Volume 38 • Number 2 • December 2018
www.worldscientific.com/worldscinet/hkpj

HONG KONG PHYSIOTHERAPY JOURNAL

World Scientific

HONG KONG PHYSIOTHERAPY JOURNAL

Volume 38 • Number 2 • December 2018

www.worldscientific.com/worldscinet/hkpj

Editor-in-Chief

Professor Marco Y.C. Pang, PhD

Department of Rehabilitation Sciences,
The Hong Kong Polytechnic University, Hong Kong

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Hong Kong Physiotherapy Journal

Aims & Scope

The Hong Kong Physiotherapy Journal (HKPJ) is the official peer-reviewed, Open Access (OA) publication of the Hong Kong Physiotherapy Association.

HKPJ publishes papers related to all areas of physiotherapy (education, research, practice, policies) and is committed to facilitating communication among educators, researchers and practitioners in the field with the aim of promoting evidence-based practice.

We are particularly interested in publishing randomized controlled trials, systematic reviews and meta-analyses. Animal studies are also welcome if the study question and findings have important relevance to physiotherapy practice.

HKPJ welcomes submissions from all over the world in the form of original research papers, reviews, editorials, treatment reports, technical notes, and correspondence.

Hong Kong Physiotherapy Journal [ISSN (print): 1013-7025 | ISSN (online): 1876-441X] is published by World Scientific Publishing Co Pte Ltd.

HONG KONG PHYSIOTHERAPY JOURNAL

Vol. 38, No. 2, December 2018

Editorial

The importance of developing evidence-based clinical examinations for low back pain
A. Wong i

Review

Musculoskeletal disorder and pain associated with smartphone use: A systematic review of biomechanical evidence
A. C. Eitivipart, S. Viriyarojanakul and L. Redhead 77

Review and Meta-analysis

Effectiveness of surgical versus conservative treatment for carpal tunnel syndrome: A systematic review, meta-analysis and qualitative analysis
D. Klokkari and I. Mamais 91

Research Paper

Reliability and validity of transfer assessment instrument version 3.0 in individuals with acute spinal cord injury in early rehabilitation phase
P. Baghel, S. Walia and M. M. Noohu 115

Agreement of clinical examination for low back pain with facet joint origin
M. Vongsirinavarat, W. Wahyuddin and R. Adisaiphaopan 125

Side-to-side elbow range of movement variability in an ulnar neurodynamic test sequence variant in asymptomatic people
M. M. Y. Tong, V. C.-H. Liu and T. Hall 133

Stratification of stroke rehabilitation: Five-year profiles of functional outcomes
B. P. H. Chung 141

Comparative effectiveness of transverse oscillatory pressure and cervical traction in the management of cervical radiculopathy: A randomized controlled study
A. O. Ojoawo and A. D. Olabode 149

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) i–iii
DOI: 10.1142/S1013702518010023

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

The importance of developing evidence-based clinical examinations for low back pain

Arnold Wong

Department of Rehabilitation Sciences
The Hong Kong Polytechnic University, Hong Kong

arnold.wong@polyu.edu.hk

Accepted 22 November 2018; Published 11 December 2018

Low back pain (LBP) is the number one cause of years lived with disability in the world.[1] Approximately 80% of people experience LBP at least once in their lifetime and many of them remain to have LBP at older ages.[1–3] Despite the high prevalence of LBP, approximately 90% of LBP are labelled as non-specific LBP because no clear etiologies can be found.[1] Given that medical imaging has limited values in diagnosing patients with LBP, lumbar imaging is recommended only when serious pathologies (e.g., malignancy, fracture, infection) are suspected.[4]

In order to prescribe treatments for patients with LBP, physical therapists need to rely on patient history and clinical examinations to inform clinical decisions. In the absence of robust evidence or knowledge on the assessment criteria for identifying patients with different underlying causes of non-specific LBP, a Delphi process is commonly used to solicit expert opinions regarding the most appropriate assessment criteria for classifying patients into different subgroups for treatment allocations.[5–7] Following the Delphi process, further studies should be conducted to evaluate the clinimetric properties of the recommended assessment criteria. Since some clinical examinations initially thought to be useful for differentiating different patient subgroups for treatment allocations may display suboptimal clinical values in some patient populations later,[8,9] any clinical examinations derived from the Delphi process must be evaluated thoroughly before applying them in clinical practice.

In this issue of Hong Kong Physiotherapy Journal, Vongsirinavarat and co-workers[10] conducted a single-group, repeated measures reliability study to evaluate the agreement of two experienced physical therapists in using nine assessment criteria derived from a Delphi study to diagnose patients with lumbar facet joint pain in a clinical setting.[7] Specifically, the assessment criteria include three subjective assessments (i.e., localized unilateral pain, referred pain above knee, and no radicular pain), three movement tests (i.e., pain reduction in flexion, pain in extension, and pain in extension with side flexion and rotation

toward the same side of the painful facet joint), and three manual assessments (i.e., replication or aggravation of pain by local pressure over a facet joint or a transverse process, localized muscle spasm over the lumbar facet joint, and reduced passive range of movement or increased stiffness on the painful facet joint during palpation). The results revealed fair to substantial agreements between the two physical therapists in using the nine criteria to distinguish patients with and without lumbar facet joint pain. The suboptimal agreement on some criteria may be attributed to poor reliability of the palpation-based assessments,[11,12] and/or lack of detailed definitions and training on some assessment criteria prior to the commencement of study. As such, the authors provided specific pragmatic solutions and suggestions for future research.

While a Delphi consensus methodology has been commonly used to solicit consensus from experts regarding the diagnosis or classification of patient with different medical conditions,[5,13] the results derived from these Delphi studies may not necessarily be adopted by clinicians. It is partly because there is no standard methodology to validate the new knowledge and to put the validated knowledge into practice. To facilitate this knowledge translation, a multistage methodology used for the validation of clinical prediction rules may be adopted.[14,15] In brief, the validation process includes: (1) a narrow validation (i.e., validation of the diagnostic criteria in one or two clinics), (2) a broad validation (i.e., validation of the criteria in separate populations), and (3) an impact analysis (i.e., evaluation of the usefulness of the diagnostic criteria in improving patient outcomes, changing clinicians' behavior, and/or reducing resource consumption). The first two stages validate the psychometric properties of the experts derived diagnostic criteria in different patient populations or different healthcare settings. The diagnostic criteria can be refined during these two stages. The third stage aims to evaluate the impact of implementing the diagnostic criteria on the diagnostic accuracy, medical costs, and patient satisfaction in clinical practice.[15] Collectively, while a Delphi consensus process can be used as an initial step to derive diagnostic criteria for a disease,[16] a standard validation process should be followed to ensure the validity of the suggested diagnostic criteria and to evaluate the benefits of using those criteria in improving clinical practice.

References

1. Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. Lancet 2012;379 (9814):482–91.
2. Wong A, Karppinen J, Samartzis D. Low back pain in older adults: Risk factors, management options and future directions. Scoliosis Spinal Disord 2017;12:14.
3. Ogunlana MO, Odole AC, Adejumo A, Odunaiya N. Catastrophising, pain, and disability in patients with nonspecific low back pain. Hong Kong Physiother J 2015;33:73–79.
4. Darlow B, Foster BB, O'Sullivan K, et al. It is time to stop causing harm with inappropriate imaging for low back pain. Br J Sports Med 2017;51:414–5.
5. Tomkins-Lane C, Melloh M, Lune J, et al., ISSLS prize winner: Consensus on clinical diagnosis of lumbar spinal stenosis: Results of an international Delphi study. Spine 2016;41:1239–46.
6. Wong A, Lauridsen H, Samartzis D, Macedo L, Ferreira P, Ferreira M. Global consensus from clinicians regarding low back pain outcome indicators for older adults: Pairwise wiki survey using crowdsourcing. JMIR Rehabil Assist Technol. [accepted].
7. Wilde VE, Ford JJ, McMeeken JM. Indicators of lumbar zygapophyseal joint pain: Survey of an expert panel with the Delphi technique. Physic Ther 2007;87(10):1348–61.
8. Mitchell U, Hurrell J. Clinical spinal stability: 10 years since the derivation of a clinical prediction rule. A narrative literature review. J Back Musculoskelet Rehabil 2018;[Epub ahead of print].
9. Kendell M, Beales D, O'Sullivan P, Rabey M, Hill J, Smith A. The predictive ability of the STarT Back Tool was limited in people with chronic low back pain: A prospective cohort study. J Physiother 2018;64:107–13.
10. Vongsirinavarat M, Wahyuddin W, Adisaiphaopan R. Agreement of clinical examination for low back pain with facet joint origin. Hong Kong Physiother J 2018;38(2):125–31.
11. Strender LE, Sjoblom A, Sundell K, Ludwig R, Taube A. Interexaminer reliability in physical examination of patients with low back pain. Spine 1997;22(7):814–20.
12. Wong A, Kawchuk G. The clinical value of lumbar posterioanterior segmental stiffness: A narrative review. PM&R 2017;9:816–30.
13. INSITE Collaborations (INTernational Study group for Identification and Treatment of Endofibrosis). Diagnosis and management of iliac artery endofibrosis: Results of a Delphi consensus study. Eur J Vasc Endovasc Surg 2016;52:90–8.

14. Falk G and Fahey T. Clinical prediction rule. BMJ 2009;339:b2899.
15. Childs JD and Cleland JA. Development and application of clinical prediction rules to improve decision making in physical therapist practice. Phys Ther 2006;86:121–31.
16. Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. Semin Arthritis Rheum 2011;41:95–105.

Review

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) 77–90
DOI: 10.1142/S1013702518300010

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

Musculoskeletal disorder and pain associated with smartphone use: A systematic review of biomechanical evidence

Aitthanatt Chachris Eitivipart[1,2,*], Sirinya Viriyarojanakul[3] and Lucy Redhead[4]

[1]*Faculty of Allied Health Sciences*
Department of Physical Therapy
Chulalongkorn University, Bangkok, Thailand

[2]*Discipline of Exercise and Sport Science, Faculty of Health Sciences*
The University of Sydney, Sydney, Australia

[3]*Faculty of Physical Therapy*
Saint Louis College, Bangkok, Thailand

[4]*School of Health Sciences, University of Brighton, Brighton, UK*

[*]*c.eitivipart28336@gmail.com*

Received 17 January 2017; Accepted 24 October 2017; Published 14 August 2018

The number of smartphone users is growing dramatically. Using the smartphone frequently forces the users to adopt an awkward posture leading to an increased risk of musculoskeletal disorders and pain. The objective of this study is to conduct a systematic review of studies that assess the effect of smartphone use on musculoskeletal disorders and pain. A systematic literature search of AMED, CINAHL, PubMed, Proquest, ScienceDirect using specific keywords relating to smartphone, musculoskeletal disorders and pain was conducted. Reference lists of related papers were searched for additional studies. Methodological quality was assessed by two independent reviewers using the modified Downs and Black checklist. From 639 reports identified from electronic databases, 11 were eligible to include in the review. One paper was found from the list of references and added to the review. The quality scores were rated as moderate. The results show that muscle activity of upper trapezius, erector spinae and the neck extensor muscles are increased as well as head flexion angle, head

[*]Corresponding author.

tilt angle and forward head shifting which increased during the smartphone use. Also, smartphone use in a sitting position seems to cause more shift in head–neck angle than in a standing position. Smartphone usage may contribute to musculoskeletal disorders. The findings of the included papers should be interpreted carefully in light of the issues highlighted by the moderate-quality assessment scores.

Keywords: Smartphone; musculoskeletal disorders; pain.

Introduction

Smartphones now have a significant role in people's everyday lives as they are being used for communication, internet browsing and gaming. In the past decade, the rate of smartphone usage, hours and frequency of use, has been increased.[1,2] A study in 2012 revealed that there were more than six billion smartphone users worldwide.[3] Additionally, research reported that over 65% of the owners in the USA spent at least 1 h per day on their phone.[4] A survey supported this trend by reporting that users spend more than 20 h weekly on texting, emailing, and using social network, representing the significant dependence on smartphones for connecting and communicating with others.[5] Consequently, the heavy reliance on the smartphone may contribute to musculoskeletal injuries in the users. Therefore, health professionals should be aware of the effect of smartphone use on physical health problems. Generally, the typical posture when using smartphones (or other touchscreen handheld devices) involves holding the tool with one or two hands below the eye level, looking down at the device and using the thumb to touch the screen.[6] This pattern of use forces the user to adopt an awkward posture such as forward neck flexion which is often maintained for long periods.[6–9] The prolonged and frequent use of smartphones, as well as the repeated movement of the upper extremities in an awkward posture, have been shown to be the main contributing factors to the incidence of musculoskeletal symptoms.[7–9] Musculoskeletal symptoms, such as discomfort and pain, in smartphone users not only occur in the neck but also in other areas of the body including shoulders, elbows, arms, wrists, hands, thumbs and fingers.[1,6,10–14]

While some research has been conducted to study the effect of smartphone use on the musculoskeletal symptoms of the neck and upper extremity, there has not been a systematic review evaluating this research. The purpose of this study is to systematically review the evidence from experimental studies and may draw a definite conclusion regarding the research that focuses on the changes in musculoskeletal symptoms caused by smartphone usage.

Methods

A search of the Cochrane Library and the databases included in this review revealed no equivalent systematic review. This systematic review was planned and accomplished based on the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement for reporting systematic review.[15]

Literature Search

A comprehensive search was performed in May 2016 by two independent researchers (AE and SV) of the following databases: AMED, CINAHL, PubMed, ProQuest and ScienceDirect. There was no date restriction. The combination of terms and keywords used were (smartphone OR mobile phone OR texting OR typing) AND (musculoskeletal disorder OR pain) AND (ergonomic OR human factor). Handsearching of the reference lists of all relevant papers was performed. Only papers written in English were included. The inclusion criteria were the following: (1) the studies must be laboratory experimental studies (pre-post, quasi-experimental, or cross-sectional study) so that the actual data relating to the change in different musculoskeletal symptoms due to the use of smartphone could be tracked in an objective way; (2) the outcome must contain at least one of the following aspects: pain, postural analysis or muscle activity; (3) the assessments of the subjects must focus on the upper extremities including neck, shoulder, elbow, wrist, hand, thumb, fingers, and upper back; and (4) the effects of smartphone use must be the main focus in the research. Studies were excluded if (1) the

research recruited subjects aged under 18; (2) the studies focused on the use of a tablet, computer, and other visual display units; and (3) the primary outcome of the research was from survey or qualitative methods.

In addition to the recruiting criteria, there is no clear and well-accepted diagnostic criteria for the term of "musculoskeletal disorders and pain". Therefore, this review was specifically designed to include the relevant papers where the participants were recruited based on one of the following indications: the participants identified themselves as having musculoskeletal disorders and pain, having participant screening processes that were able to identify those people who were symptomatic with musculoskeletal disorders and pain, having objective measurements that included but were not limited to electromyography (EMG), muscle strength or cross-sectional area of muscles that could detect change in musculoskeletal functions (either in comparison to base-line measurement or while performing the assigned task).

Data Extraction and Management

The papers were initially screened and analyzed on titles and abstracts by independent reviewers (AE and SV). Where there was any doubt, the full text was read to determine if inclusion criteria were met. Studies that failed to meet the selection criteria were excluded. The data extraction form was applied from the PECO questions on population, exposure, comparison, and outcomes.[16]

Methodological Quality

There appears no validated checklist or scale available to assess the methodological quality of the cross-sectional experimental laboratory studies in the literature.[17] Therefore, the Downs and Black checklist[18] was modified based on the previous studies[19,20] and used to assess methodological quality of the included studies. The modified Downs and Black checklist was developed that all items were scored 0 to 1, except the item number 5 with a score 0 to 2 and the item number 27 that the score was changed from a scale of 0 to 5 (unclear wording and difficult to score) to a scale of 0 to 1 (where 1 was scored if a power calculation or sample size calculation was present while 0 was scored if there was no power calculation, sample size calculation or explanation whether the number of subjects was appropriate).

Two reviewers (AE and SV) independently scored the quality of each study. Disagreements were resolved by consensus or by a third reviewer (LR). The possible range of reporting quality summary scores was 0 to 28. There is no formal cut-off point to separate the level of quality scores in the modified Downs and Black checklist. Therefore, as recommended by the previous reviews,[20] Quality scores above 19 were considered as "good," between 11 and 19 as "moderate," and below 11 as "poor".

Results

Selection of the study

The flowchart in Fig. 1 illustrates the selection process of the included studies. 639 reports were identified from the electronic databases (AMED = 64, CINAHL = 265, PubMed = 153, ProQuest = 70 and ScienceDirect = 87). Of these publications, 609 were excluded due to an irrelevant title and abstract. Duplications were also excluded, leaving 28 studies. The selection criteria of this systematic review were then applied and 17 more studies were excluded.[6,12,15,21–34] Following this selection process, 11 papers were eligible to be included in the review.[35–45] Additionally, a reference search was conducted using the reference lists of relevant papers to retrieve any missing references. Consequently, a paper written by Akkaya *et al.*[46] was added to the review. Therefore, the total number of studies included in the review was 12.[35–46]

Study characteristics

The main characteristics of the 12 studies are presented in Table 1.[35–46] All the included studies were cross-sectional experimental laboratory studies, which provided data collected form a total of 755 subjects. When considering the inclusion criteria for the studies, four papers used the term "university students," ($n = 406$),[35,37,39,42] three papers used the term "healthy (normal) adult" ($n = 214$),[36,44,46] four papers used the term "young adult" ($n = 125$)[38,40,41,45] and one paper specifically included only right-handed female subjects in their study ($n = 10$).[43]

Considering the inclusion criteria quoted in the papers, seven studies failed to provide a clear list of inclusion criteria.[35–38,43,44,46] Whereas, three

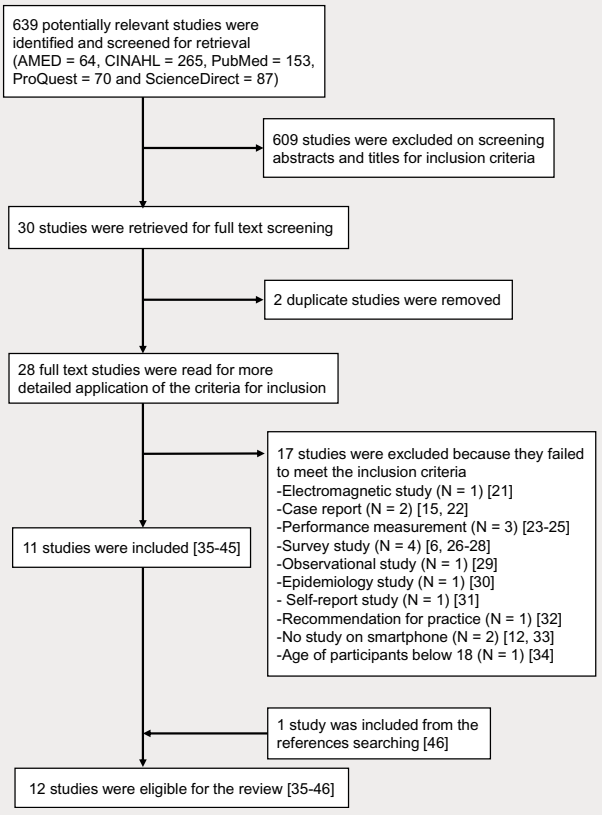

Fig. 1. Flowchart for the selection of studies.

studies indicated the amount of experience with a touch screen smartphone,[40,41,45] one study specifically included only participants aged between 18 to 29 years,[39] one study used the term "use mobile phone regularly" as an inclusion criteria.[42] Only one study by Xie *et al.*[45] demonstrated well-constructed inclusion criteria with an intention to recruit participants with similar characteristics. For the exclusion criteria, 10 studies excluded participants with experience of injury, trauma, deformity, surgery and/or any neurological condition that affected head, neck, and upper limbs.[36–39,41–46] However, participants who had any physical difficulty were excluded in Lee *et al.*,[40] but this term was not defined. There was one study which did not indicate any exclusion criteria.[35]

Regarding the study intervention, six studies had no comparison group.[35,36,39,40,42,43] Of these, two studies focused on the thumb area.[35,39] Xiong and Murasaki[35] used EMG to assess thumb performance and muscular activity of the thumb (AdP: adductor pollicis, FPB: flexor pollicis brevis, APB: abductor pollicis brevis, APL: abductor pollicis longus, FDI: first dorsal interosseous, and ED: extensor digitorum) while Eapen *et al.*[39] used ultrasound to evaluate the diameter of the thumb tendons (APL, EPB: extensor pollicis brevis, EPL: extensor pollisis longus, and FPL: flexor pollicis longus). Three studies focused on the effect of head and neck positioning during smartphone use in different positions; sitting position (lap and desk posture)[36]; standing position (using and without using smartphone)[42]; and sitting versus standing posture while using smartphone.[40] Another study used EMG to assess the neck (UT: upper trapezius) and thumb muscle (EPL and AbP: abductor pollicis) activity in sitting to compare the muscle activity between one and two hands smartphone use.[43] Four studies had a comparison group and of these; two studies compared the range of motion (ROM)[41] and muscular activity[45] in neck pain and non-pain groups; the other two studies compared the ROM between frequent and infrequent smartphone users.[44,46] The study by Inal *et al.*[37] had three groups for comparison (non, low, and high smartphone user) and used the ultrasonographic assessment of the FPL muscle and the median nerve. Another study with three-group comparison compared pain threshold and the muscle activity during smartphone use, computer use and in a control group.[38]

Table 1. Characteristic of the methodology, outcome measurement and finding included in the review.

