet noodzakelijkerwijs een betere beschrijving van het klinisch redeneerproces impliceert.

Dit onderzoek was een eerste stap om inzicht te krijgen in de volledigheid van het fysiotherapeutisch klinisch redeneren binnen RCT's bij patiënten met a-specifieke nekpijn.

In **Hoofdstuk 3** worden de resultaten gepresenteerd van een SR waarin we hebben geprobeerd gepubliceerde classificatiesystemen te identificeren met een gerichte behandelaanpak (treatment-based classificatiesystemen (TBCS's)) voor patiënten met a-specifieke nekpijn. We wilden ook de kwaliteit en effectiviteit van deze systemen beoordelen. Literatuur werd systematisch doorzocht in de databases MEDLINE, CINAHL, EMBASE en PEDro. Studies werden geïncludeerd als ze rapporteerden over classificatiesystemen waarop de behandeling was gebaseerd. Het systeem moest fysiotherapiebehandelingen omvatten bij patiënten met a-specifieke nekpijn. We hebben het door Buchbinder ontwikkelde raamwerk gebruikt om de kenmerken van een classificatiesysteem te beschrijven. Dit raamwerk bestaat uit zeven items: (1) doel van het onderzoek, (2) methode van ontwikkeling, (3) patiëntenpopulatie en setting, (4) specifieke patiëntuitsluitingen,(5)categorieën die de specifieke subgroep beschrijven, (6) criteria die worden gebruikt om patiënten toe te wijzen aan de subgroep en tenslotte (7) behandelingen die overeenkomen met de categorieën. Om de kwaliteit van de classificatiesystemen kritisch te beoordelen, hebben we gebruik gemaakt van het eveneens door Buchbinder ontwikkelde scoresysteem dat gebruik maakt van zeven criteria: doel, inhoudsvaliditeit, gezichtsvaliditeit, haalbaarheid, constructvaliditeit, (diagnostische) betrouwbaarheid en generaliseerbaarheid. Het behandeleffect van de systemen werd beoordeeld op basis van gepubliceerde effectstudies van deze TBCS's.

Dertien TBCS's werden geïdentificeerd. De algehele kwaliteit van de dertien TBCS's varieerde van laag tot matig. We vonden twee RCT's, beide met een laag risico op bias, die de effectiviteit van twee systemen, het Cleland-classificatiesysteem en het McKenzie-systeem, vergeleken met alternatieve behandelingen. De resultaten toonden aan dat behandelingen gebaseerd op beide classificatiesystemen niet superieur waren aan andere toegepaste behandelingen.

Concluderend kunnen we stellen dat de bestaande classificatiesystemen op basis van behandelingen op zijn best van matige kwaliteit zijn. Bovendien waren deze systemen niet effectiever dan alternatieve behandelingen. Daarom raden wij het gebruik van deze systemen in de dagelijkse praktijk van de fysiotherapie af.

In **Hoofdstuk 4** wordt een Delphi-onderzoek beschreven naar het klinisch redeneerproces van fysiotherapeuten bij unimodale interventies bij patiënten met a-specifieke nekpijn. Dit onderzoek had drie doelen. Ten eerste hebben we ons gericht op het verkennen van de mening van experts over de indicatie voor fysiotherapie wanneer er, behalve nekpijn, geen positieve signalen en symptomen zijn, geen positieve diagnostische tests of klachten over beperkingen in het functioneren of beperkingen in participatie. Ten tweede hebben we ons gericht op het gebruik door experts van meetinstrumenten en wanneer deze worden gebruikt om het klinische redeneerproces te ondersteunen en te objectiveren. Ten slotte wilden we consensus bereiken onder experts over het gebruik van unimodale interventies bij patiënten met niet-specifieke nekpijn, d.w.z. hun sequentiële lineaire klinische redeneren. We beschouwden meer dan 50% consensus in antwoorden als het consensusgrenspunt.