References	PECO				Recruitment criteria	Reported finding
	Participants (n)	Exposure	Comparison	Outcome		
Akkaya *et al.* (2015)[46]	Healthy adults (149; 36 male and 86 female)	Using VAS to measure thumb pain during texting and using ROM to measure MCP joint and IP joint of thumb, grip strength, pinch strength, and ultrasonographic evaluation of FPL tendon in different groups.	Frequent users versus infrequent users	VAS, pinch strength, grip strength, ROM, and ultrasonographic evaluation of FPL tendon	I: N/A E: Experience of injury of fracture to nerve, vessel and tendon and rheumatic disease.	FPL tendons were larger and had more pain on the texting side in frequent texter group.
Eapen C *et al.* (2014)[39]	Students (98)	Using ultrasound evaluation on APL, EPB, EPL, FPL and Thenar eminence.	N/A	Finklestein test, pinch strength and APL, EPB, EPL, FPL, and Thenar eminence diameter	I: Age 18 to 29 years E: Past experience of inflamative and degenerative of neuromuscular conditions of thumb, hand, and UE due to other activity instead of smartphone use.	Participants express musculoskeletal-related symptoms such as tenderness on extensor compartments, positive Finklestein test, pain on abduction and extension of thumb and increased fluid around dorsal compartment.
Guan X *et al.* (2015)[42]	University students (186; 105 male and 81 female)	Using photographic analysis to measure sagittal posture of head tilt angle, neck tilt angle, forward head shift, and gaze angle during smartphone use in different standing conditions.	Standing while using smartphone versus standing while not using smartphone	Head tilt angle, neck tilt angle, forward head shift, and gaze angle	I: Using smartphone regularly. E: Experience of craniofacial, cervical, shoulder, thoracic, and spine pain.	Head tilt angle and forward head posture were significantly increased during mobile phone use whereas neck tilt angle was decreased.
iNAL EE *et al.* (2015)[37]	University students (102; 30 male and 72 female)	Using self-report hand function questionnaire, clinical evaluation, and ultrasonographic assessment to measure hand performances in different groups.	Non-users versus low-users versus high-users	VAS, grip strength, pinch strength, median nerve ration, and FPL ratio	I: N/A E: Experience of neuropathy, radiculopathy, previous contracture, and lateral or medial epicondylitis.	High smartphone users had significantly larger median nerve CSA, less pinch strength, and hand function in dominant hand.

Table 1. (*Continued*)

References	PECO				Recruitment criteria	Reported finding
	Participants (n)	Exposure	Comparison	Outcome		
Jung SI *et al.* (2016)[44]	Healthy adults (50)	Using craniovertebral angle, scapular index, and respiratory function assessment to measure body functions in different groups.	Frequent users versus infrequent users	Craniovertebral angle, scapular index, FVC, FEV1, ratio of FEV1/FVC, and peak expiratory flow	I: N/A E: Past experience of pain, trauma, fracture or surgery to cervical, thoracic, and abdominal area, neurological disorders, lung function restriction, unstable cardiac conditions, recently smoking or smoker free within five years.	Long duration of smartphone use negatively affected the posture and respiratory function, especially peak expiratory flow.
Kim GY *et al.* (2012)[38]	Young adults (40; 17 male and 23 female)	Using pain pressure threshold and EMG to measure dominant UT, brachioradialis, FCU, and APB in different groups.	Smartphone users versus computer users versus control	Pressure pain threshold and EMG	I: N/A E: Past experience of injury, surgery or deformity of spine and UE, visual problems, dizziness, vertigo, neurological disorders and using sedative drug within 48 h.	Smartphone users showed statistically significant different in brachioradialis muscle fatigue while computer user shows statistically significant different in UT muscle fatigue. Both experimental groups showed significant reduction in pressure pain threshold of UT muscle.
Kim MS (2015)[41]	Young adults (27; 12 male and 15 female)	Using ROM to measure cervical angle during two-hand texting in sitting position.	Pain versus control	Upper and lower cervical ROM	I: At least one year experience of using smartphone. E: Past experience of neck pain, spinal trauma, cervical surgery, fibromyalgia, and systematic or connective tissue disorder.	Neck flexion angle was increased with time during smartphone use both on upper and lower cervical spines. Neck pain group was found to have greater angle.
Lee M *et al.* (2015)[43]	Right-handed female (10)	Using EMG and dolorimeter to measure muscle activity and tenderness in UT, EPL, and AbP during different conditions of smartphone use on thigh in sitting position.	One-handed smartphone use versus two-handed smartphone use	EMG and pressure pain threshold	I: N/A E: Past experience of UE ROM limitation and orthopedic problems.	One-handed smartphone use showed higher muscular activity in UT, AbP, and EPL.

Table 1. (*Continued*)

References	PECO				Recruitment criteria	Reported finding
	Participants (n)	Exposure	Comparison	Outcome		
Lee S *et al.* (2015)[40]	Young adults (18; 9 male and 9 female)	Using ROM to measure head flexion angle during text messaging, web browsing, and video watching in different position.	Sitting position versus standing position	Head flexion angle	I: At least one year experience of using smartphone. E: Physical difficulties of using smartphone.	Head flexion angle was the highest during text messaging in sitting.
Shin H & Kim K (2014)[36]	Healthy adults (15)	Using VAS, EMG, and ROM to measure cervical erector spinae during smartphone use in different posture.	Desk posture versus lap posture	Flexion relaxation ratio, ROM, and VAS	I: N/A E: Past experience of neck pain, spinal trauma, and cervical surgery.	Sustained smartphone use in lap posture could influence neck pain.
Xie Y *et al.* (2016)[45]	Young adults (40; 16 male and 24 female)	Using EMG, discomfort score, and borg scale to measure on cervical erector spinae, UT, LT, ECR, ED, FDS, and APB during smartphone and computer use in different groups.	Pain versus non-pain	EMG, discomfort score, and rate of perceived exertion	I: Right-handed users with similar texting speed who spent at least 2 h daily using smartphone for the last six months. E: Past experience of pain, trauma, fracture or surgery to cervical and UE, neurological and systematic disorders.	Participants with neck–shoulder pain showed higher muscle activity in cervical erector spinae and UT muscle during texting and typing tasks. Unilateral texting showed higher muscle loading in forearm muscles when compared to bilateral texting.
Xiong J & Murasaki S (2014)[35]	Right-handed university students (20; 10 male and 10 female)	Using pressure sensor and EMG to measure thumb performance and muscular activity during smartphone use in different button size and speed.	Small button versus large button	Thumb performance, iEMG, contraction time and iEMG/s	I: N/A E: N/A	Smaller button negatively affects thumb performance.

Notes: APB: abductor pollicis brevis; APL: abductor pollicis longus; ECR: extensor carpi radialis; ED: extensor digitorum; EMG: electromyography; EPB: extensor pollicis brevis; EPL: extensor pollicis longus; FCU: flexor carpi ulnaris; FDS: flexor digitorum superficialis; FEV1: force expiratory volume at 1 s; FPL: flexor pollicis longus; FVC: force vital capacity; iEMG: integrated electromyography; LT: lower trapezius; N/A: non-applicable; n: number; ROM: range of motion; SAS: smartphone addiction scale; UT: upper trapezius; VAS: visual analog scale.

In this systematic review, it is not possible to perform a meta-analysis due to the heterogeneity of the study designs and outcome measures.

Methodological Quality

Table 2 presents the methodological quality results from the modified Downs and Black checklist. All studies[35–46] included in this review were rated as "moderate" (ranged from 11 to 18). All studies[35–46] failed to provide information about representativeness of the population and the intervention as well as adverse events, subjects recruiting periods, blinding (both subjects and assessors) and randomization (allocation and concealment). The study by Xiong and Murasaki[35] did not provide information about the participants' characteristics. Six studies[37,38,41,44–46] partially reported information regarding principal confounders. One study[40] failed to report the descriptive statistics from the raw data percentiles was reported but not the mean and standard deviation of the measured variable and also their main confounders were not investigated. The actual p-value of the main outcomes (0.05 rather than < 0.05) was reported in eight studies.[35–39,42,45,46] Six studies[36–38,40,43,46] had no information about source of population and their recruitment processes. Compliance with the intervention was not mentioned in six studies.[35,36,39,40,42,43] Only a study by Akkaya[46] provided a statement of recruitment period. All studies with the exception of one[39] failed to conduct a power calculation.

Findings

The outcome of the studies can be divided into seven categories: EMG, ROM, Pain, finger and hand performance, tendon diameter, and subjective measures of discomfort and exertion.

Electromyography

Four studies used EMG to assess muscular activity.[35,38,43,45] Comparing between smaller buttons and larger buttons, Xiong and Muraki[35] found that using smaller buttons significantly increased the muscle activity of the FDI muscle ($p < 0.01$) and significantly decreased the muscle activity of the APB muscle ($p < 0.01$). Kim *et al.*[38] found that after a smartphone typing task, when compared to the control group, there was a statistically significant decrease in the median frequencies of the brachioradialis muscle ($p < 0.05$). Lee *et al.*[43] discovered that the muscular activity of the UT, ELP and AbP muscle was significantly higher when using the smartphone in one hand than in two hands ($p < 0.05$). Xie *et al.*[45] found that participants with neck and shoulder pain had significantly higher muscular activity in the cervical erector spinae and UT muscles than non-symptomatic participants when performing a texting and typing task. Xie *et al.*[45] also found that one-hand texting produced significantly more muscle activity of the forearm muscles than two-hand texting.

Range of motion

Five studies used ROM of the head and neck or the thumb and hand as an assessment to evaluate the change in posture during and after the smartphone use.[36,40–42,44] Shin and Kim[36] found an average change of $44 \pm 4.31°$ in ROM of cervical flexion in the lap posture when compared to the baseline measurements. Lee *et al.*[40] concluded that the cervical flexion angle was significantly larger when text messaging than when carrying out the other tasks (web browsing and video watching) ($p < 0.05$) and significantly larger in sitting than in standing ($p < 0.05$). When using the smartphone in a sitting position, one study[41] discovered that the upper and lower cervical flexion angles were significantly higher in the neck pain group than in the control group ($p < 0.05$). In addition, another study[42] compared the head and neck posture in standing with and without looking at the smartphone. They found that participants who were standing and looking at the smartphone had significantly increased the head tilt angle and forward head shift ($p < 0.05$) while significantly decreased the neck tilt angle ($p < 0.05$). Jung *et al.*[44] also found that frequent smartphone users have higher scapular index and craniovertebral angle ($p < 0.05$) compared to infrequent smartphone users.

Pain

Measures of pain were presented in five studies.[36–38,43,46] Shin and Kim[36] presented the change of mean value measured using a visual analog scale (VAS) after using a smartphone in a desk and lap posture from 0 (baseline measurement) to 1.7 and

Table 2. An assessment of methodological quality of studies assessed by modified Downs & Black checklist.

	Check list																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	
References	Study aim	Main outcome	Subject characteristics	Description of intervention	Principal confounders	Outcome data	Range of results	Adverse events	Lost to follow-up	Probability value (exact)	Source population	Representative of population	Staff, place, facility	Subjects blind to intervention	Blind assessors	Data dredging	Same length of follow-up	Appropriate statistical tests	Compliance with the intervention	Accurate outcome measure	Control recruited same	Recruitment at same time	Randomized allocation	Concealed randomization	Adjustment for confounders	Subjects lost to follow-up	Power	Total
Akkaya *et al.* (2015)[46]	Y	Y	Y	Y	P	Y	Y	N	Y	Y	N	N	N	N	N	Y	Y	Y	Y	Y	N	Y	N	N	Y	Y	0	17/28
Eapen C *et al.* (2014)[39]	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	N	N	N	Y	Y	Y	N	Y	Y	N	N	N	Y	Y	1	17/28
Guan X *et al.* (2015)[42]	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	N	N	N	Y	Y	Y	N	Y	Y	N	N	N	Y	Y	0	16/28
iNAL EE *et al.* (2015)[37]	Y	Y	Y	Y	P	Y	Y	N	Y	Y	N	N	N	N	N	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y	0	17/28
Jung SI *et al.* (2016)[44]	Y	Y	Y	Y	P	Y	Y	N	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y	0	17/28
Kim GY *et al.* (2012)[38]	Y	Y	Y	Y	P	Y	Y	N	Y	Y	N	N	N	N	N	Y	Y	Y	Y	Y	N	N	N	N	Y	Y	0	16/28
Kim MS (2015)[41]	Y	Y	Y	Y	P	Y	Y	N	Y	N	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y	0	17/28
Lee M *et al.* (2015)[43]	Y	Y	Y	Y	N	Y	Y	N	Y	N	N	N	N	N	N	Y	Y	Y	N	Y	N	N	N	N	Y	Y	0	13/28
Lee S *et al.* (2015)[40]	Y	Y	Y	Y	N	Y	N	N	Y	N	N	N	N	N	N	Y	Y	Y	N	Y	N	N	N	N	N	Y	0	11/28
Shin H & Kim K (2014)[36]	Y	Y	Y	Y	N	Y	Y	N	Y	Y	N	N	N	N	N	Y	Y	Y	N	Y	N	N	N	N	Y	Y	0	14/28
Xie Y *et al.* (2016)[45]	Y	Y	Y	Y	P	Y	Y	N	Y	Y	Y	N	N	N	N	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y	0	18/28
Xiong J & Murasaki S (2014)[35]	Y	Y	N	Y	N	Y	Y	N	Y	Y	Y	N	N	N	N	Y	Y	Y	N	Y	Y	N	N	N	Y	Y	0	15/28

Notes: *Items 1 to 27 of the modified Downs & Black checklist. "Y": the answer is yes; "N": the answer is no; "U": the answer is unable to determine; "P": the answer is partial. The question number 5 will assign a score of "0" if the answer is "No", "1" if the answer is "Partial", and "2" if the answer is "Yes". The question number 27 will assign a score of "0" if no power calculation is provided, and "1" if a power calculation is provided. All the questions except the question numbers 5 and 27 will assign a score of "0" if the answer is "No" or "Unable to determine", and "1" if the answer is "Yes". Total quality scores of studies: Less than 11 = poor 11–19 = moderate; Higher than 19 = good.

5.2, respectively. Inal *et al.*[37] found that frequent smartphone users had significantly higher VAS scores than the infrequent and non-user groups ($p < 0.05$) but found no difference between non-users and infrequent users. Two studies[38,43] concluded that the pain threshold of the UT muscle decreased significantly after smartphone use ($p < 0.01$). Lee *et al.*[43] also found that one-hand smartphone use significantly increased muscle tenderness compared to two-hand use ($p < 0.01$). Akkaya *et al.*[46] showed a statistically significant difference ($p = 0.005$) in the VAS scores between the texting side (0.3 ± 0.9) and the contralateral side (0.01 ± 0.1) in a frequent texter group.

Thumb–finger–hand performance

Four studies assessed the performance of the thumb, finger, and hand.[35,37,39,45] Xiong and Muraki[35] indicated that using a small button leads to significant shorter fatigue times than when using a large button ($p < 0.01$) in a tapping task, while the tapping speed found to be significantly slower in flexion–extension than in abduction–adduction of the thumb during a moving task ($p < 0.01$). Inal *et al.*[36] presented a correlation between pinch strength and smartphone addition scale (SAS) ($p = 0.022$, $r = -0.281$; negatively weak correlation), pinch strength and duration of smartphone use ($p = 0.288$, $r = 0.133$; weak correlation), and pinch strength with Duruoz hand index score ($p = 0.014$, $r = -0.242$; negatively weak correlation). Eapen *et al.*[39] reported the significant reduction in tip ($p = 0.002$) and lateral ($p = 0.02$) pinch grip strength in patients with thumb pain while text messaging when compared to the control group.

Tendon–nerve diameter

Three studies evaluated the thickness of the tendon and nerve in symptomatic[39] and non-symptomatic smartphone users.[37,46] Eapen *et al.*[39] applied ultrasound evaluation to the thumb area of the symptomatic subjects and found fluid around the thumb tendons at the wrist level (19%) and in the flexor muscles of the thumb (2%). Two studies[37,46] discovered that the frequent smartphone users had significantly larger FPL tendons ($p = 0.001$)[46] and median nerves ($p < 0.001$)[37] than the infrequent smartphone users.

Discomfort and exertion level

Only two studies investigated the discomfort and exertion level.[35,45] One reported[45] a significant change in the discomfort scores ($p = 0.008$) as well as the rate of perceived exertion ($p < 0.001$) after performing the texting task. This effect was greater in the symptomatic group than in the control group. Another study[35] reported that smaller button size leads to a significantly higher rating of perceived exertion (using the Borg scale) of the FDI muscle in the tapping task. Moreover, they found a significant decrease of perceived exertion score of the APB and APL muscles and a significant increase of perceived exertion score of the FDI muscle in the moving task.

Discussion

This systematic review has provided information about the change[37–39,46] and associations with musculoskeletal symptoms[35,36,40–45] in the neck, the shoulder, the upper limb, the hands and the thumb associated with smartphone use. The findings of all studies emphasized that the use of smartphone may contribute to the musculoskeletal symptoms.

Methodological Quality of Studies

The methodological quality of the studies included in this review was scored as moderate. This may be due to the nature of cross-sectional experimental laboratory studies where blinding and randomization are hard to implement.[47] In addition, more than half of the included studies[35–39,41–46] simulated the smartphone use conditions for participants to perform in the laboratory setting. Accordingly, these data may not represent the actual smartphone use in real life and therefore the studies have low external validity.[48] Half of the studies[35,36,39,40,42,43] were lacking information regarding confounding variables, source of population and how they were recruited which, therefore, exposing to high risk of selection bias (low internal validity). The presence of low internal and external validity resulted in some concerns about the applicability of the study results.[48] Moreover, half of the studies included in this review[35,36,39,40,42,43] did not provide sufficient information in order to effectively assess the comparability of the intervention and comparison groups. This notion made it

difficult to analyze whether the change and associations with musculoskeletal symptoms found in the study groups really originated from smartphone use, or from other factors. Moreover, almost all studies included in this review did not attempt to address potential sources of bias.[35–38,40–46] Finally, only one study[39] mentioned that their sample size was based on data from the pilot study while the rest of the studies[35–38,40–46] did not mention a power calculation.

Consequently, the study quality scores were moderate. However, the issues identified above must be taken into account when interpreting the results of the studies included in this review.

Overall Findings

The studies included in this review[35–46] reported their finding in three specific body regions: the head–neck, shoulder–arm, and hand–thumb.

The findings of this review suggest that using smartphone may induce musculoskeletal symptoms in the neck.[36,40–42,44,45] During smartphone use, the muscle activity of UT, erector spinae and the neck extensor muscles are increased,[43,45] especially for those who already have pain in the neck region.[45] Moreover, many studies found that neck flexion angle, head tilt angle and forward head shifting were increased during the smartphone use[36,40–42,44] and also increased with the duration of smartphone use.[40,41] Many studies suggested that people with pain in the neck region tended to adopt a more flexed posture than those who have no pain[41,44,45] which negatively affected the neck posture.[44] This could be explained by the theory that the motor control of the neck muscles was altered by prolonged poor neck posture during the use of smartphones.[49,50] In addition, the variation of the head–neck angle could possibly depend on the task, the posture and the way of holding the smartphone.[6,40] The recent review concluded that smartphone use in a sitting position seems to cause more shift in head–neck angle than in a standing position.[36,40] A possible explanation is that postural stability is associated with the head position and movement in standing, since neck flexion or extension in an upright posture in standing can alter the postural stability.[51] Therefore, when the smartphone is used in a standing position, the user tends to minimize the alternations in neck posture to avoid postural instability.[40]

For the shoulder–arm region, muscle activity increased and the pain pressure threshold decreased in the shoulder and forearm area when using a smartphone.[38,43,45] This is because the increase in muscle activity is associated directly with the rise of muscle fatigue[52,53] and the reduction of pain pressure threshold.[54,55] The repeated upper limb movements during smartphone use activate a continuous muscle contraction which may cause microscopic damage to the muscle which is the risk factor for musculoskeletal disorders.[38,43,56]

For the hand–thumb region, this review also found that one-handed smartphone use may cause more musculoskeletal symptoms in the shoulder–arm and the hand–thumb areas than using two hands to operate a smartphone.[43–45] The reason is that two-handed smartphone use allowed more effective cooperation between holding and conducting the smartphone tasks which resulted in improving the task performance and variation in movements.[25,55] Thus, less muscle activity was found in two-hand smartphone use when compare to one-hand smartphone use (less stereotypical and repetitive movements).[25,43–45] Consequently, to reduce the risk of musculoskeletal problems, using two hands to operate a smartphone is recommended.[25,43]

Furthermore, this review also revealed that the frequent smartphone users had reduced thumb performance when compared to the infrequent users,[37,39] especially, when performing sensitive tasks or tapping on a small button.[35] Additionally, this study detected changes in the tendon, nerve and space between muscular tissue in frequent smart phone users.[37,39] Practically, smartphone users naturally adjust their hand and thumb postures to fit with the phone layout which may alter their efficiency of smartphone use. The prolonged altered static posture and repetitive use of the wrist and thumb during smartphone operation may negatively impact the muscular and nervous tissue in the hand.[57] Excessive repetitive or static use of wrist and thumb movements during the smartphone use can increase the load on the joints,[1,6,57] increase carpal tunnel pressure,[58] and decrease the space available for the median nerve to move.[59] Thus, leading to the acute trauma and causing the enlargement of the median nerve[59–62] and muscular tendon (e.g., FPL tendon).[46] Accordingly, the structural changes from frequent smartphone usage may aggravate pain[36,37,43,46] which was also reported more frequently in the group of frequent smartphone users than the group of infrequent smartphone users.

Limitations of the Review

This review was based on a comprehensive search of all the evidence that relates to the research question and adheres to the inclusion and exclusion criteria set. However, there were some limitations to the data found.

This review only included publications that were published in English, leading to missing evidence that has been published in other languages. There may be some possibility of publication bias because all reports presented more positive outcomes on musculoskeletal change than null results which may indicate overestimation of the positive outcomes. In addition, the power calculations were not reported and the research design and outcome measures were different between studies. There are some issues that lower the quality of the included studies. Most studies were done on university students or young healthy adults. Consequently, the research cannot be generalized to people of all ages. Furthermore, inclusion and exclusion criteria were not explicit enough to recruit participants with similar characteristics and did not mention existing poor postures or personal habits that might affect the association between the use of smartphone and measured parameters. Additionally, the gender issue has not been addressed. The intervention and task simulations designed may not represent the use of smartphones in real life as it appears that short duration tasks and standardized posture were used in the laboratory setting. The model of smartphones used in each study were different and, moreover, the role of examiners in all studies was not clearly described and intra- and inter-rater reliability were not reported.

Implication for Further Research

Future primary research should use publication guidelines, for example, CONSORT or STROBE, to improve the reporting quality and study design. Research planning should focus initially on the issue of study quality and study validity. More clinical trials with comparison groups are needed to further improve the strength of the evidence and to identify the most suitable method of assessing the musculoskeletal changes due to the use of smartphones.

Conclusion

This systematic review revealed that the use of smartphones may contribute to the occurrence of clinical and subclinical musculoskeletal changes as well as associated factors in the head–neck, shoulder–arm, and hand–thumb areas. While there is a strong case presented in the findings of all the studies reported in this review, the evidence must be considered in the light of the moderate scores from the modified Downs and Black checklist.

Conflict of Interest

All authors declare that they have no conflict of interest.

Funding/Support

No funding was received for this systematic review.

Author Contributions

AE is the leading reviewer who contributed to the conception and design of the study. AE and SV contributed to the development of the search strategy, conducted the systematic search, extracted the data, and performed the data analysis. All authors assisted with the interpretation, prepared the manuscript, drafted and revised the final paper. LR contributed to the proof reading of the whole manuscript. All authors approved the final submitted version of the manuscript.

References

1. Jonsson P, Johnson PW, Hagberg M, Forsman M. Thumb joint movement and muscular activity during mobile phone texting — A methodological study. J Electromyogr Kinesiol 2011;21(2):363–70.
2. Goggin G. Cell phone culture: Mobile technology in everyday life. New York: Routledge, 2012.
3. International Telecommunication Union. Measuring the Information Society, 2012. Available at http://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2012/MIS2012_without_Annex_4.pdf. Accessed September 2016.
4. Khalaf S. Mobile use grows 115% in 2013, propelled by messaging apps. Flurry Analytics 2014.

5. Madge C, Meek J, Wellens J, Hooley T. Facebook, social integration and informal learning at university: 'It is more for socialising and talking to friends about work than for actually doing work'. Learn Media Technol 2009;34(2):141–55.
6. Berolo S, Wells RP, Amick BC. Musculoskeletal symptoms among mobile hand-held device users and their relationship to device use: A preliminary study in a Canadian university population. Appl Ergon 2011;42(2):371–8.
7. Gold JE, Driban JB, Yingling VR, Komaroff E. Characterization of posture and comfort in laptop users in non-desk settings. Appl Ergon 2012;43 (2):392–9.
8. Maniwa H, Kotani K, Suzuki S, Asao T. Changes in posture of the upper extremity through the use of various sizes of tablets and characters. In: Int Conf Human Interface and the Management of Information. Berlin, Heidelberg: Springer, 2013;89–96.
9. Bababekova Y, Rosenfield M, Hue JE, Huang RR. Font size and viewing distance of handheld smart phones. Optom Vis Sci 2011;88(7):795–7.
10. Kwon M, Lee JY, Won WY, et al. Development and validation of a smartphone addiction scale (SAS). PloS One 2013;8(2):e56936.
11. Raffle AE, Mackenzie EF. Management of cervical dyskaryosis. No easy answer. BMJ 1994;309(6949):270.
12. Gustafsson E, Johnson PW, Hagberg M. Thumb postures and physical loads during mobile phone use — A comparison of young adults with and without musculoskeletal symptoms. J Electromyogr Kinesiol 2010;20(1):127–35.
13. Barr AE, Barbe MF, Clark BD. Work-related musculoskeletal disorders of the hand and wrist: Epidemiology, pathophysiology, and sensorimotor changes. J Orthop Sports Phys Ther 2004;34(10):610–27.
14. Ming Z, Pietikainen S, Hänninen O. Excessive texting in pathophysiology of first carpometacarpal joint arthritis. Pathophysiology 2006;13(4):269–70.
15. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Ann Intern Med 2009;151(4):264–9.
16. Mattos CT, Ruellas AC. Systematic review and meta-analysis: What are the implications in the clinical practice? Dental Press J Orthod 2015;20(1):17–9.
17. Zeng X, Zhang Y, Kwong JS, et al. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: A systematic review. J Evid-Based Med 2015;8(1):2–10.
18. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomised studies of health care interventions. J Epidemiol Community Health 1998;52:377e84.
19. Eng J, Teasell R, Miller W, et al. Spinal cord injury rehabilitation evidence: Method of the SCIRE systematic review. Top Spinal Cord Inj Rehabil 2007;13(1):1–10.
20. Methajarunon P, Eitivipart C, Diver CJ, Foongchomcheay A. Systematic review of published studies on aquatic exercise for balance in patients with multiple sclerosis, Parkinson's disease, and hemiplegia. Hong Kong Physiother J 2016;35:12–20.
21. Vecsei Z, Csathó Á, Thuróczy G, Hernádi I. Effect of a single 30 min UMTS mobile phone-like exposure on the thermal pain threshold of young healthy volunteers. Bioelectromagnetics 2013;34(7):530–41.
22. Gilman L, Cage DN, Horn A, Bishop F, Klam WP, Doan AP. Tendon rupture associated with excessive smartphone gaming. JAMA Intern Med 2015;175(6):1048–9.
23. Kietrys DM, Gerg MJ, Dropkin J, Gold JE. Mobile input device type, texting style and screen size influence upper extremity and trapezius muscle activity, and cervical posture while texting. Appl Ergon 2015;50:98–104.
24. Lee J, Seo K. The comparison of cervical repositioning errors according to smartphone addiction grades. J Phys Ther Sci 2014;26(4):595–8.
25. Trudeau MB, Asakawa DS, Jindrich DL, Dennerlein JT. Two-handed grip on a mobile phone affords greater thumb motor performance, decreased variability, and a more extended thumb posture than a one-handed grip. Appl Ergon 2016;52:24–8.
26. Stalin P, Abraham SB, Kanimozhy K, Prasad RV, Singh Z, Purty AJ. Mobile phone usage and its health effects among adults in a semi-urban area of Southern India. J Clin Diagn Res 2016;10(1):LC14.
27. Kim HJ, Kim JS. The relationship between smartphone use and subjective musculoskeletal symptoms and university students. J Phys Ther Sci 2015;27(3):575.
28. Shan Z, Deng G, Li J, Li Y, Zhang Y, Zhao Q. Correlational analysis of neck/shoulder pain and low back pain with the use of digital products, physical activity and psychological status among adolescents in Shanghai. PloS One 2013;8(10):e78109.
29. Sharan D, Ajeesh PS. Risk factors and clinical features of text message injuries. Work 2012;41 (Supplement 1):1145–8.
30. Oftedal G, Wilen J, Sandström M, Mild KH. Symptoms experienced in connection with mobile phone use. Occup Med 2000;50(4):237–45.
31. Korpinen LH, Pääkkönen RJ. Self-report of physical symptoms associated with using mobile phones and other electrical devices. Bioelectromagnetics 2009;30(6):431–7.
32. Gustafsson E. Ergonomic recommendations when texting on mobile phones. Work 2012; 41(Supplement 1): 5705–6.

33. Gustafsson E, Johnson PW, Lindegård A, Hagberg M. Technique, muscle activity and kinematic differences in young adults texting on mobile phones. Ergonomics 2011;54(5):477–87.
34. Kee IK, Byun JS, Jung JK, Choi JK. The presence of altered craniocervical posture and mobility in smartphone-addicted teenagers with temporomandibular disorders. J Phys Ther Sci 2016;28(2):339.
35. Xiong J, Muraki S. An ergonomics study of thumb movements on smartphone touch screen. Ergonomics 2014;57(6):943–55.
36. Shin H, Kim K. Effects of cervical flexion on the flexion-relaxation ratio during smartphone use. J Phys Ther Sci 2014;26(12):1899.
37. İNal EE, ÇetİntürkA, Akgönül M, Savaş S. Effects of smartphone overuse on hand function, pinch strength, and the median nerve. Muscle Nerve 2015;52(2):183–8.
38. Kim GY, Ahn CS, Jeon HW, Lee CR. Effects of the use of smartphones on pain and muscle fatigue in the upper extremity. J Phys Ther Sci 2012;24(12):1255–8.
39. Eapen C, Kumar B, Bhat AK, Venugopal A. Extensor pollicis longus injury in addition to De Quervain's with text messaging on mobile phones. J Clin Diagn Res 2014;8(11):LC01.
40. Lee S, Kang H, Shin G. Head flexion angle while using a smartphone. Ergonomics 2015;58(2):220–6.
41. Kim MS. Influence of neck pain on cervical movement in the sagittal plane during smartphone use. J Phys Ther Sci 2015;27(1):15.
42. Guan X, Fan G, Wu X, et al. Photographic measurement of head and cervical posture when viewing mobile phone: A pilot study. Eur Spine J 2015;24(12):2892–8.
43. Lee M, Hong Y, Lee S, et al. The effects of smartphone use on upper extremity muscle activity and pain threshold. J Phys Ther Sci 2015;27(6):1743.
44. Jung SI, Lee NK, Kang KW, Kim K, Lee DY. The effect of smartphone usage time on posture and respiratory function. J Phys Ther Sci 2016;28(1):186.
45. Xie Y, Szeto GP, Dai J, Madeleine P. A comparison of muscle activity in using touchscreen smartphone among young people with and without chronic neck–shoulder pain. Ergonomics 2016;59(1):61–72.
46. Akkaya N, Dogu B, Ünlü Z, et al. Ultrasonographic evaluation of the flexor pollicis longus tendon in frequent mobile phone texters. Am J Phys Med Rehabil 2015;94(6):444–8.
47. Mann CJ. Observational research methods. Research design II: Cohort, cross-sectional, and case-control studies. Emerg Med J 2003;20(1):54–60.
48. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. Plast Reconstr Surg 2010;126(2):619.
49. Szeto GP, Straker LM, O'Sullivan PB. A comparison of symptomatic and asymptomatic office workers performing monotonous keyboard work — 1: Neck and shoulder muscle recruitment patterns. Man Ther 2005;10(4):270–80.
50. Szeto GP, Straker LM, O'Sullivan PB. A comparison of symptomatic and asymptomatic office workers performing monotonous keyboard work — 2: Neck and shoulder kinematics. Man Ther 2005;10(4):281–91.
51. Buckley JG, Anand V, Scally A, Elliott DB. Does head extension and flexion increase postural instability in elderly subjects when visual information is kept constant? Gait Posture 2005;21(1):59–64.
52. Enoka RM, Stuart DG. Neurobiology of muscle fatigue. J Appl Physiol 1992;72(5):1631–48.
53. Allen DG, Lannergren J, Westerblad H. Muscle cell function during prolonged activity: Cellular mechanisms of fatigue. Exp Physiol 1995;80(4):497–528.
54. Persson AL, Hansson GÅ, Kalliomäki J, Moritz U, Sjölund BH. Pressure pain thresholds and electromyographically defined muscular fatigue induced by a muscular endurance test in normal women. Clin J Pain 2000;16(2):155–63.
55. Cook DB, O'Connor PJ, Eubanks SA, Smith JC, Lee MI. Naturally occurring muscle pain during exercise: Assessment and experimental evidence. Med Sci Sports Exerc 1997;29(8):999–1012.
56. Haaland KY, Mutha PK, Rinehart JK, Daniels M, Cushnyr B, Adair JC. Relationship between arm usage and instrumental activities of daily living after unilateral stroke. Arch Phys Med Rehabil 2012;93(11):1957–62.
57. Ko K, Kim HS, Woo JH. The study of muscle fatigue and risks of musculoskeletal system disorders from text inputting on a smartphone. J Ergon Soc Korea 2013;32(3):273–8.
58. Trudeau MB, Young JG, Jindrich DL, Dennerlein JT. Thumb motor performance varies with thumb and wrist posture during single-handed mobile phone use. J Biomech 2012;45(14):2349–54.
59. Gelberman RH, Hergenroeder PT, Hargens AR, Lundborg GN, Akeson WH. The carpal tunnel syndrome. A study of carpal canal pressures. J Bone Joint Surg Am 1981;63(3):380–3.
60. Bower JA, Stanisz GJ, Keir PJ. An MRI evaluation of carpal tunnel dimensions in healthy wrists: Implications for carpal tunnel syndrome. Clin Biomech 2006;21(8):816–25.
61. Wieslander G, Norbäck D, Göthe CJ, Juhlin L. Carpal tunnel syndrome (CTS) and exposure to vibration, repetitive wrist movements, and heavy manual work: A case-referent study. Br J Ind Med 1989;46(1):43–7.
62. Harris-Adamson C, Eisen EA, Kapellusch J et al. Biomechanical risk factors for carpal tunnel syndrome: A pooled study of 2474 workers. Occup Environ Med 2015;72(1).

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) 91–114
DOI: 10.1142/S1013702518500087

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

Effectiveness of surgical versus conservative treatment for carpal tunnel syndrome: A systematic review, meta-analysis and qualitative analysis

Diony Klokkari[1] and Ioannis Mamais[2,3,*]

[1]*Private Physiotherapist, MSc, Nicosia, Cyprus*

[2]*Department of Hygiene, Epidemiology and Medical Statistics*
Medical School, National and Kapodistrian University of Athens, Greece

[3]*Department of Health Science & Department of Life Science*
European University of Cyprus, Cyprus

mamais@ucy.ac.cy

Received 9 October 2016; Accepted 6 October 2017; Published 2 July 2018

Background: Carpal tunnel syndrome (CTS) is the most common entrapment neuropathy of the upper limb. Treatment options include physiotherapy, splinting, steroid injections or surgery.
Objective: To compare the effectiveness of surgical versus conservative treatment for CTS for symptom and functional improvement and improvement of neurophysiological parameters.
Methods: Systematic searches of PubMed and EBSCO host were conducted to identify the studies published between 1990 and 2016, comparing any surgical treatment to any conservative treatment. Participants were adults with a diagnosis of CTS, with symptom duration ranging from 8 months to 3 years. A meta-analysis and a qualitative analysis were conducted to summarize the results of the included studies and establish any agreement between the two.
Results: A total of 15 studies were included in the study and 10 were included in the meta-analysis, involving 1787 wrists. The qualitative and quantitative analyses were consistent with the results of both indicating that surgical treatment leads to a greater improvement of symptoms at six months (mean difference: 0.52, 95%CI 0.27 to 0.78) and a greater improvement of neurophysiological parameters [distal motor latency (mean difference: 0.31, 95%CI 0.06 to 0.56), sensory nerve conduction velocity (mean difference: 3.71 m/s, 95%CI 1.94 to 5.49)]. At 3 months and 12 months, the results were not significant in favor of surgery or conservative treatment.

*Corresponding author.

Conclusion: Conservative treatment for CTS should be preferred for mild and short-term CTS. Surgery is more effective than conservative in CTS, and should be considered in persisting symptoms, taking into account the complications, which are more severe after surgery. Further research should focus on the field of manual therapy and compare it to surgical treatment for CTS.

Keywords: Carpal tunnel syndrome; median nerve entrapment; surgical treatment; conservative treatment; systematic review; meta-analysis.

Introduction

Carpal tunnel syndrome (CTS) remains more common among non-computer-related jobs,[1] despite the increasing usage of computers in recent years, affecting approximately 1–5% of the general population[2–4] and approximately 34% of hospital workers.[5] CTS is a result of compression of the median nerve at the wrist due to the confined space of the carpal tunnel which restricts the movement of the tissues. Any increase in internal or external canal pressure results in neurological impairment with numbness and tingling along the distribution of the nerve. CTS often results from repetitive strain in manual jobs but it is also associated with conditions like rheumatoid arthritis,[6] pregnancy due to water retention,[7] and diabetes mellitus, which increases the likelihood of a symptomatic response in an already compressed nerve.[8] Conservative treatment with physiotherapy, wrist splints, corticosteroid injections, diuretics, vitamin B6 is proposed initially, whereas surgical treatment is reserved for more severe cases with thenar muscle atrophy or after failure of conservative treatment.[9]

Previous systematic reviews have reported an advantage of surgery over conservative treatment for symptom functional improvement.[10,11] These reviews were published in the last decade. More recent trials have been published since, some of which contradict the results suggested in these reviews. Given a substantial number of recent studies, an updated systematic review is justified to incorporate data brought to light since 2011. Furthermore, this review includes both a meta-analysis and a qualitative analysis to formulate conclusions regarding the relative effectiveness of the interventions.

Therefore, this review sought to investigate whether surgical treatment for CTS can lead to greater symptom improvement, greater functional improvement and greater improvement of neurophysiological parameters than conservative treatment, both short term (3 months) and long term (12 months). An additional aim was to establish the extent of agreement between a qualitative analysis of the studies and a meta-analysis.

Material and Methods

Identification and selection of literature

Systematic searches of PubMed and EBSCO host were conducted to identify studies written in English and published between 1990 and 2016 (see Appendix A for full search strategy). Manual searches of reference lists from previous studies were also conducted to ensure that all relevant studies were captured. Studies eligible for inclusion were randomized controlled trials, clinical trials (CTs), prospective and retrospective studies comparing any surgical intervention to any conservative intervention for CTS patients. The outcome measures short-term and long-term improvement of symptoms, functional status and improvement of neurophysiological parameters. Case control studies were excluded. Studies comparing surgical interventions and studies comparing conservative interventions were also excluded as the aim was to compare the relative effectiveness of the two interventions.

Two independent investigators (KD and MI) screened the titles and read the abstracts and the relevant papers were obtained in full text to assess further eligibility. Data extraction was performed by the two investigators independently. The following data were extracted: authors, year of publication, study design, description of the sample, description of the surgical and conservative intervention, duration of study, study outcomes, assessment times, study results and study conclusions.

Quality assessment

The methodological quality of the randomized controlled trials and CTs included in this systematic review were assessed using the CBRG methodological criteria scale proposed by van Tulder *et al.*[12] This scale consists of 11 items and is similar to the PEDro scale, which has good levels of validity and reliability,[13,14] but with clearer operationalization criteria. It addresses the internal validity of the studies in order to minimize the risk of systematic bias (selection bias, performance bias, attrition bias and performance bias). Items relate to the adequacy of the randomization, treatment allocation concealment, baseline similarity of treatment groups, patient, treatment provider and assessor blinding, similarity of co-interventions, adequacy of compliance, adequacy and description of the dropout rate, similar assessment timing across groups and analysis according to intention-to-treat. Each item is accompanied by a strict list and was evaluated with a "yes", "no", or "unclear" (if it did not apply or if it was not mentioned). Each positive answer scored 1 point. The studies were regarded as high quality if the total score of positive answers in the criteria list was seven and above. In consequence, if the total score was below seven, studies were regarded as low quality.

The methodological quality of the prospective and retrospective studies was assessed using the methodological criteria scale proposed by Moga *et al.*[15] using a modified Delphi technique. This scale is made up of 18 items addressing methodological quality (i.e., patient characteristics, adequacy of eligibility criteria, adequacy of intervention, similarity of co-interventions, relevancy and timing of outcome measures) and statistical reporting (i.e., suitability of statistical tests, length and loss of follow-up, random variability, adverse effect and competing interest reporting). Each item was accompanied by a strict list and was evaluated with a "yes" or "no" and each positive answer scored 1 point. The studies were regarded as high quality if the total score of positive answers was 14 and above. If the total score was below 14, then the studies were regarded as low quality. The methodological quality result was used for the formulation of conclusions in the qualitative analysis, which is described further below.