Aan dit onderzoek namen vijftien internationale experts deel. De Delphi bestond uit drie rondes. Pijn alleen werd volgens alle deskundigen niet als indicatie voor fysiotherapie beschouwd. Door patiënten gerapporteerde uitkomstmaten werden voornamelijk gebruikt voor evaluatieve doeleinden en fysieke tests voor diagnostische en evaluatieve doeleinden. Achttien verschillende varianten van sequentieel lineair klinisch redeneren werden onderzocht in onze Delphi-studie. Slechts 6 van de 18 varianten van sequentieel lineair klinisch redeneren bereikten een consensus van meer dan 50%.

Er is inzicht verkregen over de indicatie voor fysiotherapie en wanneer en welke meetinstrumenten worden ingezet. Er was weinig consensus over sequentieel klinisch redeneren.

Hoofdstuk 5 bevat een review die de volledigheid van de beschrijving van manipulatie- en mobilisatie interventies onderzocht in gerandomiseerde gecontroleerde onderzoeken van proefpersonen met a-specifieke nekpijn. Het doel van deze studie was om te onderzoeken of de kwaliteit van de beschrijving van manipulatie- en mobilisatie-interventies voldoende is om deze interventies in de klinische praktijk te repliceren.

Er is systematisch gezocht naar literatuur in de databases MEDLINE, CINAHL en PEDro. Een studie werd opgenomen als deze voldeed aan de volgende criteria: een full-text origineel artikel gepubliceerd in het Engels, volwassen proefpersonen (> 18 jaar oud) met niet-specifieke nekpijn, monodisciplinaire fysiotherapeutische interventie en gerandomiseerde gecontroleerde studie. Om te beoordelen of de mobilisaties/manipulaties interventies volledig beschreven waren, hebben we gebruik gemaakt van de zogenaamde TIDieR checklist. Deze checklist bestaat uit 12 items.

Voor de analyses zijn 67 artikelen geïncludeerd. Slechts één artikel beschreef een van de mobilisatie/manipulatie technieken ‘voldoende’ om reproduceerbaar te zijn. Omdat deze techniek echter slechts een deel van de interventie is, werden in geen van de artikelen alle items op de TIDieR-lijst beschreven, waardoor de interventie niet te repliceren is.

Concluderend, mobilisatie- of manipulatie-interventies worden slecht gerapporteerd in RCT's, waardoor de externe validiteit van RCT's in gevaar komt, waardoor het voor clinici en onderzoekers moeilijk wordt om deze interventies te repliceren.

In **Hoofdstuk 6** werd het diagnostisch fysiotherapeutisch proces onderzocht met betrekking tot een beperkte ROM van de nek. Ten eerste hebben we de diagnostische nauwkeurigheid onderzocht van een zelf-gerapporteerde test door de proefpersonen, als onderdeel van de anamnese, voor het bepalen van een bewegingsbeperking van de cervicale wervelkolom (= indextest). De proefpersonen werd gevraagd de volgende vraag te beantwoorden: "In welke mate voelt u zich beperkt in het bewegen van uw nek?" gescoord met een 0-10 score op een numerieke ratingschaal (NRS) (0 betekent "geen beperking" en 10 betekent "volledig beperkt"). Voor onze hoofdanalyses hebben we dit gedichotomiseerd in 'niet beperkt' (0) en 'beperkt' (1). Ten tweede onderzochten we de diagnostische nauwkeurigheid van negen groepen van functietesten, gericht op het vaststellen van een artrogene disfunctie, voor het bepalen van een bewegingsbeperking van de cervicale wervelkolom (= indextest), als onderdeel van het lichamelijk onderzoek.

De referentietest was het Cervical Range of Motion-apparaat (CROM). De CROM is een valide meetinstrument voor het meten van de Active Range of Motion (AROM) van de cervicale wervelkolom. Een recente systematische review presenteerde gepoolde normatieve Actieve Range of Motion (AROM) waarden voor elke richting, gestratificeerd naar leeftijdscategorie. Voor de huidige studie classificeerden we een beweging als beperkt als de AROM kleiner was dan de gepoolde gemiddelde normwaarde minus 1 standaarddeviatie. De referentietest (bij lichamelijk onderzoek) werd enkele minuten na de indextest (bij anamnese) afgenomen.