Participants

Studies involving patients diagnosed with CTS irrespective of the cause, the way it was diagnosed, other associated conditions, the age or the sex of the person. Patient characteristics such as age, sex, duration of symptoms were recorded in order to assess heterogeneity between studies.

Intervention

The included studies compared any surgical intervention such as open carpal tunnel release (OCTR) or endoscopic carpal tunnel release (ECTR) to any conservative intervention such as steroid injections, wrist splints, physiotherapy with electrotherapy, exercise or manual therapy or a combination of different modalities.

Outcome measures

The primary outcome was the patient self-reported improvement in symptoms and function measured using the symptom severity scale and functional status scale of the Boston questionnaire (BQ). Secondary outcome measures used to evaluate the effectiveness of the intervention were improvement of neurophysiological parameters measured using electrodiagnostic studies, and side effects reported.

Data analysis

Data extraction was performed by one investigator (KD) and cross-checked by another (MI). Data were documented on a customized table in order to compare patient demographics, parameters of intervention, duration and outcome measures of each study post-intervention. Where sufficient information was obtainable and the outcome measures were comparable, meta-analyses were performed, allowing a quantitative analysis of the studies. The pooled estimations regarding outcomes were expressed as dichotomous or as continuous variables. These were calculated using a random effect model or a fixed effect model. For dichotomous data, the pooled odds ratio (OR) was calculated. The pooled mean difference was estimated to assess continuous data. Statistical analyses were performed using the Review Manager (RevMan) Version 5.0 software (The Nordice Cochrane Center, The Cochrane Collaboration, Copenhagen,

Denmark, 2008) and STATA Version 13. $P < 0.05$ was considered significant.

Heterogeneity analysis

The existence of statistical heterogeneity between the included studies was assessed using the I^2 test. The heterogeneity was considered low, moderate or high if the I^2 was 25%, 50% or >75%, respectively.[16] If p-value was less than 0.05, the random effect model was adopted or vice versa. The between-trial heterogeneity was assessed using the Q test and the I^2 statistic.[17] Subgroup analyses by month were conducted in order to explore potential sources of the between-trial heterogeneity and potential effect modifiers in this study.

Publication bias assessment

To assess asymmetry, funnel plots were formulated. The Begg's rank correlation and Egger's linear regression tests were used to detect potential publication bias.[18] A two-tailed p-value < 0.10 for Egger regression indicated the presence of publication bias.

Sensitivity analyses

The influence of individual studies, from which the meta-analysis estimates are derived, was examined by omitting low quality studies to see the extent to which inferences depend on a particular study or group of studies (sensitivity analysis).

Qualitative analysis

Based on the methodological quality score of each study, a qualitative analysis was also performed to formulate conclusions thus allowing a wider inclusion of studies. This was done using the Best Evidence Synthesis,[12] which was modified to include observational studies. This method consists of five levels of scientific evidence, presented in Box 1. Consistency was defined *a priori* at 60% (i.e., if 60% or more of studies agreed in the same direction of results).

Box 1. Synthesis of results for the qualitative analysis.

Strong	Consistent findings among two or more, high quality RCTs
Moderate	Consistent findings among one high quality RCT and one or more low quality RCTs and/or CTs or one high quality observational study
Limited	Consistent findings from one high quality RCT, or one low quality RCT or CT, or one high quality observational study
Conflicting	Inconsistent findings among multiple studies (RCTs, CTs and/or observational studies)
No evidence from studies	No studies found

Results

Selection of studies

From the search strategy, 459 potentially relevant studies were identified and 252 duplicate paper were removed after checking titles and abstracts. Out of these studies, 15 fulfilled the inclusion criteria and were eligible for data analysis. Of these, nine were Randomized CTs (RCTs), two were CTs, two were prospective studies and two were retrospective studies. The meta-analysis only included 10[17–26] of these studies with similar outcome measures, whereas all the studies were included in the qualitative analysis. The flow of studies through the selection process is presented in Fig. 1. (see Appendix B for excluded papers). A summary of the studies is presented in Table 1.

Study characteristics

Quality

The methodological quality of the eligible studies was low to moderate. There was a mean CBRG score of 5.1 out of 11 for the 9 RCTs and the 2 CTs. CTs had a lower quality score (2/11) due to lack of randomization, whereas the maximum score in RCTs was 8/11. Due to the nature of the intervention (surgery), neither the patient nor the therapist could be blinded (criteria 4 and 5), so the highest score that could be expected was 9/11. Only 4 out of 11 trials had 80% retention rates for short-term follow-up and 70% for long-tem follow-up and compliance was not clearly stated in the studies with the exception of one.[27] There was a mean score of 13 out of 18 of the criteria proposed by Moga *et al.*[15] for 2 prospective and the 2 retrospective studies (range 11 to 17). Quality scores are presented in Tables 2 and 3.

Participants

The review included 1787 wrists with a clinical diagnosis of CTS. The sample sizes of the 15 eligible

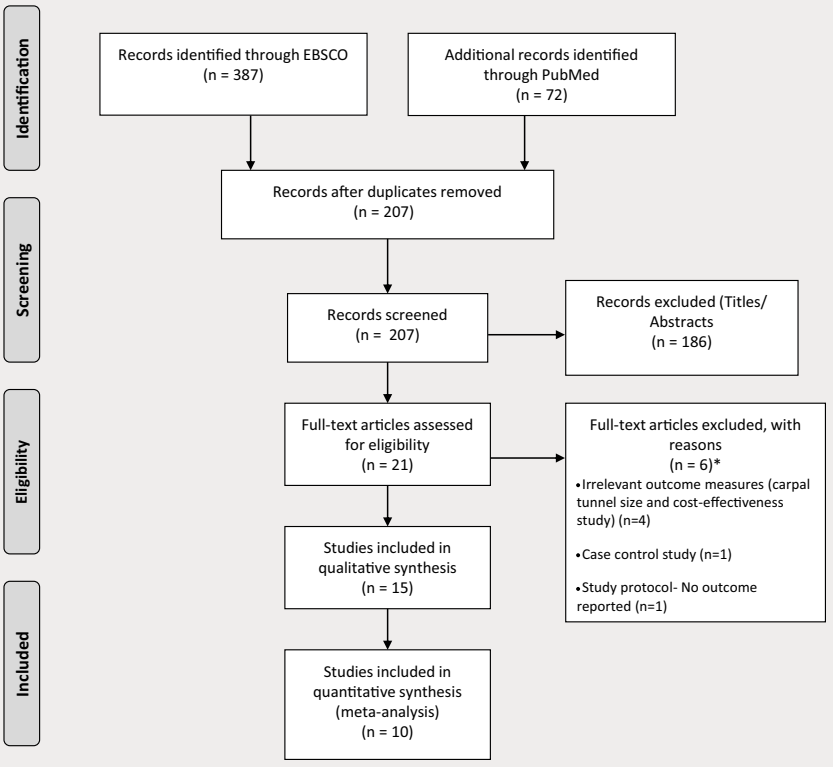

Fig. 1. Selection of studies for inclusion in the systematic review.

*Papers may have been excluded for failing to meet more than one inclusion criteria.

studies ranged from 40 to 429 participants. Participants were mostly females (79%) and one study[28] included only women. The weighted mean age in the studies included was 48.8 years ranging from 20 years old[19] to 88.5 years old[29] although one study concerned an elderly population,[29] affecting the overall mean age. The duration of symptoms ranged from 8 months[20] to 3 years.[27]

Intervention

All eligible studies compared surgery (OCTR or ECTR) to a conservative intervention. Conservative interventions involved the use of steroid injections in six studies,[20–25] splinting in two studies,[22,26] Low Level Laser Therapy (LLLT) in one study,[30] manual physical therapy in one study[28] and multimodality in four studies.[19,27,29,31] The treatment period in the CTs ranged from a single application to three months.

Outcome measures

Symptom severity and functional status were assessed using the BQ in six studies,[22,23,26,28,29] which comprises of the Symptom Severity Scale (11 items) and the Functional Status Scale (8 items). The BQ is a self-administered CTS-specific tool measuring symptom severity and functional status on a scale of 1 to 5, where 1 = no symptoms or no difficulty and 5 = very severe symptoms or so difficult and could not do activity) and the overall score is the mean of all items out of 5. High scores are indicative of more severe symptoms or functional limitation. The BQ's validity and reliability have been previously assessed.[32] One study[27] used the CTS assessment questionnaire (CTSAQ) to assess the symptoms (11 item scale) and functional status (9 item scale) on a 1–5 scale similar to the BQ. Symptom severity was assessed with the Global symptom score (GSS) in two studies.[24,25] This scoring system rates symptoms on a scale of 0 (no symptoms) to 10 (severe) in five categories: pain, numbness, paresthesia, weakness/clumsiness and nocturnal awakening, and the result is the sum of the scores out of 50. Three publications of the same study[20,33,34] assessed the functional status on a 100 mm Visual Analogue Scale (VAS), where 0 = no functional impairment and 100 = the most

Table 1. Characteristics of included studies ($n = 15$).

Study	Design	Participants	Intervention	Outcome measures
Ref. 19	Retrospective	Incl = Clinical diagnosis of CTS confirmed by EDS $n = 265$ patients Sex = 43%F/57%M (114F/151M) Age [μ (range)] = 45 (20–90) yrs; Gp A: 49 yrs; Gp B: 42 yrs	A = OCTR or ECTR (95 wrists/77 patients) B = Multimodality (patient education, wrist splinting, vitamin B, NSAIDs, steroid injections and job change or modification) (188 patients) All groups = educational videotape and brochure, wrist splints, 100 mg Vit B6 and 50 mg B2 daily and NSAIDs.	• Patient satisfaction = Subjective • Repeat history, physical examination and nerve conduction studies = Objective findings • Follow-up = 3 to 9 month intervals (for group B)
Ref. 21	Prospective	Incl = Clinical diagnosis of CTS confirmed by EDS $n = 125$ wrists Sex = Gp A: 76%F/24%M; Gp B: 81%F/19%M; GP C: 79%F/21%M Age [median (range)] = Gp A: 57.5 (30–88) yrs; Gp B: 58.6 (28–87) yrs; Gp C: 57.6 (28–87) yrs Symptom duration (μ) = Gp A: 23.3 mo; Gp B: 22.9 mo; Gp C: 20.9 mo	A = OCTR or ECTR (33 wrists) B = 1–3 doses steroid injections (56 wrists) C = No local treatment. NSAIDs or vascular drugs (36 wrists)	• Improvement of neurophysiological parameters for median nerve: SNCV • Follow-up = Month: 1, 6 and 12
Ref. 31	Prospective	Incl = Clinical diagnosis of CTS > 1 mo $n = 429$ patients Sex = Gp A: 191F/79M; Gp B: 102F/23M Age ($\mu \pm$ SD) = (3 subgroups) ($\mu \pm$ SD): A1: 68.0 ± 9.1 yrs; A2: 42.0 ± 7.3 yrs; A3: 39.0 ± 8.1 yrs; B1: 64.0 ± 7.0 yrs; B2: 41.0 ± 8.9 yrs; B3: 37.0 ± 8.8 yrs	A1, A2 and A3 = OCTR or ECTR (270 patients) B1, B2 and B3 = Multimodality (125 patients) [NSAIDs (96), wrist splints (115), physiotherapy (42), work modification (70), steroid injections (48), Vitamin B6 (13)]	• Symptoms = SSS — 11 items (5-point scale) • Function = FSS — 8 items (5-point scale) • Satisfaction = satisfaction scale 7 items (5-point scale) • Health status = SF-36: 36 questions assessing general health-related quality of life • Follow-up = Month: 6, 18 and 30
Ref. 23	CT	Incl = Clinical diagnosis of unilateral CTS > 6 mo, confirmed by EDS $n = 90$ patients Sex = Gp A: 38F/6M; Gp B: 42F/4M Age (μ ± SD) = Gp A: 48.0 ± 8.4 yrs; Gp B: 45.3 ± 9.9 yrs	A = OCTR (44 patients) B = 2 doses steroid injections 2 wks apart of 6.4 mg betamethasone (46 patients) 23-gauge needle at anterior wrist flexion crease, angulation 45° distally and 45° radially	• Symptoms = BQ — 11 items (5-point scale) • Function = BQ — 8 items (5-point scale) • Improvement of neurophysiological parameters for median and ulnar nerves = DML, MNCV, CMAP, DSL, SNCV, SNAP using the Nihon Kohden-Neuropack MEB 5504 K • Follow-up = Month: 3 and 6

Table 1. (*Continued*)

Study	Design	Participants	Intervention	Outcome measures
Ref. 26	RCT	Incl = Pain or paraisthesias in the median nerve distribution, clinical diagnosis of CTS, confirmed by EDS, age > 18 yrs, ability to fill out questionnaires in Dutch $n = 176$ Sex = Gp A: 66F/11M; Gp B: 77F/2M Age ($\mu \pm$ SD) = Gp A: 49 ± 11; Gp B: 49 ± 12 Symptom Duration [median (IQR)] = Gp A: 40 (16–104) wks; Gp B: 52 (24–104) wks Characteristics = small heterogeneity in sex and symptom duration	A = OCTR (87) B = Splinting in neutral position of wrist for 6 wks min at night and during the day as preferred (89)	• General improvement = 6-point ordinal transition scale (success = "completely recovered" or "much improved") • No. of nights the patient woke due to symptoms in the last week • Symptoms = 11-point scale • Symptoms = SSS — 11 items (5-point scale) • Function = FSS — 8 items (5-point scale) • Severity of CTS-related complaints-Physiotherapist assessment on 11-point scale • Improvement of neurophysiological parameters for median nerve = DML, DSL, Median-Ulnar DSL difference • Follow-up = Month: 3, 6 and 12
Ref. 24	RCT	Incl = Clinical diagnosis of CTS > 3 mo but < 1 yr confirmed by EDS $n = 50$ patients Sex = 48F/2M; Gp A: 24F/1M; Gp B: 24F/1M Age ($\mu \pm$ SD) = Gp A: 50.8 ± 11.6 yrs; Gp B: 48.2 ± 6.5 yrs Characteristics = no large heterogeneity in age, symptom severity, baseline measurements	A = OCTR (25 patients) B = 1 dose steroid injection 15 mg methylprednisolone acetate with 25-gauge needle 30° angle medial to Palmaris longus tendon (25 patients)	• Symptoms = GSS — 11-point scale for five symptom categories, total score/50 • Improvement of neurophysiological parameters for median nerve = DML, SNCV • Grip strength = JAMAR hydraulic hand dynamometer • Follow-up = week: 6 and 20
Ref. 20	RCT	Incl = Age ≥ 18 yrs, Clinical diagnosis of CTS > 3 mo, confirmed by EDS, consecutive referral and unresponsiveness to ≥ 2 wks NSAIDs treatment $n = 163$ wrists/101 patients Sex: 93F/8M Age ($\mu \pm$ SD) = Gp A: 50.53 ± 10.87 yrs; Gp B: 53.17 ± 13.93 yrs ($p = 0.213$) Symptom duration ($\mu \pm SD$): Gp A: 31.12 ± 7.27 wks; Gp B: 33.25 ± 8.17 wks ($p = 0.723$)	A = OCTR (80 wrists) B = 1 or 2 doses steroid injection 20 mg in 1 mL paramethasone acetonide with 22-gauge needle 45° angle distally 1–2 cm depth medial to Palmaris longus tendon (83 wrists)	• Nocturnal paresthesias = 100 mm VAS • Diurnal pain = 100 mm VAS • Functional impairment = 100 mm VAS • Follow-up = month: 3, 6 and 12

Table 1. (*Continued*)

Study	Design	Participants	Intervention	Outcome measures
Ref. 29	Retrospective	Incl = Age $\geq$ 70 yrs, Recent clinical diagnosis of CTS, with or without normal EDS, county residents n = 96 wrists/60 patients Sex = 43F/27M Age [μ (range)]: 77.0 (70.2–88.5) yrs Characteristics = no large heterogeneity in age, symptom duration or symptom severity	A = OCTR or ECTR (47 wrists) B = Multimodality (NSAIDs, splinting, steroid injections) (41 wrists)	• Symptoms = BQ — 11 items (5-point scale) • Function = BQ — 8 items (5-point scale) • Patient expectations and satisfaction = MODEMSQ — Musculoskeletal Outcome Data Evaluation and Management System Questionnaire — 5 items • Health status = SF-36: 36 questions assessing general health-related quality of life • Follow-up = N/A
Ref. 22	RCT	Incl = Clinical diagnosis of mild to moderate CTS $>$ 6 mo confirmed by EDS n = 57 wrists Sex = 53F/4M Age ($\mu \pm$ SD) = Gp A: 45.27 $\pm$ 13.19 yrs; Gp B: 44.50 $\pm$ 7.24 yrs; Gp C: 44.46 $\pm$ 8.52 yrs ($p = 0.976$) Symptom duration ($\mu \pm$ SD): Gp A: 21 $\pm$ 11 mo; Gp B: 15.26 $\pm$ 7.19 mo; Gp C: 19.13 $\pm$ 13 mo ($p = 0.869$)	A = OCTR (11) B = Splinting (23) in neutral position of the wrist for 3 mo night and day C = Splinting & 1 dose steroid injection (23)	• Improvement of neurophysiological parameters for median nerve = DML, PML, MNCV, CMAP wrist and elbow, SNCV, SNAP using the Nihon-Cohden Neuropack • Symptoms = BQ — 11 items (5-point scale) • Function = BQ — 8 items (5-point scale) • Satisfaction = 5-point scale • Follow-up = month: 3 and 6
Ref. 30	CT	Incl = Clinical diagnosis of CTS n = 60 wrists/54 patients Sex = Gp A: 23F/4M; Gp B: 25F/2M Age ($\mu \pm$ SD) = Gp A: 42.65 $\pm$ 8.0 yrs; Gp B: 49.11 $\pm$ 7.23 yrs Symptom duration ($\mu \pm$ SD) = Gp A: 36.17 $\pm$ 4.38 mo; Gp B: 28.21 $\pm$ 7.03 mo	A = OCTR (30 wrists) B = LLLT Neon (He–Ne) (632.8 nm, Level Laser M300) continuous wave (CW), $\geq$ 12 mW, 30 cm from skin, 3 J/cm^2, 2x/wk for 12 sessions (30 wrists)	• % patients with • Symptom relief • Return to normal activities • Adverse effects from treatment • Positive nerve conduction tests • Follow-up = month: approximately 6

Table 1. (*Continued*)

Study	Design	Participants	Intervention	Outcome measures
Ref. 27	RCT	Incl = Age $\geq$ 18 yrs, Clinical diagnosis of CTS $>$ 2 wks confirmed by EDS, +ve "flick test" or nocturnal pain, failure of 2 wks conservative treatment n = 116 patients Sex = Gp A: 28F/29M; Gp B: 34F/25M Age ($\mu \pm$ SD) = Gp A: 50.2 $\pm$ 10.3 yrs; Gp B: 51.2 $\pm$ 8.9 yrs ($p = 0.213$) Symptom duration [median (IQR)] = Gp A: 3.2 (1.3–5.5) yrs; Gp B: 3.4 (1.0–8.7) yrs	A = OCTR or ECTR (57 patients) B = Multimodality [200 mg ibuprofen 3x/day, 6 sessions hand therapy over 6 wks, educational booklet, hand exercises, splinting night and day, work modifications] (59 patients)	• Function = CTSAQ — 9 items (5-point scale) • Symptoms = CTSAQ — 11-items (5-point scale) • Pain interference with work or activities = 11-point scale • Health status = SF-36: 36 questions assessing general health-related quality of life • Additional treatments = Patient diary • Follow-up = month: 6 and 12
Ref. 33	RCT	Incl = Age $\geq$ 18 yrs, Clinical diagnosis of CTS $>$ 3 mo, confirmed by EDS, consecutive referral and unresponsiveness to $\geq$ 2 wks NSAIDs treatment n = 163 wrists/101 patients Sex: 93F/8M Age ($\mu \pm$ SD) = Gp A: 50.53 $\pm$ 10.87 yrs; Gp B: 53.17 $\pm$ 13.93 yrs ($p = 0.213$) Symptom duration ($\mu \pm$ SD): Gp A: 31.12 $\pm$ 7.27 wks; Gp B: 33.25 $\pm$ 8.17 wks ($p = 0.723$)	A = OCTR (80 wrists) B = 1 or 2 doses steroid injection 20 mg in 1 mL paramethasone acetonide with 22-gauge needle 45° angle distally 1–2 cm depth medial to Palmaris longus tendon (83 wrists)	• Nocturnal paresthesias = 100 mm VAS • Diurnal pain = 100 mm VAS • Functional impairment = 100 mm VAS • Follow-up = month: 3, 6, 12 and 24
Ref. 34	RCT	Incl = Age $\geq$ 18 yrs, Clinical diagnosis of CTS $>$ 3 mo, confirmed by EDS, consecutive referral and unresponsiveness to $\geq$ 2 wks NSAIDs treatment n = 163 wrists/101 patients Sex: 93F/8M Age ($\mu \pm$ SD) = Gp A: 50.53 $\pm$ 10.87 yrs; Gp B: 53.17 $\pm$ 13.93 yrs ($p = 0.213$) Symptom duration ($\mu \pm$ SD): Gp A: 31.12 $\pm$ 7.27 wks; Gp B: 33.25 $\pm$ 8.17 wks ($p = 0.723$)	A = OCTR (80 wrists) B = 1 or 2 doses steroid injection 20 mg in 1 mL paramethasone acetonide with 22-gauge needle 45° angle distally 1–2 cm depth medial to Palmaris longus tendon (83 wrists)	• Nocturnal paresthesias = 100 mm VAS • Diurnal pain = 100 mm VAS • Functional impairment = 100 mm VAS • Follow-up = month: 3, 6 and 12 • Improvement of neurophysiological parameters for median nerve = DML, MA, SNCV, SA • Follow-up = month: 12

Table 1. (*Continued*)

Study	Design	Participants	Intervention	Outcome measures
Ref. 25	RCT	Incl = Clinical diagnosis of CTS confirmed by EDS, symptoms for > 3 mo $n = 40$ patients Sex = 29F/11M ($p = 0.723$) Age ($\mu \pm$ SD) = Gp A: 43.8 ± 10.98 yrs; Gp B: 46.9 ± 12.33 yrs ($p = 0.406$) Symptom duration ($\mu \pm$ SD) = Gp A: 12.5 ± 8.76 mo; Gp B: 10.15 ± 6.75 mo ($p = 0.348$)	A = OCTR (20 patients) B = Steroid injection with 40 mg of methylprednisolone (20 patients)	• Symptoms = GSS — 11-point scale for five symptom categories, total score/50 • Follow-up = week: 2, 4 and 12
Ref. 28	RT	Incl = Clinical diagnosis of CTS confirmed by EDS (sensory and motor deficit), symptoms for > 12 mo $n = 120$ patients Sex = 120F/0M Age ($\mu \pm$ SD) = Gp A: 46 ± 9 yrs; Gp B: 47 ± 10 yrs Symptom duration ($\mu \pm$ SD) = Gp A: 3.5 ± 3.1 yrs; Gp B: 3.1 ± 2.7 yrs	A = OCTR or ECTR (60 patients) B = Manual therapy 3 sessions of 30' 1x/wk (60 patients)	• Pain intensity = NPRS — 0–10 (11-point scale where 0 = no pain and 10 = worst possible pain) • Symptoms = BQ — 11 items (5-point scale) • Function = BQ — 8 items (5-point scale) • Self-Perceived improvement = GROC (from −7 (worse) to +7 (better)) • Follow-up = Month: 1, 3, 6 and 12

Notes: Incl = inclusion criteria, n = number of patients randomized, Gp = group, EDS = electrodiagnostic studies, OCTR = Open carpal tunnel release, ECTR = Endoscopic carpal tunnel release, BQ = Boston Questionnaire, SF-36 = Short Form 36, DML = Distal Motor Latency, MNCV = Motor Nerve Conduction Velocity, CMAP = Compound Muscle Action Potential, DSL = Distal Sensory Latency, SNCV = Sensory Nerve Conduction Velocity, SNAP = Sensory Nerve Action Potential, SSS = Symptom severity scale, FSS = Functional status scale, GSS = Global symptom score, VAS = Visual analogue scale, MODEMSQ = Musculoskeletal Outcome Data Evaluation and Management System Questionnaire, PML = Proximal motor latency, CTSAQ = Carpal tunnel syndrome assessment questionnaire, MA = Motor amplitude, SA = Sensory amplitude, RCT = Randomized controlled trial, CT = Clinical trial, RT = Randomized trial, NPRS = Numerical Pain Rating Scale, GROC = Global rating of change.

Table 2. Quality scores for CTs ($n = 11$).

Study	Randomization adequacy	Allocation concealment	Baseline comparability	Participant blinding	Therapist blinding	Assessor blinding	Cointervention avoidance	Compliance	Dropout rate < 20%	Timing	Intention to treat analysis	Total (0 to 11)
Ref. 23	N	N	Y	N	N	N	U	U	N	Y	N	2
Ref. 26	Y	Y	Y	N	N	Y	Y	N	N	Y	Y	7
Ref. 24	Y	Y	Y	N	N	Y	Y	U	Y	Y	N	7
Ref. 20	Y	Y	Y	N	N	N	U	U	Y	Y	Y	6
Ref. 22	Y	U	Y	N	N	N	U	U	U	Y	N	3
Ref. 30	N	N	Y	N	N	N	N	U	U	Y	N	2
Ref. 27	Y	Y	Y	N	N	Y	U	Y	N	Y	Y	7
Ref. 33	Y	Y	Y	N	N	N	U	U	N	Y	Y	5
Ref. 34	Y	Y	Y	N	N	N	U	U	N	Y	N	4
Ref. 25	Y	Y	Y	N	N	U	N	U	Y	Y	N	5
Ref. 28	Y	Y	Y	N	N	Y	Y	U	Y	Y	Y	8

Table 3. Quality scores for prospective and retrospective studies ($n = 4$).

Study	Hypothesis stated	Description of characteristics	Multicenter study	Appropriate eligibility criteria	Consecutive recruitment	Similar stage of condition	Clearly described intervention	Clearly described cointerventions	Outcome measure description	Outcome measure suitability	Outcome measure timing	Statistical test suitability	Length of follow-up	Loss to follow-up	Random variability	Adverse events	Conclusion supported by result	Competing interest	Total (0 to 18)
Ref. 19	Y	Y	N	Y	N	N	Y	Y	Y	N	Y	N	Y	N	N	Y	Y	Y	11
Ref. 21	Y	Y	N	Y	N	Y	N	Y	Y	Y	Y	Y	Y	N	N	N	Y	N	11
Ref. 31	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	17
Ref. 29	Y	Y	N	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	N	13

severe functional impairment. For meta-analysis purposes, the GSS scores and the VAS scale scores were divided by 10 and 20, respectively, coming up with a common denominator of 5, in order to be coherent with the BQ and CTSAQ scores. The improvement of neurophysiological parameters was carried out by physicians and neurologists with nerve conduction studies. The most commonly assessed parameters were the Distal Motor Latency (DML) and the Sensory Nerve Conduction Velocity (SNCV).

Effect of intervention on symptom improvement

The meta-analysis pooled data from 6 studies with a total of 805 wrists. One study[26] was not included in the pooled analysis as it presented the standardized mean difference in contrast to the other studies, which presented the mean score and standard deviation (SD). The results demonstrate that surgical treatment leads to a greater symptom improvement at six months by 0.52 points lower score on a 5-point symptom severity scale compared to conservative treatment 95%CI (0.27 to 0.78). There was a statistically significant high heterogeneity ($I^2 = 82\%$, $p < 0.0001$) (Fig. 2). The results remain similar, if only high quality studies[17,18,22,26] were included in the meta-analysis (MD 0.56, 95%CI 0.16 to 0.96, $I^2 = 89\%$, $p < 0.00001$) (see Fig. C.1 in Appendix C). Meta-analysis results at three months further increased the total heterogeneity, therefore it was not astute to present them. Results for 18 or > 18 months could not be calculated due to insufficient data from the included studies.

The qualitative analysis agreed with the meta-analysis, with strong evidence in favor of surgery for symptom improvement at 6 months, conflicting evidence at 12 months, moderate evidence at 18 months and limited evidence for the time period longer than 18 months (Table 4, see also Tables C.1–C.5 in Appendix C for details of included studies).

Effect of intervention on functional improvement

The total meta-analysis results pooled data from 6 studies with a total of 918 wrists. One study[26] was not included in the pooled analysis for reasons explained previously. The results demonstrate that surgery was superior to conservative treatment for functional improvement but the result was not significant (MD 0.06, 95%CI −0.10 to 0.22, $I^2 = 84\%$, $p < 0.00001$). Due to the high heterogeneity, a subgroup analysis was fitting. At 3 months and 12 months, there was no statistically

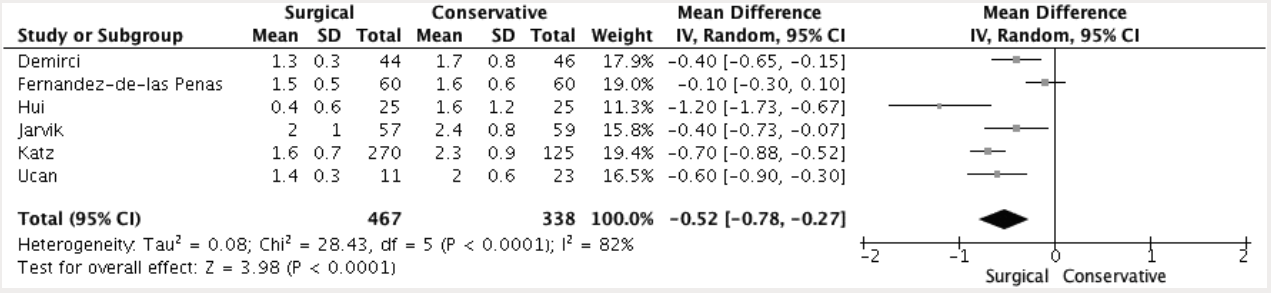

Fig. 2. MD (95%CI) of effect of surgical and conservative treatment on symptom improvement at six months of treatment by pooling data from six studies ($n = 805$).

Table 4. Qualitative analysis for symptom improvement.

	3 months	6 months	12 months	18 months	> 18 months
Symptom improvement	Conflicting evidence (6 studies)	Strong evidence fav. Surg (8 studies)	Conflicting evidence (4 studies)	Moderate evidence fav. Surg (2 studies)	Limited evidence fav. Surg (2 studies)

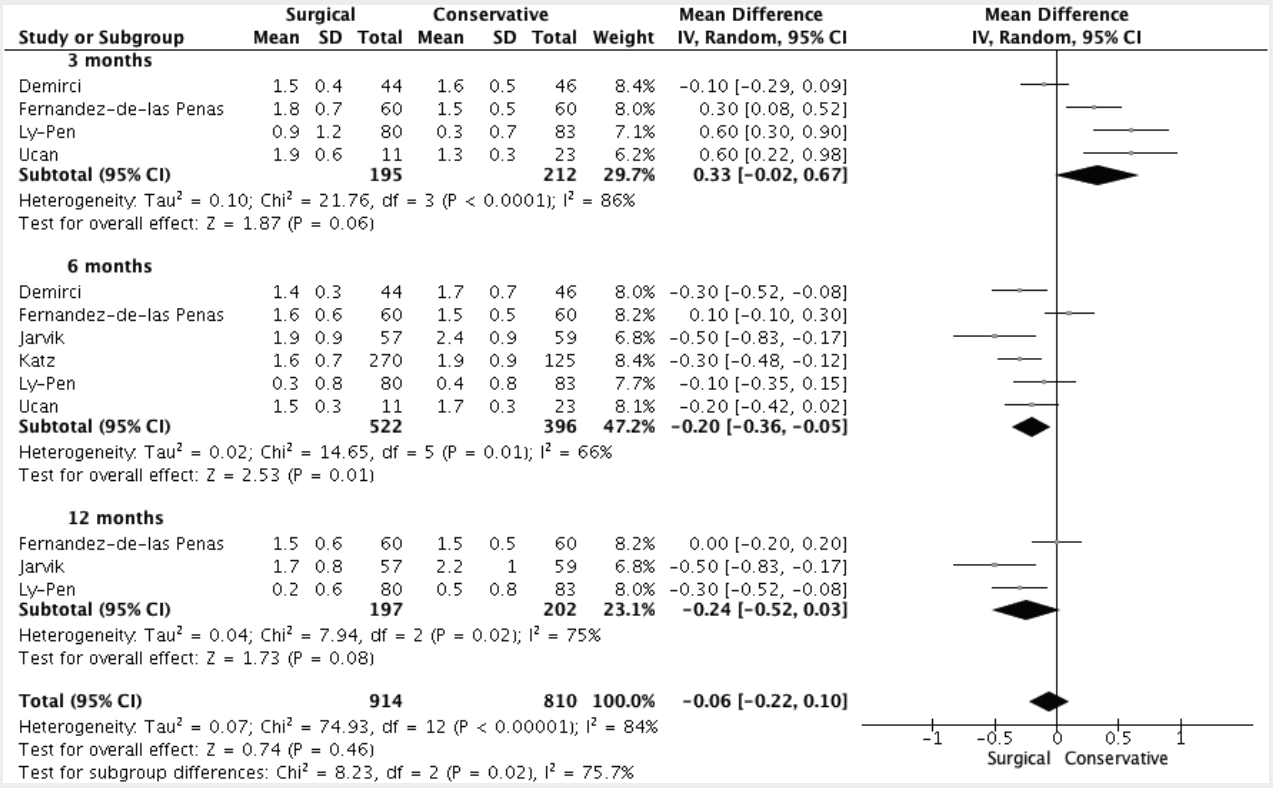

Fig. 3. MD (95%CI) of effect of surgical and conservative treatment on functional improvement at 3, 6 and 12 months of treatment by pooling data from six studies ($n = 918$).

significant difference between the surgical and conservative treatment for function. Surgery proved more effective than conservative treatment at six months (MD 0.20, 95%CI 0.05 to 0.36, $I^2 = 66\%$, $p = 0.01$) (Fig. 3, but if only high quality studies[17,18,26] were included, the results were not statistically significant at six months (MD −0.22, 95%CI −0.55 to 0.12, $I^2 = 85\%$, $p = 0.001$). The results remain similar at 12 months (MD −0.23, 95%CI −0.72 to 0.26, $I^2 = 85\%$, $p = 0.01$) with data from two studies[17,18] (see Fig. C.2 in Appendix C).

The qualitative analysis favored surgery with strong evidence at 6 and 12 months and limited evidence for the time period longer than 18 months (Table 5, see also Tables C.6–C.10 in Appendix C for details of included studies).

Effect of intervention on improvement of neurophysiological parameters

The included studies presented results for different neurophysiological parameters. The meta-analysis was only applicable for the DML and SNCV. Studies assessed the outcomes for the DML between 5 and 12 months. One study[24] performed the follow-up measurement at 5 months, another[34] at 12 months and two more studies performed the

Table 5. Qualitative analysis for functional improvement.

	3 months	6 months	12 months	18 months	> 18 months
Functional improvement	Moderate evidence fav. Conserv. (5 studies)	Strong evidence fav. Surg (8 studies)	Strong evidence fav. Surg. (4 studies)	Conflicting evidence (2 studies)	Limited evidence fav. Surg (2 studies)

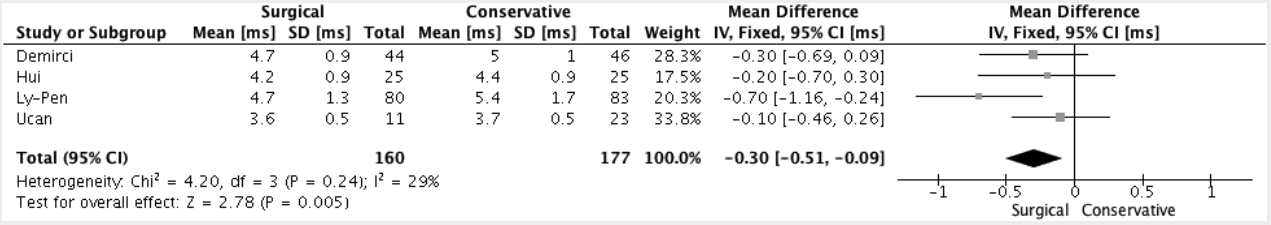

Fig. 4. MD (95%CI) of effect of surgical and conservative treatment on improvement of DML between 5 and 12 months of treatment by pooling data from four studies ($n = 337$).

follow-up at 6 months.[22,23] Pooled data utilizing 337 wrists demonstrated that surgery provides a greater improvement of the DML compared to conservative treatment with a mean difference 0.30 ms less time delay in the surgical group compared to the conservative (95%CI 0.09 to 0.51) and low to moderate non-significant heterogeneity ($I^2 = 29\%$, $p = 0.24$) (Fig. 4).

In the total meta-analysis results for the SNCV were in favor of surgery (MD 2.73 m/s, 95%CI 0.71 to 4.75) with a total of 80 wrists for the surgical and 94 for the conservative group. There was a moderate non-significant heterogeneity of 52% ($p = 0.08$), which was affected by the results at 3 months, so a subgroup analysis was appropriate (Fig. 5). At 6 months, the results were in favor of surgery (MD 3.71 m/s, 95%CI 1.94 to 5.49, $I^2 = 0\%$, $p = 0.42$).

In contrast to the meta-analysis, the qualitative analysis took into consideration the overall effectiveness from the nerve conduction studies reported in the included studies. The data synthesis demonstrated that there was a benefit in favor of surgery with moderate evidence at 6 and 12 months and insufficient data to formulate conclusions for the long-term effectiveness (Table 6, see

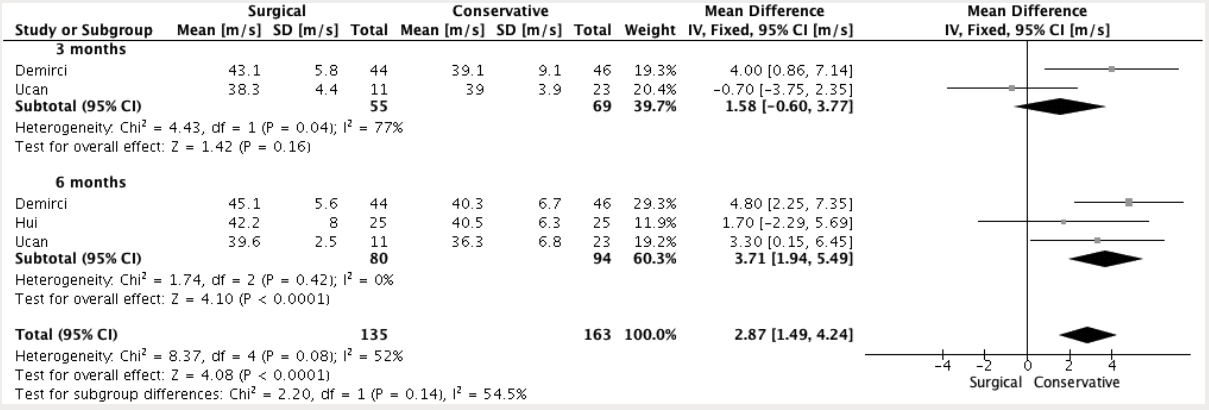

Fig. 5. MD (95%CI) of effect of surgical and conservative treatment on improvement of SNCV at six months of treatment by pooling data from three studies ($n = 174$).

*Left = results in favor of surgery, Right = results in favor of conservative treatment.

Table 6. Qualitative analysis for improvement of neurophysiological parameters.

	3 months	6 months	12 months	18 months	> 18 months
Improvement of neurophysiological parameters	Conflicting evidence (2 studies)	Moderate evidence fav. Surg (3 studies) Limited evidence fav Surg. for DML (1 study)	Moderate evidence fav. Surg (3 studies)	N/A	N/A

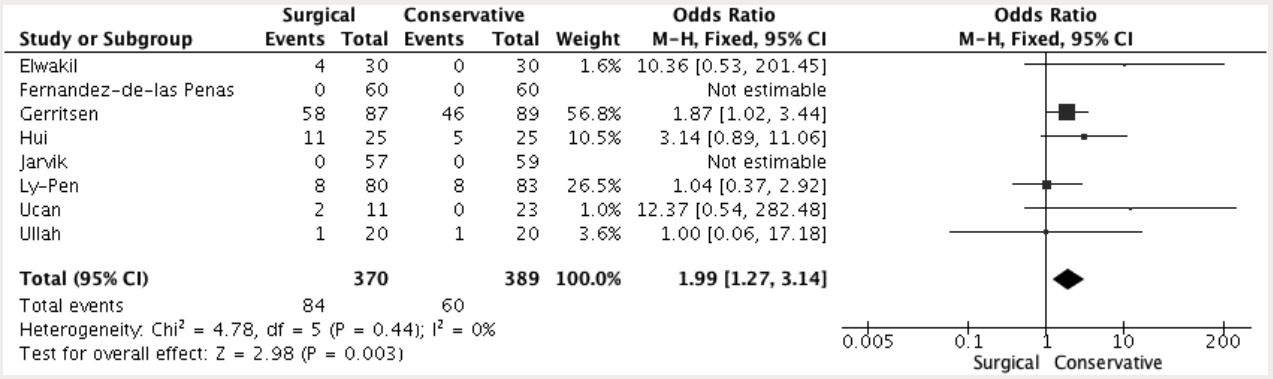

Fig. 6. OR (95%CI) of effect of surgical and conservative treatment on complications reported by pooling data from eight studies ($n = 759$).

also Tables C.11–C.13 in Appendix C for details of included studies).