Vijf fysiotherapeuten verzamelden de gegevens in de dagelijkse praktijk tijdens het reguliere traject fysiotherapie. Proefpersonen werden geïncludeerd wanneer ze > 18 jaar oud waren, a-specifieke nekpijn (acuut en chronisch) hadden en de Nederlandse taal voldoende verstonden om de PROM te voltooien. In aanmerking komende proefpersonen werden geïnformeerd en vervolgens gevraagd om deel te nemen aan het onderzoek. Als de proefpersoon ermee instemde om deel te nemen aan het onderzoek, ondertekende de proefpersoon een geïnformeerde toestemmingsverklaring voorafgaand aan het verzamelen van de gegevens. Het totale aantal deelnemende proefpersonen was 128. De diagnostische nauwkeurigheid van de zelfgerapporteerde test was laag tot matig op basis van de 'area under the curve' (AUC), positief voorspelde waarde, negatief voorspelde waarde, Likelihood Ratio (LR)+ en LR -. De diagnostische nauwkeurigheid bleef laag bij het combineren van de zelfgerapporteerde test en de beste fysieke test per bewegingsrichting.

Geconcludeerd kan worden dat de algehele diagnostische nauwkeurigheid van lichamelijk onderzoek beperkt is (in vergelijking met de CROM-meting). Daarom moet in de dagelijkse fysiotherapiepraktijk een meetinstrument worden gebruikt om te beoordelen of een bewegingsrichting beperkt is.

In **Hoofdstuk 7** wordt een verkennend, praktijkgericht onderzoek gedaan naar op elkaar afgestemde behandelingen bij patiënten met a-specifieke nekpijn. Het doel van dit onderzoek was 1) het vaststellen van de meetfout van de gebruikte accelerometer; 2) het vaststellen van de verschillende toegepaste behandelingen bij patiënten met a-specifieke nekpijn met een geconstateerde beperking in Range of Motion (ROM) in de eerstelijns fysiotherapiepraktijken; en 3) onderzoeken of de cervicale ROM, pijn, (ervaren) beperking en motorische controle verbeterden na één behandeling.

Om de effecten op de cervicale ROM klinisch te kunnen interpreteren is inzicht in de meetfout van de ROM, gemeten met de Sensamove Cervical Trainer accelerometer (SCT), noodzakelijk. Daarom is er eerst een reproduceerbaarheidsonderzoek uitgevoerd op de SCT. De meetfout wordt berekend als de standaard meetfout (SEM) en het kleinste detecteerbare verschil (SDD).

Voor de pilotstudie zijn proefpersonen geworven uit drie eerstelijns fysiotherapiepraktijken. Het belangrijkste inclusiecriterium was een geconstateerde beperking in de cervicale Range of Motion (ROM) (gemeten met een accelerometer). Vier manueel therapeuten voerden de behandelingen uit. Omdat in de literatuur geen evidentie bestaat welke specifieke behandelingen leiden tot een verbetering van de ROM, werd de keuze voor de specifieke behandelingen, afgestemd op de individuele patiënt, overgelaten aan de beoordeling van de manueeltherapeuten op basis van hun klinisch redeneerproces. Dit sluit aan bij hoe manueeltherapeuten handelen in de dagelijkse praktijk en daarom bestempelen wij dit onderzoek als praktijkgericht onderzoek.

Metingen vonden plaats voor en na de behandeling en na een week. Uitkomstmaten waren: cervicale ROM (flexie/extensie, links/rechts rotatie en links/rechts laterale flexie), een motorische controletaak (beide gemeten met een accelerometer), pijn met een Numeric Pain Rating Score (NPRS) en ervaren beperking met de Patiëntspecifieke functieschaal (PSFS). Na een week werd ook een algemeen waargenomen effect (GPE) gemeten.

De SCT is een betrouwbare accelerometer voor het meten van nek-ROM, met een kleine meetfout. Acht verschillende behandelingen werden uitgevoerd door de 4 fysiotherapeuten. De NPRS, de PSFS en de linker- en rechterrotatie lieten een klinisch relevante verbetering zien (meer dan de meetfout). Drieëntwintig van de vierentwintig deelnemers ervoeren verbetering gemeten met de GPE.