Side effect and complication

Eight of the trials included reported side effects from treatment. However, there was a wide range of side effects and complications reported. Serious complications were only reported in the surgery group and included reflex sympathetic dystrophy[25,26] and complex regional pain syndrome.[22] Mild side effects from both treatments included symptoms like pain, swelling, discomfort and were reported in most CTs.[20,22,24,26,30,33,34] The results from 370 wrists, which underwent surgery and 389 wrists treated conservatively, showed that conservative treatment was more beneficial than surgery with almost half the complications reported with conservative compared to surgical treatment (OR = 1.99, 95%CI 1.27 to 3.14, $I^2 = 0\%$, $p = 0.44$) (Fig. 6). The results remain similar, if only high quality studies[17,18,22,24] were included in the meta-analysis (OR 2.07, 95%CI 1.20 to 3.57, $I\% = 0\%$, $p = 0.47$) (see Fig. C.3 in Appendix C).

Publication bias

There was symmetry in the funnel plots about the standard error. Therefore, no publication bias was identified because the funnel plots for symptom improvement, functional improvement or improvement of neurophysiological parameters, were symmetric. No publication bias was noted on funnel plots and Egger regression ($p > 0.05$).

Sensitivity analyses

Sensitivity analyses showed that the pooled estimates of outcome measures did not vary substantially with the exclusion of low quality study.

Qualitative analysis and meta-analysis agreement

A comparison of a qualitative analysis and a quantitative analysis (meta-analysis) was only possible where sufficient data were available for a meta-analysis. It is worth noting that there was a considerable agreement between the qualitative and the quantitative analysis, despite the inclusion of different studies in each type of analysis. In particular, the outcomes with conflicting results in the qualitative analysis showed non-significant results in the meta-analysis at the same intervals of reassessment. Furthermore, the outcomes demonstrating statistically significant results in the meta-analysis had strong evidence in favor of a treatment.

Discussion

Even with the inclusion of more recent studies, the results of this systematic review are consistent with the previous systematic reviews with regards to the direction of results. Surgical treatment outweighed conservative treatment in all outcomes. Conservative treatment however caused fewer complications than surgery. Both treatments were effective in improving symptoms and function at six months.

There was some concordance in the results of this systematic review compared to the latest one,[10] with statistically significant results at six months for symptom improvement. There was a disagreement regarding the results at 12 months, with the current meta-analysis showing no significant differences in symptoms and function, in contrast to the previous systematic review, which clearly supported surgery. These differences are attributed to the inclusion of both additional and different studies in the meta-analysis. One trial[26] was excluded for reasons previously mentioned, and the most recent study[28] showed no difference between the two interventions at 12 months.

The exclusion of low quality trials, for meta-analysis purposes, was only possible for 6 months symptom improvement, 6 and 12 months functional improvement and complication and side effects. The results were only differentiated in functional improvement but this increased the heterogeneity by almost 20%.

Since the heterogeneity of the studies included in the meta-analysis was high, a qualitative analysis was carried out in addition to the meta-analysis. The qualitative analysis allowed for a wider inclusion of studies, upon which to draw conclusions. In addition, it provided a classification of studies according to their quality, with the highest quality studies having a greater effect on the overall outcome. The qualitative analysis took into account the result of each study for each outcome, the methodological quality of each study and the number of studies in favor of an intervention for each outcome.

In the results for symptom improvement, it was evident that surgery was superior to conservative treatment at six months. Regarding functional status at six months, the surgery group had greater functional improvement, which spread to the 12-month re-evaluation (only in the qualitative analysis). However, the results for symptoms and function at three months did not favor one of the interventions as the qualitative analysis showed conflicting evidence and the meta-analysis showed a trend towards surgery, which was not statistically significant.

The reason for this discrepancy between the short- and long-term efficacy for symptoms and function is the use of steroid injections as conservative treatment intervention in the above studies. Steroid injections can reduce wrist joint effusion and vascular congestion[35] around a median nerve which often appears swollen when examined with ultrasound.[36] Injections can provide short-term analgesia but the mechanical compression persists,[24] resulting in a gradual recurrence of symptoms.

It was previously reported that 50% of the cases treated with steroid injections are worse regarding their clinical presentation and neurophysiological studies over a period of six months.[37] This is reflected in the studies using steroid injections used in this review.[20,22,25,33,34] Ly-Pen *et al.*[33] revealed that as the time passes from the injection until the re-evaluation, the failures of the intervention increase. They attributed the need for review of the original study to this event, and monitored the patients for another year after their initial intervention, where the results were clearly in favor of surgery. On the contrary, surgery provides a more permanent solution, as the resection of the flexor retinaculum decompresses the median nerve. This fact, combined with the complications of surgery, the postoperative discomfort and the patient's reluctance to move after surgery to avoid irritation of the wound, results in a conservative treatment with injections appearing more effective than surgery in the short term.

An additional reason for this discrepancy was the inclusion of the most recent study,[28] which affected the overall result at 12 months as it showed no significant differences at 6 and 12 months. The unique feature of this study, which had the highest methodological score (8/11), was the nature of the intervention, which was not concentrated on the wrist, unlike the other conservative interventions. This manual therapy intervention included desensitization maneuvers, across the continuum of the median nerve from the cervical spine to the wrist. Specifically, treatment targeted all possible locations of entrapment of the median nerve along its path, prior to the application of gliding movements of the nerve to improve its glide in relation to the adjacent tissues. This study only included female participants but it is unlikely that the results will be different if participants of both genders were included.

The improvement of neurophysiological parameters was also superior with surgical treatment. The meta-analysis data for DML were extracted from studies with different time intervals, but this was the only way to group the studies together for the meta-analysis.

Complications and side effects were the only occasion where conservative treatment outweighed surgery with almost twice as many complications in the surgery group compared to the group of conservative treatment. Severe complications occurred only with surgery. Most reported side effects were related to pain or tenderness at the incision site, which are common signs and symptoms after open surgery. In addition, there was great heterogeneity in the reported side effects. Only a few studies reported serious side effects, while others reported all adverse reactions regardless of severity. Because there was great variation in the severity of complications reported in each study, it is not possible to verify the real advantage of conservative versus surgical treatment in this outcome.

Additional heterogeneity existed, regarding the chronicity and severity of symptoms, the period of re-evaluation, the outcome measures of each study, and how these were measured. Analysis into subgroups according to the severity of symptoms or the time of reassessment was not possible due to the limited number of studies classifying the patients accordingly. Perhaps, this should be addressed in future trials comparing the two interventions.

The severity of the symptoms could not be accurately determined as there was great variation in the assessment methods. Moreover, some studies[29–31] did not confirm the diagnosis through electrodiagnostic studies, and as a result, some of the patients included may not have been suffering from CTS alone. The need for neurophysiological studies however is not universally accepted. Some studies, indicate that the nerve conduction studies are not necessary for the diagnosis of CTS, since it can lead to false negative or false positive results.[38–40] Other studies indicate that there is no correlation between clinical symptoms and results of neurophysiological studies in CTS. Researchers propose the use of neurophysiological studies as an additional independent tool for diagnosis and assessment of the severity of CTS.[41]

Regarding the severity of the condition, most studies excluded patients with severe atrophy of the thenar muscles, or previous surgery for CTS excluding in this way severe cases of CTS. Six RCTs and CTs reported that they included people with mild to moderate CTS.[20,23,24,26,33,34] One study included patients with mild, moderate or severe CTS[28] and the remaining eight studies did not make such a reference. Instead, some presented baseline measurements for the severity of symptoms and functional status. Comparing these baseline values from the BQ showed that the severity of symptoms and functional status were comparable between studies, where the severity of CTS is mentioned (see Tables C.14 and C.15 in Appendix C). Therefore, we can deduce that this systematic review refers mainly to people with mild to moderate CTS.

An additional issue of concern was the fact that the search strategy for this review was limited to electronic databases and the gray literature (unpublished studies) was not searched. This could have affected the results of the publication analysis.

From the studies included, many were limited by lack of randomization, lack of standardized outcome measures and retrospective design, which lacked information on patient baseline measurements, so a comparison of the severity of initial symptoms was not possible, nor was an estimate of the improvement from baseline until reassessment. The incorporation of prospective and retrospective studies however allowed for an evaluation of the outcomes in the longer term.

In the absence of randomization, these observational studies are considered lower quality than RCTs and CTs for the collection of data on the efficacy of an intervention because they can be affected by various types of bias such as selection, detection, performance, attrition, reporting and publication.[15] However, their inclusion constituted a strength of this review. In these studies, treatment was pre-determined according to the severity of the condition. Patients with severe CTS underwent surgery. Since the surgical intervention was more effective than conservative, and in more severely impaired patients, one can assume that the effect size of surgery might be higher than the one of the conservative. On the other hand, a long-term observational study[42] presented evidence that in some cases, CTS may improve spontaneously, causing these patients to undergo unnecessary surgery. In addition, severely impaired patients might find it easier to score positively on subjective outcome measures like VAS and GSS. These remarks make the formulation of a clear conclusion difficult and future studies should address them.

An evaluation of 331 hands identified 5 factors of poor prognosis and the need for surgery: age above 50 years, symptom for over 10 months,

consecutive symptoms of paraesthesia, flexor tenosynovitis and a positive Phalen's test for less than 30 s.[43] It was reported that when these factors did not exist, 2/3 of patients healed with conservative treatment. When 4 or 5 of these factors are present, they recommend surgical treatment. Conversely, other researchers studied 45 hands conservatively with steroid injections and concluded that there is no correlation between signs and symptoms of CTS and the final result.[44] They argued that chronicity is the most significant factor for the final result.[44] Therefore, conservative treatment is a feasible option for mild and short-term symptoms, but surgery can provide a more permanent solution to persisting symptoms.

Conclusion

The results of this review demonstrate that surgery leads to a greater improvement of symptoms and neurophysiological parameters at six months, compared to conservative treatment. The decision however, about the choice of treatment, needs careful consideration, taking into account the complications reported with surgical treatment and the fact that in some cases, CTS may be resolved spontaneously. However, the conclusions derived from this review are based on a number of underpowered studies. Therefore, high quality prospective studies are needed in order to identify the characteristics of individuals where CTS has promising path to avoid unnecessary surgery. In addition, further research should focus on exploring the field of manual therapy and compare it to the surgical intervention for CTS. Research should also address the long-term effectiveness of the two interventions beyond 12 months.

Conflict of Interest

The authors declare that they have no conflict of interest.

Funding/Support

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Author Contributions

DK was responsible for the data collection, data analysis, data interpretation and writing the manuscript, IM was responsible for revising the manuscript, statistical analysis and project management. Both authors participated in editing and completion of the manuscript.

Appendix A.

Summary of search strategy.

#	Keywords	PubMed (Title/ Abstract)	EBSCO (Abstract)
#1	((surgical*) OR (surgical intervention) OR (open carpal tunnel release) OR (OCTR) OR (endoscopic carpal tunnel release) OR (ECTR))	822.822	1.713.262
#2	((conservative) OR (conservative intervention) OR (corticosteroid injections) OR (steroid injections) OR (wrist splints) OR (physiotherapy) OR (electrotherapy) OR (exercise) OR (manual therapy))	322.464	1.579.763
#3	((Carpal tunnel syndrome) OR (CTS) OR (median nerve entrapment) OR (nerve compression))	13.486	61.818
#4	((RCT) OR (random*) OR (randomized controlled trial) OR (controlled trial) OR (cohort study) OR (clinical trial) OR (controlled clinical trial) OR (retrospective) OR (prospective))	1.595.404	6.814.297
#5	**#1 AND #2 AND #3 AND #4**	**72**	**387**

Appendix B.

List of excluded papers

Study	Year	Title	Reason for exclusion
Martin *et al.*	2005	RCT of surgery versus conservative therapy for CTS	Study protocol
Schrijver *et al.*	2005	Correlating nerve conduction studies and clinical outcome measures on CTS: Lessons from a randomized controlled trial	Study comparing nerve conduction and clinical improvement using the data from the study of Ref. 26
Korthals-de Bos *et al.*	2006	Surgery is more cost-effective than splinting for CTS in the Netherlands: Results of an economic evaluation alongside a randomized controlled trial	Cost-effectiveness study using the data from Ref. 26
Pomerance *et al.*	2009	The cost-effectiveness of non-surgical versus surgical treatment for CTS	Cost-effectiveness study
Vogelin *et al.*	2010	Sonographic follow-up of patients with CTS undergoing surgical or non-surgical treatment: Prospective cohort study	Different outcome measures. Study measuring the size of the carpal tunnel after the intervention.
Onuma *et al.*	2013	Bilateral CTS due to gouty tophi: Conservative and surgical treatment in different hands of the same patient	Case-control study

Appendix C.

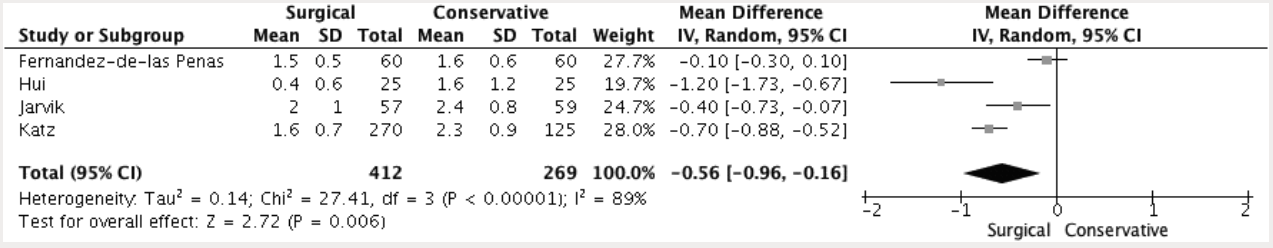

Fig. C.1. Detailed meta-analysis results for symptom improvement if only high quality studies[17,18,22,26] were included.

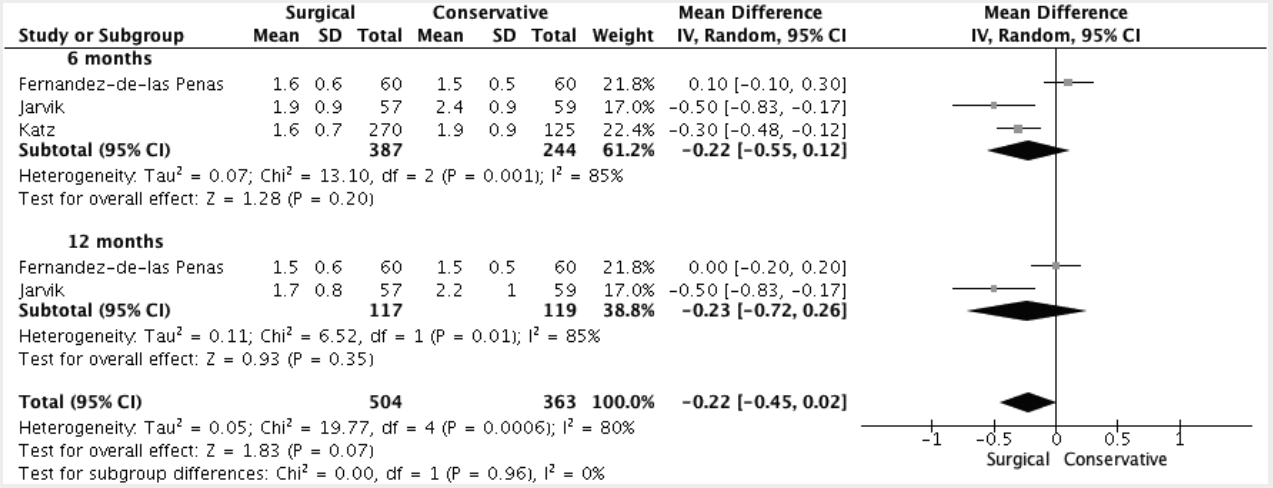

Fig. C.2. Detailed meta-analysis results for functional improvement at 6 and 12 months if only high quality studies[17,18,26] were included.

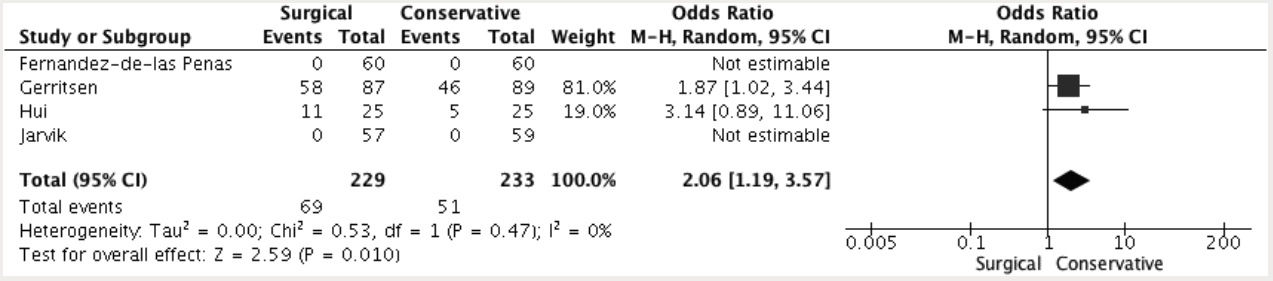

Fig. C.3. OR (95%CI) of effect of surgical and conservative treatment on complications reported if only high quality studies[17,18,26] were included.

Table C.1. Qualitative analysis (symptom improvement — 3 months).

	Favors surgery	No difference	Favors conservative
High quality	Ref. 26	Ref. 28	
Low quality	Ref. 25	Ref. 26	Refs. 20 and 22

Table C.2. Qualitative analysis (symptom improvement — 6 months).

	Favors surgery	No difference	Favors conservative
High quality	Refs. 24, 26 and 27	Ref. 28	
Low quality	Refs. 22, 26 and 30	Ref. 20	

Table C.3. Qualitative analysis (symptom improvement — 12 months).

	Favors surgery	No difference	Favors conservative
High quality	Refs. 26 and 27	Ref. 28	
Low quality		Ref. 20	

Table C.4. Qualitative analysis (symptom improvement — 18 months).

	Favors surgery	No difference	Favors conservative
High quality	Refs. 26 and 31*		
Low quality			

Note: *Observational study.

Table C.5. Qualitative analysis (symptom improvement — >18 months).

	Favors surgery	No difference	Favors conservative
High quality	Ref. 31*		
Low quality	Refs. 29 and 33	Ref. 21	

Note: *Observational studies.

Table C.6. Qualitative analysis (functional improvement — 3 months).

	Favors surgery	No difference	Favors conservative
High quality		Ref. 26	Ref. 28
Low quality		Ref. 23	Refs. 20 and 22

Table C.7. Qualitative analysis (functional improvement — 6 months).

	Favors surgery	No difference	Favors conservative
High quality	Refs. 24, 26, 27 and 31*	Ref. 28	
Low quality	Ref. 23	Refs. 20 and 22	

Note: *Observational study.

Table C.8. Qualitative analysis (functional improvement — 12 months).

	Favors surgery	No difference	Favors conservative
High quality	Refs. 26 and 27	Ref. 28	
Low quality	Ref. 20		

Table C.9. Qualitative analysis (functional improvement — 18 months).

	Favors surgery	No difference	Favors conservative
High quality	Ref. 31*	Ref. 26	
Low quality			

Note: *Observational study.

Table C.10. Qualitative analysis (functional improvement — >18 months).

	Favors surgery	No difference	Favors conservative
High quality	Ref. 31		
Low quality	Refs. 29* and 33		

Note: *Observational studies.

Table C.11. Qualitative analysis (improvement of neurophysiological parameters — 3 months).

	Favors surgery	No difference	Favors conservative
High quality			
Low quality		Ref. 23 (apart from SNCV and DSL) and Ref. 22	

Table C.12. Qualitative analysis (improvement of neurophysiological parameters — 6 months).

	Favors surgery	No difference	Favors conservative
High quality		Ref. 24 (apart from SNCV)	
Low quality		Ref. 23 (apart from SNCV and DSL), Refs. 22 and 30	

Table C.13. Qualitative analysis (improvement of neurophysiological parameters — 12 months).

	Favors surgery	No difference	Favors conservative
High quality	Ref. 26 (apart from DSL)	Ref. 26 (apart from DML)	
Low quality	Refs. 21* and 34		

Note: *Observational study.

Table C.14. Symptom severity baseline measurements.

Study	Measurement tool	Intervention	Baseline measurement
Ref. 31	Symptom severity scale	Surg:	3.2 ± 0.8
		Cons:	2.6 ± 0.8
Ref. 23	BQ	Surg:	3.4 ± 0.7
		Cons:	3.3 ± 0.7
Ref. 26	Symptom severity scale	Surg:	2.5 (1.9–3.1)
		Cons:	2.4 (1.8–2.9)
Ref. 24	GSS/10	Surg:	2.86 ± 1.10
		Cons:	2.52 ± 1.05
Ref. 22	BQ	Surg:	3.09 ± 0.5
		Cons (Splinting):	2.66 ± 0.35
		Cons (Splint + injection):	2.79 ± 0.63
Ref. 27	CTSAQ	Surg:	2.95 ± 0.77
		Cons:	3.01 ± 0.64
Ref. 25	GSS/10	Surg:	3.545 ± 0.74
		Cons:	3.48 ± 0.81
Ref. 28	BQ	Surg:	2.7 ± 0.6
		Cons:	2.5 ± 0.7

Table C.15. Functional status baseline measurements.