Hoofdstuk 8 bevat de algemene discussie. Deze geeft een overzicht van dit proefschrift en bespreekt de sterke punten en beperkingen van de studies en de mogelijke implicaties van de resultaten en aanbevelingen voor toekomstig onderzoek.

We hebben een eerste voorzichtige stap gezet in het onderzoeken van geschikte behandelingen (de ‘wat?’-vraag) voor een ROM-beperking (de ‘waarom?’-vraag) van de nek. We hopen dat deze voorlopige opzet van het verkennend onderzoek, maar ook de andere onderzoeken, onderzoekers enigszins zal overtuigen en enthousiasmeren om meer fysiotherapeutisch valide, en mogelijk ook meer praktijkgericht wetenschappelijk, onderzoek naar de fysiotherapie te doen.

Dankwoord

Volgens schrijver Mark Twain "moet men reizen om te leren".
Voetballer Tim Cahill beweert dat je "een reis beter kan uitdrukken in het aantal vrienden dan in de afstand". Eigenlijk heb ik niets met reizen, maar dit proefschrift was zeker een reis waar ik veel van geleerd heb.
Over de "afstand" heb ik het niet graag, maar over die "vrienden" des te liever.

Dit proefschrift is het product van samenwerking met ongelofelijk gedreven, slimme, maar vooral ook fijne mensen. Promotieteam, promotiecommissie, co-auteurs, collega's bij mijn verschillende werkkringen en niet te vergeten vrienden en familie: allemaal hebben ze op een of andere manier hun steentje bijgedragen. Daarom benoem ik deze mensen hier in mijn dankwoord.

Mijn promotieteam bestond uit professor doctor Raymond Ostelo en doctoren Jan Pool en Harriët Wittink.
Raymond, het begon allemaal aan jouw keukentafel. Op voorspraak van Jan mocht ik mijn onderzoek aan je presenteren. Aanvankelijk ging dat met horten en stoten, maar algauw stelde je me op mijn gemak door geïnteresseerd vragen te stellen. Vragen die richting gaven aan mijn onderzoeksplannen. Je begreep snel dat ik, als clinicus, weinig waarde hechtte aan epidemiologisch goede, maar praktisch weinig bruikbare onderzoeken. Je was een analytisch scherpe promotor, maar altijd ook een fijn mens. Onze intense gesprekken waren eigenlijk altijd een feestje. Ik zal ze missen.
Jan, eigenlijk begon het toen jij bij Hogeschool Utrecht mijn collega werd. Ik deelde mijn ideeën over de klinische zin en onzin van wetenschappelijk onderzoek. Je bood hulp om dit pad verder te ontdekken. De rest is geschiedenis. Onze autoritten van en naar Gouda bleken een prima manier om de voortgang te bespreken. Helaas haalde COVID-19 een streep door die ritjes. Hoe dan ook, zonder jouw internationale netwerk was het Delphi-onderzoek zeker niet geworden wat het nu is. Ik zal je hulp en begeleiding nooit vergeten.
Harriët, we kennen elkaar van toen ik nog Fysiotherapiewetenschap studeerde. Vanaf dat moment begon je me wetenschappelijk te vormen, tot op de dag van vandaag. Zo herkende je meteen de praktische relevantie van mijn premature onderzoeksvoorstel. Bovendien leidden jouw bijdragen tot de HU-voucher, waardoor mijn onderzoek überhaupt mogelijk werd. Jij was, bent en blijft mijn inspiratiebron bij het verder onderbouwen van de fysiotherapie. Want het is een mooi vak, maar onzin hoort er niet in thuis; daar zijn we allebei stellig van overtuigd. Je was een geweldige begelei-

der, stond altijd klaar voor goede raad of een hart onder de riem. Want behalve voor het onderzoek, had je oog voor mij als persoon. Gelukkig blijven we samenwerken.

De leden van de promotiecommissie, dr. S.M. Rubinstein, prof.dr. J.H. van Dieen , prof. dr. B. Cagnie, dr. J.B. Staal en prof.dr. L. Hooft wil ik hartelijk danken voor het lezen en beoordelen van dit proefschrift. Ook dank aan de toenmalige wetenschapscommissie van EMGO VU, zodat mijn onderzoek, na toetsing, ingebed kon worden binnen dat instituut. Een speciale dank gaat naar de Hogeschool Utrecht voor het toekennen van de onderzoeksvoucher.

Mijn dank gaat ook uit naar prof.dr. Nico van Meeteren en prof.dr. Raoul Engelbert voor hun belangrijke bijdragen aan mijn wetenschappelijke scholing.
Beste Nico, jij wist me voor de Fysiotherapiewetenschap te winnen. En ondertussen diagnosticeerde je mijn chronische aandoening, namelijk "wiskunde-deficiëntie" ("kuren" bij het James Boswell instituut heeft me genezen…).
Beste Raoul, zonder jou zou ik gestopt zijn met de studie Fysiotherapiewetenschap. Deelname aan het JASS-project onder jouw bezielende leiding leverde een artikel op in het Nederlands Tijdschrift voor Fysiotherapie. Dat inspireerde me om door te gaan. En daar heb ik nooit spijt van gehad.

Ook co-auteurs Edwin, Jurgen†, Eric, Nanna, Marloes, Paul en Glenn wil ik hier noemen. Jullie bijdragen aan de artikelen in dit boekje waren stuk voor stuk gulle giften. Marloes, jij verdient extra dank voor jouw Lime-survey-programma, je hulp bij het verzamelen van data en het meedenken bij de laatste twee studies. Ook Melissa, Ilona, Tamara, Glenn, Koen, Guido en Erik bedankt voor jullie hulp bij het verzamelen van de data. En prof.dr. Arianne Verhagen voor haar hulp bij het verbeteren van één van de manuscripten.

Hogeschool Utrecht wil ik bedanken voor het in mij gestelde vertrouwen. De HU voelt als een warm bad. Het is onmogelijk om alle collega's binnen het Instituut voor bewegingsstudies hier persoonlijk te bedanken. Een aantal collega's noem ik toch bij naam. Peter, Norman, Sijmen en Jaap J. voor het brainstormen en omdat jullie de onderzoeksresultaten in het onderwijs benutten. De collega's van het Lectoraat Leefstijl en Gezondheid, Edwin, Ryan, Erik-Jan, Han, Marlies, Martine, Else, Eline, Stefan, Janke, Michiel, Martine, Tim, Jan, Manon, Marleen, Marike, Imke, Hannelies, Claudia, Henri, Jacqueline N., Jacqueline O., Kristel, Karlijn en Barbara, voor de gezellige review sessies. De "oude" hoofden van de Master Fysiotherapie, Ina, Jacqueline N., Rutger, Jorrit, Roland en Rob dank ik voor de mooie tijd op de Bolognalaan en jullie warme belangstelling voor mijn project. Een belangstelling die ik ook mocht ervaren van de "nieuwe" hoofden: Linda, Barbara, Brenda, Stefan, Els en Henri als "capo di tutti capi".

Verder alle collega's uit de HU-teams waar ik in mocht werken: Jaap D., Kitty, Marielle, Janke, Miriam en Sijmen van het (oude en nieuwe) coördinatieteam van de generieke leerlijn; Jan C., Marjolein, Marielle, Milou en Mark van het beweegzorgteam en Annemarie, Marc, Martine, Sabrine, Esther en Casper van mijn huidige OMT-team. Bedankt voor de samenwerking en jullie support. Ook dat ene bijzondere 'team' moet hier genoemd worden: het maandagochtend-theegroepje. Carla en Barbara, jaren hebben we op de vroege maandagochtend lief en leed gedeeld over onder andere mijn traject. Dank voor jullie interesse en steun en laten we vooral doorgaan met die gezellige start van de week.

Ook collega's uit mijn andere werkkringen horen in dit dankwoord thuis: Adri, Caroline, Karin en Saskia van de FysioPraxis-redactie; Nico, Jordy Danielle en overige collega's van fysiotherapie "Westwijk" en Annelies, Stefan, Lennard, Gerard, Nathan, Sandra, Marianne, Paul, Radha en Diek van het NVMT-bestuur. Ook jullie bedankt voor alle steun en begrip.

Rest me, mijn basis in het leven te bedanken: vrienden en familie.