Study	Measurement tool	Intervention	Baseline measurement
Ref. 31	Functional status scale	Surg:	2.7 ± 0.9
		Cons:	2.1 ± 0.9
Ref. 23	BQ	Surg:	3.3 ± 1.0
		Cons:	3.0 ± 0.8
Ref. 26	Functional status scale	Surg:	2.3 (1.5–3.0)
		Cons:	2.0 (1.5–2.9)
Ref. 20	VAS scale/20	Surg:	1.95 ± 1.40
		Cons:	1.895 ± 1.318
Ref. 22	BQ	Surg:	2.7 ± 0.62
		Cons (Splinting):	2.47 ± 0.65
		Cons (Splint + injection):	2.19 ± 0.51
Ref. 27	CTSAQ	Surg:	2.42 ± 0.82
		Cons:	2.53 ± 0.82
Ref. 28	BQ	Surg:	2.4 ± 0.6
		Cons:	2.3 ± 0.5

References

1. Mediouni Z, Bodin J, Dale AM, et al. Carpal tunnel syndrome and computer exposure at work in two large complementary cohorts. BMJ Open 2015; 5(9):e008156.
2. Tanaka S, Wild DK, Seligman PJ, Behrens V, Cameron L, Putz-Anderson V. The US prevalence of self-reported carpal tunnel syndrome: 1988 National Health Interview Survey data. Am J Public Health 1994;84(11):1846–8.
3. de Krom MC, Knipschild PG, Kester AD, Thijs CT, Boekkooi PF, Spaans F. Carpal tunnel syndrome: Prevalence in the general population. J Clin Epidemiol 1992;45(4):373–6.
4. Atroshi I, Gummesson C, Johnsson R, Ornstein E, Ranstam J, Rosen I. Prevalence of carpal tunnel syndrome in a general population. Jama 1999; 282(2):153–8.
5. Castro Ado A, Skare TL, Nassif PA, Sakuma AK, Barros WH. Sonographic diagnosis of carpal tunnel syndrome: A study in 200 hospital workers. Radiol Bras 2015;48(5):287–91.
6. Kostopoulos D. Treatment of carpal tunnel syndrome: A review of the non-surgical approaches

with emphasis in neural mobilization. J Bodyw Mov Ther 2004;8(1):2–8.
7. Ibrahim I, Khan WS, Goddard N, Smitham P. Carpal tunnel syndrome: A review of the recent literature. Open Orthop J 2012;6:69–76.
8. Rempel D, Dahlin L, Lundborg G. Pathophysiology of nerve compression syndromes: Response of peripheral nerves to loading. J Bone Joint Surg Am 1999;81(11):1600–10.
9. Bland JD, Rudolfer SM. Clinical surveillance of carpal tunnel syndrome in two areas of the United Kingdom, 1991–2001. J Neurol Neurosurg Psychiatry 2003;74(12):1674–79.
10. Shi Q, MacDermid JC. Is surgical intervention more effective than non-surgical treatment for carpal tunnel syndrome? A systematic review. J Orthop Surg Res 2011;6:17.
11. Verdugo RJ, Salinas RA, Castillo JL, Cea JG. Surgical versus non-surgical treatment for carpal tunnel syndrome. Cochrane Database Syst Rev 2008(4):CD001552.
12. 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):1290–9.
13. 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):713–21.
14. Macedo LG, Elkins MR, Maher CG, Moseley AM, Herbert RD, Sherrington C. There was evidence of convergent and construct validity of Physiotherapy Evidence Database quality scale for physiotherapy trials. J Clin Epidemiol 2010;63(8):920–5.
15. Moga C, Guo B, Schopflocher D, Harstall C. Development of a quality appraisal tool for case series studies using a modified Delphi technique. Institute of Health Economics: Edmonton AB; 2012.
16. Sedgwick P. Meta-analyses: Heterogeneity and subgroup analysis. Bmj 2013;346(2):f4040.
17. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21 (11):1539–58.
18. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Bmj 1997;315(7109):629–34.
19. Harter BT, Jr., McKiernan JE, Jr., Kirzinger SS, Archer FW, Peters CK, Harter KC. Carpal tunnel syndrome: Surgical and nonsurgical treatment. J Hand Surg 1993;18(4):734–9.
20. Ly-Pen D, Andreu JL, de Blas G, Sanchez-Olaso A, Millan I. Surgical decompression versus local steroid injection in carpal tunnel syndrome: A one-year, prospective, randomized, open, controlled clinical trial. Arthritis Rheum 2005;52(2):612–9.
21. Seror P. Nerve conduction studies after treatment for carpal tunnel syndrome. J Hand Surg 1992;17 (6):641–5.
22. Ucan H, Yagci I, Yilmaz L, Yagmurlu F, Keskin D, Bodur H. Comparison of splinting, splinting plus local steroid injection and open carpal tunnel release outcomes in idiopathic carpal tunnel syndrome. Rheumatol Int 2006;27(1):45–51.
23. Demirci S, Kutluhan S, Koyuncuoglu R, et al. Comparison of open carpal tunnel release and local steroid treatment outcomes in idiopathic carpal tunnel syndrome. Rheumatol Int 2002; 22(1):33–7.
24. Hui AC, Wong S, Leung CH, et al. A randomized controlled trial of surgery vs steroid injection for carpal tunnel syndrome. Neurology 2005;64 (12):2074–8.
25. Ullah I. Local steroid injection or carpal tunnel release for carpal tunnel syndrome — which is more effective? J Postgrad Med Instit (Peshawar–Pakistan), North America, 27, Mar. 2013.
26. Gerritsen AA, de Vet HC, Scholten RJ, Bertelsmann FW, de Krom MC, Bouter LM. Splinting vs surgery in the treatment of carpal tunnel syndrome: A randomized controlled trial. Jama 2002;288(10):1245–51.
27. Jarvik JG, Comstock BA, Kliot M, et al. Surgery versus non-surgical therapy for carpal tunnel syndrome: A randomised parallel-group trial. Lancet 2009;374(9695):1074–81.
28. Fernandez-de-Las Penas C, Ortega-Santiago R, de la Llave-Rincon AI, et al. Manual physical therapy versus surgery for carpal tunnel syndrome: A randomized parallel-group trial. J Pain 2015; 16(11):1087–94.
29. Ettema AM, Amadio PC, Cha SS, Harrington JR, Harris AM, Offord KP. Surgery versus conservative therapy in carpal tunnel syndrome in people aged 70 years and older. Plast Reconstr Surg 2006;118(4):947–58; discussion 959–60.
30. Elwakil TF, Elazzazi A, Shokeir H. Treatment of carpal tunnel syndrome by low-level laser versus open carpal tunnel release. Lasers Med Sci 2007;22(4):265–70.
31. Katz JN, Keller RB, Simmons BP, et al. Maine carpal tunnel study: Outcomes of operative and nonoperative therapy for carpal tunnel syndrome in a community-based cohort. J Hand Surgery 1998;23(4):697–710.
32. Leite JC, Jerosch-Herold C, Song F. A systematic review of the psychometric properties of the boston carpal tunnel questionnaire. BMC Musculoskelet Disord 2006;7:78.
33. Ly-Pen D, Andreu JL, Millan I, de Blas G, Sanchez-Olaso A. Comparison of surgical decompression and local steroid injection in the treatment

of carpal tunnel syndrome: 2-year clinical results from a randomized trial. Rheumatology 2012;51(8): 1447–54.
34. Andreu JL, Ly-Pen D, Millan I, de Blas G, Sanchez-Olaso A. Local injection versus surgery in carpal tunnel syndrome: Neurophysiologic outcomes of a randomized clinical trial. Clin Neurophysiol 2014;125(7):1479–84.
35. Werner RA, Andary M. Electrodiagnostic evaluation of carpal tunnel syndrome. Muscle Nerve 2011;44(4):597–607.
36. Wong SM, Griffith JF, Hui AC, Tang A, Wong KS. Discriminatory sonographic criteria for the diagnosis of carpal tunnel syndrome. Arthritis Rheum 2002;46(7):1914–21.
37. Girlanda P, Dattola R, Venuto C, Mangiapane R, Nicolosi C, Messina C. Local steroid treatment in idiopathic carpal tunnel syndrome: Short- and long-term efficacy. J Neurol 1993;240(3):187–90.
38. Finsen V, Russwurm H. Neurophysiology not required before surgery for typical carpal tunnel syndrome. J Hand Surg 2001;26(1):61–4.
39. Braun RM, Jackson WJ. Electrical studies as a prognostic factor in the surgical treatment of carpal tunnel syndrome. J Hand Surg 1994;19(6):893–900.
40. Redmond MD, Rivner MH. False positive electrodiagnostic tests in carpal tunnel syndrome. Muscle Nerve 1988;11(5):511–8.
41. Glowacki KA, Breen CJ, Sachar K, Weiss AP. Electrodiagnostic testing and carpal tunnel release outcome. J Hand Surg 1996;21(1):117–21.
42. Padua L, Padua R, Aprile I, Pasqualetti P, Tonali P, Italian CTSSGCts. Multiperspective follow-up of untreated carpal tunnel syndrome: A multicenter study. Neurology 2001;56(11):1459–66.
43. Kaplan SJ, Glickel SZ, Eaton RG. Predictive factors in the non-surgical treatment of carpal tunnel syndrome. J Hand Surg 1990;15(1):106–8.
44. Irwin LR, Beckett R, Suman RK. Steroid injection for carpal tunnel syndrome. J Hand Surg 1996; 21(3):355–7.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) 115–123
DOI: 10.1142/S1013702518500099

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

Reliability and validity of transfer assessment instrument version 3.0 in individuals with acute spinal cord injury in early rehabilitation phase

Preeti Baghel[1], Shefali Walia[1] and Majumi M Noohu[2,*]

[1]*ISIC Institute of Rehabilitation Sciences, New Delhi, India*

[2]*Centre for Physiotherapy and Rehabilitation Sciences, Jamia Millia Islamia, New Delhi, India*

mnoohu@jmi.ac.in

Received 9 January 2017; Accepted 28 July 2017; Published 14 August 2018

Background: Transfers are very important in functional activities of subjects with spinal cord injury (SCI). The transfer assessment instrument (TAI) was the first tool to standardize the assessment of transfer technique.
Objective: The purpose of this study was to establish the reliability and validity of TAI 3.0 in people with SCI in early rehabilitation phase.
Methods: Thirty subjects with acute traumatic SCI were recruited from a tertiary care center for SCI management. Four raters assessed the quality of transfer using TAI 3.0 and a fifth rater used global assessment of transfer scale (VAS). TAI 3.0's intraclass correlation coefficient (ICC) for intrarater and interrater reliability, standard error of measurement (SEM), minimal detectable change (MDC), limits of agreement and concurrent validity was determined.
Results: The intrarater ICC was 0.93 to 0.98 and interrater ICC was 0.99, indicating high levels of reliability. The SEMs among the raters for TAI 3.0 total was from 0.23 to 0.28. The MDC among the raters TAI 3.0 total was from 0.54 to 0.86. Correlation for different raters between the TAI 3.0 and VAS ranged between 0.88 and 0.90.
Conclusion: TAI 3.0 is a reliable and valid tool to assess the transfer skill in individuals with SCI in early rehabilitation phase.

Keywords: Spinal cord injury; transfers; activities of daily living.

*Corresponding author.

Introduction

Wheelchairs are the primary mode of mobility in individuals with spinal cord injury (SCI). The transfers are vital in performing activities of daily living in wheelchair users. A transfer is considered as a movement of oneself from one surface to another in one direction. If a full-time wheelchair user is unable to transfer, or transfers become more difficult, the individual's quality of life may be further affected.[1]

Transfers often cause excessive loading of the arms. This may lead to upper extremity pain and injuries, such as rotator cuff tears, elbow pain, and carpal tunnel syndrome.[2] The transfers may also place joints of upper limb in extreme positions and expose to high internal joint forces. Therefore, it seems that prevention may be critical in reducing upper limb pain and overuse injuries. Learning the transfer methods that reduces forces and awkward joint motions is an important strategy for preventing impairments in upper limb of wheelchair users.[3,4]

There is a wide variation in the amount and the type of transfer training and no uniform way to evaluate transfer quality in early stage of rehabilitation after SCI. Observation by a therapist and qualitative assessment was the standard method of evaluating transfers. Therefore, McClure *et al.* developed a tool called transfer assessment instrument (TAI) to assess the quality of transfer in full-time wheelchair users. It is a safe and easy tool to administer outcome measure to assess transfers in wheelchair users. It was found to have an acceptable interrater and intrarater reliability.[5] Tsai *et al.*, introduced the refined TAI version 3.0. It was found to have high reliability (0.74 to 0.88) among raters of different clinical backgrounds and experience in people using wheelchairs for more than a year.[6]

In early stage of rehabilitation following SCI, transfer evaluations are not done objectively and in a consistent manner. The transfer assessment can be influenced by the subjective experience of the therapists. It may lead to less accurate evaluations and variability in transfer skill assessment. TAI was the first tool to standardize the evaluation of transfer technique. The items included in the TAI were based on clinical practice guidelines, available information on literature, and best clinical practices for transfers.[5] There is a lack of valid assessment tool to evaluate transfers in early stage of rehabilitation following SCI. Therefore, the purpose of this study was to establish the reliability and validity of TAI 3.0 in people with SCI in early stage of rehabilitation.

Methods

Sample

A sample of convenience of 30 subjects with acute traumatic SCI who met the inclusion criteria and were willing to participate in the study was included. The sample was selected from rehabilitation department of Indian spinal injuries center, New Delhi. The subjects were in the early rehabilitation phase. The transfer training was started one week before the date of evaluation. The design was methodological research-repeated measure study. The study was approved by institutional ethical committee where the study was carried out.

Sample size

A sample size of 30 subjects with 2 observations per subject achieves 91% power to detect an intraclass correlation of 0.70 under the alternative hypothesis when the intraclass correlation under the null hypothesis is 0.30 using an F-test with a significance level of 0.05 was required for reliabilty testing. A sample size of 25 achieves 81% power with the alternative hypothesis correlation of 0.60 by using a two-sided hypothesis test with a significance level of 0.05 which was needed for validity testing.[7,8] The sample size was calculated using PASS 2008 software.

Subjects included were first time manual wheelchair users who were potential full-time wheelchair users with tetraplegia, high paraplegia (T2–T7) or low paraplegia (T8–L4). They were able to sit with or without hand support for 30 s. Subjects with following problems were excluded, such as, unhealed pressure ulcers, > 19 score on Beck depression inventory-II,[9,10] wheelchair user's shoulder pain index score > 8.5,[11] weight relief raises, musculoskeletal deformities of upper extremities, unstable medical condition (e.g., angina, seizures), respiratory distress, cardiovascular, emotional or psychiatric problems and with significant visual impairments.

Procedure

The subjects were recruited by convenient sampling method who met the inclusion criteria. The

purpose of the study was explained to the subjects who gave the consent to participate in the study. Then their age (years), gender, height (cm), weight (kg), level of injury and American spinal injury association (ASIA) impairment level,[12] area of body affected, spinal cord independence-III measure (SCIM) score,[13] modified functional reach test score,[14] type of transfer, duration since injury (months), body mass index (BMI) were collected.

Five physiotherapists who had an experience of 3–4 years in rehabilitation of the people with spinal cord injury (SCI) were the raters for transfers done by the participants. Four raters evaluated the transfer skill of participants on TAI 3.0 while the fifth therapist evaluated the transfer on global assessment of transfer scale (VAS). Before the testing, the raters were explained how to administer TAI 3.0 using the text instruction manual of the assessment tool. The instructions contained the details and how to score each item. In addition, general recommendations were provided including where to stand when observing different aspects of transfer. The raters were also told about the instructions to be given to participants during transfers. The study investigator ensured that each rater was trained properly using TAI 3.0.[6] The raters got individualized instruction on items with difficulty in scoring.

Participants were asked to perform four transfers each in two sessions (sessions 1 and 2). The transfers were done from their own wheelchair, to and fro from a wheelchair level surface bed or a height adjustable hospital bed with their usual way of transfer. The height of the hospital bed was adjusted, depending on the individual's preference. An attendant was there with subject to prevent any fall or provide assistance during transfers. If participants needed assistance for transfers, they were permitted to use transfer device (e.g., transfer board or lift) or the attendant to provide assistance. In session 1, while participants performed transfer, four raters used TAI 3.0 to score and evaluate their transfer skills. For session 2, all participants were asked to return after 72 h to perform the transfer again. Then the same four raters evaluated them for second time using TAI 3.0. Each session lasted for 30 to 45 min per subject.

TAI scoring (TAI 3.0)

The TAI contains two parts. Part 1 comprises of 15 items and is scored as follows: "yes", 1 point, "no", 0 points or "not applicable" (N/A), which means a removed item. Part 1 is completed after each transfer and item scores are averaged to produce a single representative item score. The part 1 is the summation of each item's score multiplied by 10 and then divided by the number of applicable items, ranging from 0 to 10. The items in part 2 are completed after all transfers have been performed. The 12 items in part 2 are scored on a likert scale ranging from 0 (strongly disagree) to 4 (strongly agree). The part 2 score is the summation of each item's score multiplied by 2.5 and then divided by the number of applicable items, resulting in a range of scores from 0 to 10. The final score of TAI is the average of the part 1 and part 2. The items of the instrument and what is evaluated during a transfer are given in Table 1. All the recruited subjects completed both the sessions (session 1 and session 2) of the study.[6]

Global assessment of transfer scale

This is a likert scale which rates the overall transfer quality on a 10-point scale. The participants' transfer from a wheelchair was evaluated and graded from poor (0) to excellent (10). The criteria for rating a transfer as poor were if the individual does not make use of equipment when needed, do not make transfer easy and safe and inappropriate placement of the hand and feet. A transfer was rated excellent if the transfer was appropriately done without transfer devices, easy and safe, placing the hand and feet on right places and using human assistance when need. During session 1, while four raters evaluated the transfer skill of the participants, the fifth therapist who has not seen the TAI rated the study participants transfer skill on a VAS. The VAS evaluation was done only in session 1.[5]

Data Analysis

Data analysis was performed using SPSS 21.0 software. Descriptive statistics were calculated for the subject's demographic data including age, gender, type of transfer, type of disability, BMI. The intraclass correlation coefficients (ICCs) within each rater and between raters in part 1, part 2 and final TAI 3.0 scores were calculated to assess reliability. The limits of agreement (LOA) analysis were done by plotting Bland and Altman (B&A)

Table 1. The items and evaluation component of the TAI 3.0.

	Item no.		What is being evaluated
Part 1		1	Distance between the wheelchair and object to which he/she is transferring on to. The subject's wheelchair is within 3 inches of the object to which he is transferring on to.
		2	The angle between the subject's wheelchair and the surface to which he is transferring.
		3	Whether the subject attempts to place his chair to perform the transfer forward of the rear wheel.
		4	If possible, the subject removes his armrest independently or with assistance.
		5	Level or downhill transfer.
		6	Placement of feet in a stable position.
		7	Scoots to the front edge of the wheelchair seat before he transfers.
		8	Hands' position.
		9	Handgrip of the leading arm.
		10	Handgrip of the trailing arm.
		11	Control over flight.
		12	Head–hip relationship.
		13	Positioning of the lead arm.
		14	The landing phase of the transfer.
		15	The assistant supporting the subject's arms during the transfer.
Part 2	Weight bearing arm position	1	The lead arm position.
	Set-up phase	2	Sets up for a safe and easy transfer.
		3	Change the height of the object he is transferring to/from to make the transfer level.
		4	Gets close to the object that he is transferring on to.
		5	Uses handgrips when necessary.
	Conservation	6	Uses a transfer device when necessary.
		7	Alternate the leading/trailing arm over the course of the assessment.
	Quality	8	Transfer is smooth and well controlled.
		9	Clearly communicate his needs in transfer.
		10	Does not allow the assistant to pull on his arms during the transfer.
		11	The subject corrects the assistant.
		12	The subject is able to correctly direct his care in an assertive and polite manner.

plots with graph pad prism software (Prism version 6.00).[15] To determine variability of TAI 3.0 scores, standard error of measurement (SEM) and minimal detectable change (MDC) were analyzed. Statistical significance was set at $p \leq 0.05$ with confidence interval of 95%. To establish convergent validity, Pearson correlation coefficients were calculated for each rater to evaluate the correlation of TAI 3.0 scores (total) with global assessment of transfer scores.

Reliability Testing

Interrater and intrarater reliability

For calculating intrarater and interrater reliability, ICC coefficient value (ICC 3,1 : two-way mixed effect and consistency) was calculated separately for part 1, part 2 and total score of TAI3.0. ICCs higher than 0.80 were considered strong, between 0.60 and 0.79 were acceptable, between 0.40 and 0.59 were moderate, and lower than 0.40 were weak.[16]

Standard error of measurement

The SEM provides a value for measurement error in the same units as the measurement itself, it is a measure of absolute reliability.[17] The SEM was calculated for part 1, part 2 and final TAI scores using the formula: $\text{SEM} = \text{SD} \times (1 - r)^{1/2}$, where SD is the standard deviation of the dataset and r is the reliability coefficient or ICC value. The SEM was calculated for individual rater based on ICC values in case of intrarater reliability analysis.