Edwin, al was je formeel geen lid van het promotieteam, voor mij hoorde je er gewoon bij. We volgden samen de opleiding Fysiotherapie, werkten in dezelfde Rotterdamse wijk, studeerden later allebei Fysiotherapiewetenschap. Daar herkenden we ons allebei niet in de resultaten van wetenschappelijk onderzoek. We dachten na over hoe we ons klinisch handelen konden objectiveren en zo ontstond onze liefde voor de klinimetrie. Jij bent de drijvende kracht achter de ideeën waarop dit proefschrift is gebaseerd. Maar voor mij nog veel belangrijker zijn jouw nooit aflatende steun in lief en leed en jouw oprechtheid: een echte vriend.
Marieke, jij bent net zo oprecht als Edwin. Jullie zijn een mooi stel. Gelukkig wilde je de vormgeving van mijn proefschrift op je nemen, zoals je dat eerder ook voor Edwin had gedaan. Zonder jou hulp en jullie ervaring was het niet zo'n mooi boekje geworden.

Beste Mohamed, sommige mensen kosten energie, van anderen krijg je energie. Jij bent zo'n gever. Dank voor je enthousiasme tijdens het brainstormen over mijn onderzoek en dank dat je mijn paranimf wilde zijn. Dat we nog lang mogen blijven samenwerken.

Alfred en Marcia, Erik en Gemma veel dank als goede vrienden voor jullie steun en enthousiasme.
Mijn moeder, schoonmoeder, zwager Jaco, schoonzus Petra, nicht Chrétiènne en zwager André (tijdens de thuiswedstrijden van Sparta) wilden constant op de hoogte worden gehouden over het wel en wee van mijn traject. Dank voor jullie betrokkenheid.

Als laatste mijn hometeam. Lisanne en Annelotte, bedankt voor jullie geduld met mij. Ik zal namelijk niet altijd de gezelligste vader geweest zijn. Ondertussen groeiden jullie uit tot volwassen dochters met jullie eigen carrières. Ik ben trots op jullie! En dan, last but not least, Nicolette. Dankjewel voor alles. Samenwonen met een promovendus is nu eenmaal geen sinecure. Jij bent en blijft mijn ultieme steun en toeverlaat. Op naar nieuwe rust en regelmaat!

Publications

Meijer S, Suir I, Maissan JF, Nuysink J.
Reliability of using a smartphone application to objectify skull deformation.
Pediatr Phys Ther; May, 2022

Verwoerd MJ, Wittink H, Goossens MEJB, Maissan F, Smeets RJEM.
Physiotherapists' knowledge, attitude and practice behavior to prevent chronification in patients with non-specific, non-traumatic, acute- and subacute neck pain: A qualitative study. Musculoskelet Sci Pract, 2022: 57

de Raaij EJ, Wittink H, Maissan JF, Twisk J, Ostelo RWJG.
Illness perseptions; exploring mediators and/or moderators in disabling persistent low back pain. Multiple baseline single-case experimental design.
BMC Musculoskelet Di, 2022; 23,140

de Raaij EJ, Wittink H, Maissan JF, Westers P, Ostelo RWJG.
Limited predictive value of illness perceptions for the short-term poor recovery in musculoskeletal pain. A multi-center longitudinal study. Bmc Musculoskelet Di, 2021; 22,522

de Oude K , Maissan JF.
Predictiemodel voor vallen bij ouderen met een verstandelijke beperking.
Nederlands Tijdschrift voor Geriatrie. 2021; Juni

JF Maissan, E de Raaij.
Praktijkgericht wetenschappelijk onderzoeks in de manuele therapie. FysioPraxis. 2021; Maart

A

de Raaij EJ, Ostelo RWJG, Maissan JF, Pool J, Westers P, Wittink H.
Illness perceptions associated with patient burden with musculoskeletal pain in outpatient physical therapy practice, a cross-sectional study. Musculoskelet Sci Prakt. 2020; 45

Thoomes-de Graaf M, Thoomes E, Falla D, Fernandez-de-Las-Penas C, Maissan JF, Cleland JA.
Does the patient and clinician perception of restricted range of cervical movement agree with the objective quantification of movement in people with neck pain? And do clinicians agree in their interpretation? Musculoskelet Sci Pract. 2020;

Verwoerd M, Wittink H, Maissan JF, Smeets RJEM.
Consensus of potential modifiable prognostic factors for persistent pain after a first episode of nonspecific idiopathic, non-traumatic neck pain: results of nominal group and Delphi technique approach. BMC Muscloskelet disord. 2020; 21:656

Maissan JF, Pool J, de Raaij E, Wittink H, Ostelo RWJG.
Treatment based classification systems for patients with non-specific neck pain.
A systematic review. Musculoskelet Sci Pract. 2020; 47

Pool J, Maissan JF, de Waele N, Wittink H, Ostelo RWJG.
Completeness of the description of manipulation and mobilisation techniques in randomized controlled trials in neck pain; A review using the TiDieR checklist.
Musculoskelet Sci Pract. 2020; 45

Verwoerd M, Wittink H, Maissan JF, de Raaij EJ, Smeets RJEM.
Prognostic factors for persistent pain after a first episode of nonspecific idiopathic, non-traumatic neck pain: A systematic review. Musculoskelet Sci Pract. 2019; 42, 13-37

Jorna-Lakke S, Maissan JF.
Master-eis voor alle manueel therapeuten in zicht. FysioPraxis. 2018; 27(10), 38-39

Maissan JF, Pool J, Stutterheim E, Wittink H, Ostelo RWJG.
Clinical reasoning in unimodal interventions in patients with non-specific neck pain in daily physiotherapy practice, a Delphi study. Musculoskelet Sci Pract. 2018; 37, 8-16

Maissan JF, Pool J, de Raaij E, Mollema J, Ostelo RWJG, Wittink H.
The clinical reasoning process in randomized clinical trials with patients with non-specific neck pain is incomplete: A systematic review.
Musculoskelet Sci Pract. 2018; 35, 8-17

de Raaij EJ, Ostelo RWJG, Maissan JF, Mollema J, Wittink H.
Illness Perceptions' Association and Prognosis With Pain and Physical Function in Patients With Non-Cancer Musculoskeletal Pain: A Systematic Literature Review.
J Orthop Sports Phys Ther. 2018; 10, 1-48

Zwerus EL, Somford MP, Maissan JF, Heisen J, Eygendaal D, van den Bekerom MP. Physical examination of the elbow, what is the evidence? A systematic literature review. Br J Sports Med. 2018; 52 (19), 1253-1260

de Raaij EJ, Pool J, Maissan JF, Wittink H.
Illness perceptions and activity limitations in osteoarthritis of the knee: A case report intervention study. Man Ther. 2014; 19(2), 169-172

de Raaij EJ., Schröder C., Maissan JF.J., Pool J.J., Wittink H.
Cross-cultural adaptation and measurement properties of the Brief Illness Perception Questionnaire-Dutch Language Version; Manual Therapy, 2012; 17(4), 330-335

Pekaric J.,Berg, van den, MJM., Maissan, JF., Rijn, van, M.
Klinisch relevante veranderingen bij een patiënt met laterale epicondyalgie na toepassing van manuele mobilisaties en excentrische spierkrachttraining.
Tijdschrift Manuele Therapie. 2011; 2: 13-17

Arentsen R, de Raaij E.J., Maissan J.F.
Zorgt evidence-based practice voor een spagaat binnen de fysiotherapie?
FysioPraxis. 2010; 33

Brueren MM, de Raaij EJ, Maissan JF.
Facilitatie en inhibitie van musculaire dysfuncties in beenmusculatuur bij mensen met een amputatie van de onderste extremiteit.
Nederlands Tijdschrift voor Fysiotherapie 2010; (120) 57-65

Maissan JF, van Maris RF, Aerts MGF, de Raaij EJ.
Cervicale Range of Motion en manuele therapie. Tijdschrift voor Manuele Therapie. 2010; 6 (1), 31-32

Maissan JF, Pruijssers P, de Raaij EJ.
Klinische predictieregels en manuele therapie: een succesvol huwelijk?
Tijdschrift voor Manuele Therapie. 2009; 5(2), 13-16

de Raaij EJ, Maissan JF, Kingma JJ, van Maanen CJ, Vossen HPLM.
Onderzoeksvragen op het gebied van diagnostiek en prognostiek onder de Nederlandse deelnemers van het IFOMT- 2008 congres manuele therapie.
Tijdschrift Manueel Therapie 2009; 18(2), 28-32

A

van de Water ATM , Maissan JF.
Extracorporale Shockwave Therapie gecombineerd met actieve oefeningen bij een patiënt met chronische fasciosis plantaris: een case study. FysioPraxis 2009; 28-32

van Peppen RPS, Maissan JF, van Genderen FR, van Dolder R, van Meeteren NLU. Outcome measures in physiotherapy management of patients with stroke: a survey into self-reported use, and barriers to and facilitators for use.
Physiotherapy Research International, 2008; 13(4), 255-270

Maissan JF, de Raaij EJ, Prins JCM, van Genderen FR.
Is er een klinisch relevante benadering van betrouwbaarheidsonderzoek geldend voor de fysiotherapeut? Nederlands Tijdschrift voor Fysiotherapie 2008. 2, 42-46

Maissan JF, de Raaij EJ. KNGF-richtlijnen als procesindicator en de uitkomstmaat als uitkomstindicator; een kans of bedreiging? FysioPraxis 2007, 16 (12), 18-20

Maissan JF, Passchier EA, Viehoff PB, van Genderen FR, Engelbert RHH, van Meeteren NLU. De invloed van training op het betrouwbaar meten van actieve gewrichtsmobiliteit en spiersterkte; een pragmatische studie. Nederlands Tijdschrift voor Fysiotherapie 2006; 116 (3), 50-55

Under review

Maissan JF, Pool J, de Raaij EJ, Thoomes- de Graaf M, Ostelo RWJG, Wittink H.
The diagnostic accuracy of self-reported limitations and physical tests for measuring range of motion of the neck.

Accepted

Maissan JF, Pool J, de Raaij EJ, Thoomes-de Graaf M, Westers P, Kroon G, Wittink H, Ostelo RWJG.
An exploratory, practice-oriented pilot study into matched treatments in patients with non-specific neck pain.

About the author

Francois Maissan was born in Gorinchem, the Netherlands, on September 11th 1965. He completed his physiotherapy bachelor education at the University of Applied Sciences Rotterdam in 1990 and his Manual Therapy specialization at the SOMT university in 1996. In 2002 he started his master of Science education in Health Sciences, during which he specialised in Physiotherapy-Research. In 2006 he received his Master of Science degree with a thesis named “Effect of rater training on reliability of Myoton® and goniometer measurements, a clinical randomized trial”.

It was during this Health Sciences master that the idea of this PhD trajectory arose. It was striking that the results of scientific research into physiotherapeutic interventions did not correspond with the results he observed in clinical practice. In November 2015, Francois received a research voucher from the HU University of Applied Sciences for his PhD research named “Non-specific neck pain: To match or not to match”.

Francois has been working in various physiotherapy practices as a physiotherapist since 1990 and as a manual therapist since 1996. He has been an internship supervisor for physiotherapy and manual therapy students. He started working for the Master Physiotherapy education at HU University of Applied Sciences in 2006. There he held various positions such as faculty member, program coordinator, vice-chairman of the examination committee, member of the participation council and he recently started as head of the manual therapy educational program. In 2012, he started his preparatory work for his PhD research at the Lifestyle and Health research group under the supervision of professor Harriët Wittink.

Other activities that Francois has carried out were:
co-owner of “Fysiometrics”; a company that provides courses in applied clinimetrics for paramedical professions (from 2008), editor for the Fysiopraxis journal, a Dutch journal for physiotherapists (2011-2019) and board member of the Dutch manual therapy association (2012-2021).

Amsterdam
Movement
Sciences

VU VRIJE UNIVERSITEIT AMSTERDAM

Amsterdam UMC
University Medical Centers

Amsterdam Movement Sciences conducts scientific research to optimize physical performance in health and disease based on a fundamental understanding of human movement in order to contribute to the fulfillment of a meaningful life.