Minimal detectable change

The MDC was calculated for individual rater for intrarater reliability analysis. It is an estimate of the smallest change in score that can be detected

objectively for a subject, it is the amount by which a subject's score needs to change and ensure that the change is greater than measurement error.[18] MDC was analyzed based on 95% confidence interval (MDC = $1.96 \times 2^{1/2} \times$ SEM) for part 1, part 2 and final TAI scores.

Limits of agreement

LOA between session 1 and 2 were determined by B&A method.[19]

Validity Testing

Convergent validity

Pearson correlation coefficient was calculated for each rater by correlating their final TAI 3.0 scores with global assessment of transfer scores (VAS) of session 1.

Results

Demographic characteristic of the subjects, mean ± SD, such as age (years), height (cm), weight (kg), BMI, SCIM score, modified functional reach test score, and frequency distribution of area of body affected, ASIA impairment level, gender, type of transfer, and duration since injury are shown in Table 2. The mean ± SD of part 1 (TAI 1), part 2 (TAI 2) and total score (TAI total) of TAI at two different time points for all four raters is given in Table 3. The mean ± SD global assessment of transfer scores was 7.9 ± 1.48.

Reliability

The intrarater ICCs ranged from 0.93 to 0.95 for TAI 1, 0.97 to 0.98 for TAI 2 and TAI total, suggesting high levels of reliability (Table 4). The interrater ICC of TAI 1 at first time was 0.98 and for TAI 2 and TAI total was 0.99 and for the second time, ICC of TAI 1, TAI 2 and TAI total was 0.99 (Table 5).

SEM and MDC

The SEM among the raters for TAI 1 was from 0.34 to 0.43, TAI 2 was from 0.23 to 0.27 and for TAI total was from 0.23 to 0.28. The MDC among the raters for TAI 1 ranged from 1.19 to 0.94, TAI 2 was from 0.69 to 0.86 and for TAI total was from 0.54 to 0.86 (Table 3). Between the raters, SEM and MDC is given in Table 5.

Table 2. Demographic characteristics of the sample.

Variable		Mean ± SD/*n* (%)
Age (years)		31.9 ± 12.3
Height (cm)		163.78 ± 9.64
Weight (kg)		62.61 ± 13.49
BMI (kg/m^2)		23.33 ± 2.04
Duration(months)		1.33 ± 0.47
SCIM-III score		40.32 ± 11.64
mFR (cm)		10.27 ± 3.05
Gender	Male	25 (83.3)
	Female	5 (16.7)
Area of body affected	Tetraplegia	6 (20)
	High paraplegia (T2–T7)	5 (16.7)
	Low paraplegia (T8–L4)	19 (63.3)
ASIA impairment level	A	11 (36.7)
	B	11 (36.7)
	C	5 (16.7)
	D	3 (10)
Type of transfer	Independent sitting pivot	20 (66.7)
	Assisted sitting pivot	10 (33.3)

Notes: SCIM: spinal cord independence measure; mFR: modified functional reach; BMI: body mass index; SD: Standard deviation, *n*: number.

Table 3. Mean ± SD of TAI at two different time points.

Raters	Time 1			Time 2		
	TAI 1	TAI 2	TAI total	TAI 1	TAI 2	TAI total
1	8.42 ± 1.56	7.62 ± 1.90	7.91 ± 1.68	8.30 ± 1.45	7.764 ± 1.89	8.03 ± 1.61
2	8.05 ± 1.57	7.571 ± 1.81	7.81 ± 1.65	8.23 ± 1.43	7.77 ± 1.85	8.00 ± 1.60
3	8.21 ± 1.45	7.63 ± 1.89	7.92 ± 1.61	8.23 ± 1.52	7.83 ± 1.80	8.03 ± 1.61
4	8.12 ± 1.69	7.71 ± 1.87	7.91 ± 1.71	8.16 ± 1.52	7.96 ± 1.78	8.06 ± 1.60

Notes: TAI 1: Part 1 of TAI; TAI 2: Part 2 of TAI; TAI total: TAI total score.

Table 4. Intrarater reliability analysis for TAI 3.0.

Raters	ICC (95 CI)			MDC			SEM		
	TAI 1 (95 CI)	TAI 2 (95 CI)	TAI total (95 CI)		TAI 2	TAI total		TAI 2	TAI total
1	0.95 (0.89–0.97)	0.98 (0.94–0.99)	0.98 (0.94–0.98)	0.94	0.75	0.64	0.34	0.27	0.23
2	0.94 (0.88–0.97)	0.98 (0.96–0.99)	0.98 (0.94–0.98)	1.02	0.69	0.40	0.37	0.25	0.23
3	0.94 (0.86–0.96)	0.98 (0.96–0.99)	0.98 (0.94–98)	1.02	0.72	0.86	0.37	0.26	0.23
4	0.93 (0.85–0.96)	0.97 (0.94–0.98)	0.97 (0.93–0.98)	1.19	0.86	0.78	0.43	0.23	0.28

Notes: TAI 1: Part 1 of TAI; TAI 2: Part 2 of TAI; TAI total: TAI total score; ICC Intraclass correlation coefficient; MDC: minimum detectable change; SEM: standard error of measurement; CI: confidence interval.

Table 5. Interrater reliability analysis for TAI 3.0.

Item	ICC ($_{95}$CI)		MDC		SEM	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Part 1	0.98 (0.97–0.99)	0.99 (0.94–0.99)	0.44	0.42	0.16	0.15
Part 2	0.99 (0.98–0.99)	0.99 (0.94–0.99)	0.53	0.50	0.19	0.18
Total	0.99 (0.96–0.99)	0.99 (0.96–0.99)	0.47	0.44	0.17	0.16

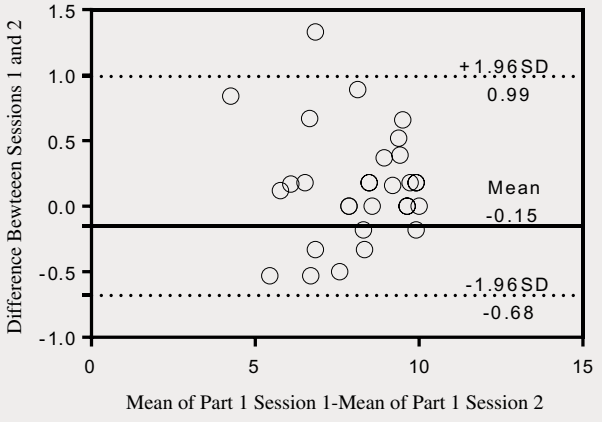

Fig. 1. B&A plot of agreement between sessions 1 and 2 for part 1 score. The figure reveals that only one data point lies outside ±1.96 SD.

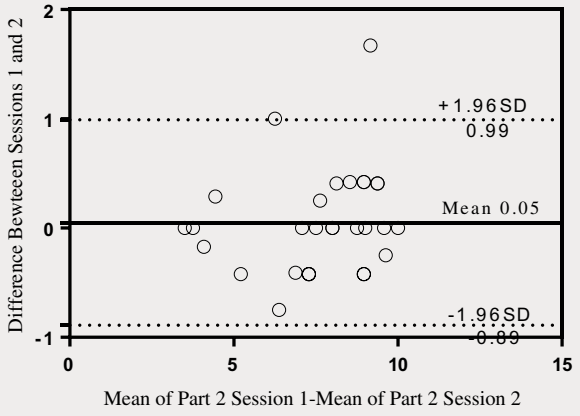

Fig. 2. B&A plot of agreement between sessions 1 and 2 for part 2 score. The figure reveals that only one data point lies outside ±1.96 SD.

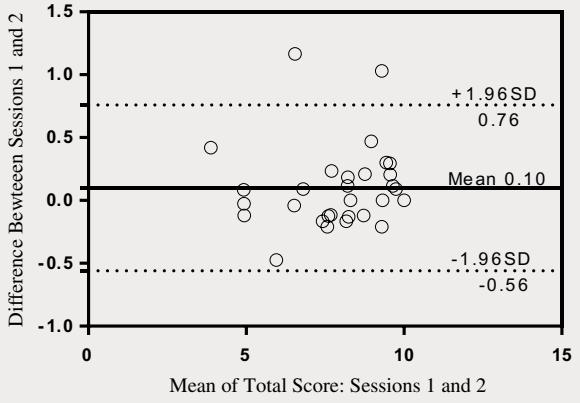

Fig. 3. B&A plot of agreement between sessions 1 and 2 for total score. The figure reveals that only two data points lie outside ±1.96 SD.

LOA Plots

The LOA plot showed that there was 1 data point for part 1 and part 2 outside +1.96 SD (Figs. 1 and 2) and for total score, there was 2 data points outside +1.96 SD (Fig. 3).

Table 6. Correlation of total TAI score with the criterion standard (VAS).

Rater	Pearson correlation coefficient	*p*
1	0.89	0.001
2	0.89	0.001
3	0.88	0.001
4	0.90	0.001

Convergent validity

Pearson correlation coefficients ranged from 0.88 to 0.90 among the raters with the VAS score ($p = 0.001$) as given in Table 6.

Discussion

Measurements of functional outcomes are an integral part of any goal-orientated, interdisciplinary rehabilitation program. It is important for quantifying the success of rehabilitation program. A good clinical assessment tool should be both reliable and valid.[20,21] The study results showed that TAI 3.0 has higher levels of intrarater (ICCs ranged from 0.93 to 0.98) and interrater reliability (ICC — 0.99). For convergent validity, correlations ranged between 0.88 and 0.90. Results

showed significant correlations between TAI and global assessment of transfer skills. The findings are similar, but psychometric properties appear better than the previous reliability and validity analysis done in wheelchair users with various disabilities with the same scale.[5,6] The reason may be that all the subjects were with SCI and from the same rehabilitation care setup. This might have made the sample more homogenous and another component might be the uniformity in transfer training.

The reason for choosing ICCs to analyze the reliability of the TAI is because the ICC measures the association and agreement. The ICC can be used to assess reliability for more than two raters and can be used to analyze ordinal type of data.[6] The mean difference determined by LOA analysis was very small. Only total of four data points (one each for parts 1 and 2 and 2 for total score) were outside the 95% confidence interval limits. Hence, there was an agreement with the two sessions of measurements.[22] B&A plot is the quantification of the agreement between two measurements by plotting it graphically, the mean difference and constructing LOA. The difference of the two paired measurements is plotted against the average of the two measurements. The LOA recommended by B&A is that 95% of the data points should lie within ± 2SD of the mean difference. The part 1 and part 2 scores meet the LOA criteria whereas total score is slightly less at 93.33%.[19,23]

Scores may vary, given expected variability of individual performance and measurement error. A measure of absolute variability provides useful information to delineate the "expected" changes from "true" changes in performance. Statistically, absolute reliability is determined by the SEM. Clinically useful mechanism for looking at absolute reliability is the MDC score.[24] Results showed that SEMs ranged from 0.23 to 0.28 within raters and 0.15 to 0.19 between raters. The smaller the SEM, the more accurate are the assessments that are being made. The smaller SEM in this study further indicates the accuracy of measurements with TAI 3.0.[25] This study found that the MDC ranged from 0.64 to 0.86 within raters and 0.47 to 0.44 between raters. Minimum of 0.86 point change would be needed to identify a true difference in transfer skills that is not a measurement error. The MDC is relatively easy to calculate which provides clinically relevant information. The limitation of MDC is that it assumes that detectable changes are uniform throughout the scale, but the measurement error can vary at different points in the scale.[25] The raters were given handouts with an explanation of each item, a description of different scoring scenarios and a short practice session. Instructions administering TAI 3.0 might have improved the consistency among raters.[6]

Currently, no other outcome measure exists to assess transfer quality in population with SCI population in early rehabilitation phase. Hence, global assessment (VAS) was used to evaluate convergent validity. Using a non-validated tool to evaluate convergent validity is not only a preferred option, but also cannot be avoided because of a lack of a comparable criterion standard. The results showed an excellent convergent validity. The VAS was also previously used to establish the concurrent validity of TAI.[5]

The objective evaluation of transfers may help the clinicians to improve the transfer training, identify and correct improper transfer techniques. The identification of improper transfer techniques may prevent musculoskeletal injuries and pain in the upper extremities. The evaluation was done only on participants in early phase of rehabilitation, who performed independent or assisted sitting pivot transfers, so results cannot be generalized to all full-time wheelchair users. Future researches should be done to find out the effect of transfer training program on changes in TAI score in people with SCI.

Conclusion

The TAI is a reliable and valid tool which can be used as an outcome measure to evaluate transfer quality in people with acute SCI in early rehabilitation phase.

Conflict of Interest

We hereby declare that there is no conflict of interest involved in this study in terms of monetary benefits or in any other form.

Funding/Support

The study did not receive any funds and it was supported by Indian Spinal Injuries Center, New Delhi, India, by providing space to carry out the study. We acknowledge the support of Chitra

Kataria, Principal, and all the supporting staff of ISIC institute of rehabilitation sciences, New Delhi, India for this study.

Author Contributions

The conception and design of the study, or acquisition of data were made by Preeti Baghel, Shefali Walia and Majumi. The analysis and interpretation of data were carried out by Shefali Walia and Majumi. Drafting the paper or revising it critically for important intellectual content was carried out by Preeti Baghel, Shefali Walia and Majumi. All authors approved the final submitted version of the manuscript.

References

1. Rice LA, Smith I, Kelleher AR, Greenwald K, Hoelmer C, Boninger ML. Impact of the clinical practice guideline for preservation of upper limb function on transfer skills of persons with acute spinal cord injury. Arch Phys Med Rehabil 2013;94:1230–46.
2. Preservation of Upper Limb Function Following Spinal Cord Injury: A Clinical Practice Guideline for Health-Care Professionals. Washington, DC: Paralyzed Veterans of America, 2005.
3. Tsai, CY, Hogaboom NS, Boninger ML, Koontz AM. The relationship between independent transfer skills and upper limb kinetics in wheelchair users. BioMed Res Int 2014:984526.
4. Mohammed K, Dunn JA. Shoulder pain in tetraplegia. Orthop Trauma 2014:28:27–32.
5. McClure LA, Boninger ML, Ozawa H, Koontz AM. Reliability and validity analysis of the transfer assessment instrument. Arch Phys Med Rehabil 2011;92:499–508.
6. Tsai CY, Rice LA, Hoelmer C, Boninger ML, Koontz AM. Basic psychometric properties of the transfer assessment instrument (version 3.0). Arch Phys Med Rehabil 2013;94:2456–64.
7. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998;17:101–10.
8. Zar JH, Biostatistical Analysis. 2nd ed. Engelwood Cliffs, NJ:Prentice-Hall, 1984.
9. Arnau RC, Meagher MW, Norris MP, Bramson R. Psychometric evaluation of the beck depression inventory-II with primary care medical patients. Health Psychol 2001;20:112–9.
10. Shin JC, Goo HR, Yu SJ, Kim DH, Yoon SY. Depression and Quality of Life in Patients within the First 6 Months after the Spinal Cord Injury. Ann Rehabil Med 2012;36:119–25.
11. Curtis KA, Roach KE, Applegate EB et al., Reliability and validity of the Wheelchair User's Shoulder Pain Index (WUSPI). Paraplegia 1995;33: 595–601.
12. Maynard FM Jr, Bracken MB, Creasey G et al., International standards for neurological and functional classification of spinal cord injury. Spinal Cord 1997;35:266–74.
13. Itzkovich M, Gelernter I, Biering-Sorensen F et al., The Spinal Cord Independence Measure (SCIM) version III: Reliability and validity in a multi-center international study. Disabil Rehabil 2007;29: 1926–33.
14. Lynch SM, Leahy P, Barker SP. Reliability of measurements obtained with a modified functional reach test in subjects with spinal cord injury. Phys Ther 1998;78:128–33.
15. Prism version 6.00 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com.
16. Krebs DE. Declare your ICC type. Phys Ther 1986;66:1431.
17. Harvill LM. Standard error of measurement. Educ Meas 1991;10:33–41.
18. Finch E, Brooks D, Stratford PW, Mayo NE (eds.) Physical rehabilitation outcome measures: A guide to enhanced clinical decision making. 2nd ed. Hamilton: Canadian Physiotherapy Association, BC Decker Inc, 2002.
19. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
20. Emerich L, Parsons KC, Stein A. Competent care for persons with SCI and dysfuction in acute inpatient rehabilitation. Top Spinal Cord Inj Rehabil 2012;18:149–66.
21. Sullivan GM. A primer on the validity of assessment instruments. J Grad Med Edu 2011;3:119–20.
22. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: An illustration of appropriate statistical analyses. Clin Rehabil 1998;12: 187–99.
23. Giavarina D. Understanding Bland Altman analysis. Biochem Med 2015;25:141–51.
24. Donoghue D. How much change is true change? The minimum detectable change of the berg balance scale in elderly people. J Rehabil Med 2009;41:343–6.
25. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland–Morris questionnaire. Phys Ther 1996;76:359–65.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) 125–131
DOI: 10.1142/S1013702518500105

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

Agreement of clinical examination for low back pain with facet joint origin

Mantana Vongsirinavarat*, Wahyuddin Wahyuddin and Ratchaneewan Adisaiphaopan

Faculty of Physical Therapy, Mahidol University, Nakornpathom 73170, Thailand

**mantana.von@mahidol.ac.th*

Received 31 January 2017; Accepted 28 July 2017; Published 14 August 2018

Background: Low back pain (LBP) with facet joint origin is a common diagnosis of patients referred to physical therapy clinic. An expert consensus of diagnostic criteria has been proposed. However, the reliability of the assessment has not been proved.
Objective: To test the degrees of agreement between two physical therapists for nine physical examination items and the diagnosis of facet joint origin.
Methods: The examination according to diagnostic criteria was performed independently by two physical therapists in 45 patients with chronic LBP. The percent agreements and Kappa coefficients of each examination item and diagnostic conclusion were calculated.
Results: The percent agreements of nine examined items ranged from 73.3–91.1%. The Kappa coefficients, widely ranged from 0.250–0.690 ($p = 0.48$ to < 0.001), showed statistically significant agreements for all examination items. The low level of agreements was partly due to improper distributions of test results. The agreement of conclusion was 86.7% and Kappa coefficient was 0.492 ($p = 0.001$) which reflected good agreement of facet diagnosis.
Conclusion: There were adequate agreements for clinical examination of LBP with facet joint origin. The low level of agreement suggested the clinicians to have operational definition and rigorous training sessions although the examinations seemed to be routinely performed.

Keywords: Low back pain; facet joint; assessment; reliability; agreement.

Introduction

Among musculoskeletal complaints, low back pain (LBP) is the symptom most frequently leading to physician and physical therapy visits.[1] Several anatomical sites in low back were considered the source of pain including facet joint, intervertebral

*Corresponding author.

disc, ligaments, nerve root and dura, muscles, and fascia.[2] A number of factors such as neurological, mental stress, and social status are also reportedly contributing to LBP.[3,4]

Facet or zygapophyseal joint was indicated to be the cause in 42% of persons with LBP using controlled comparative local anesthetic diagnostic blocks for diagnostic confirmation,[5] especially in the chronic cases.[6] A community-based survey in older adults showed evident facet degeneration on CT imaging associated with LBP.[7]

A study reported that when physical therapists use the combination of McKenzie lumbar spine assessment algorithms with a series of pathoanatomical diagnostic tests in patients with chronic LBP, the most frequent diagnosis (49%) was facet joint.[8] To diagnose the facet joint pain in clinic, the literatures suggested the uses of symptoms,[9] symptom changes with movement and activities,[10] local symptoms without nerve root pain,[11,12] and unilateral pain without referred pain lower than knee, and no symptom along dermatome or myotome.[13,14] The movements activating pain by increasing pressure on the joint and stretching the capsule are extension, rotation, and lateral flexion.[10,11] In rotation and extension, the pain mechanics is supposed as the inferior articular process slips on superior articular process which could activate nociceptors in the capsule.[15,16]

A study in 2007 presented the consensus of experts about the clinical features of facet joint pain.[17] The three round Delphi survey resulted in 12 indicators relevant to pathoanatomical mechanism of pain. The criteria included "(1) unilateral local pain, (2) activate pain with unilateral pressure on the joint or transverse process, (3) no nerve root pain, (4) less pain in flexion, (5) referred pain not beyond knee, (6) pressure on the joint found decreased range or increased resistance on the painful side, (7) muscle spasm same side to the facet, (8) pain in extension, (9) pain in extension with lateral flexion and rotation to the same side of facet, (10) the injection into the joint relief pain, (11) pain improves with fluoroscopically guided double-anesthetic blocks into the medial branch of the dorsal ramus which innervated the joints, and (12) could not diagnosed from X-ray".

Clinically, the physical therapists could evaluate the criteria 1 to 9 to determine whether the pain is caused by facet joint. Some studies attempted to validate clinical features of pain from facet joint origin showed controversial results.[18,19] However, these criteria are valuable in clinical reasoning and support the biomechanic evidences of pathology of the facet joints.[17] Use of these criteria for evaluating the patients would result in more homogenous of subjects and the manual therapy specifically on facet joints would be more effective. However, the reliability of using these criteria in clinic has not been reported. Therefore, this study aimed to test the agreement of physical therapists in using the criteria to examine and diagnose patients with facet joint problem causing LBP.

Methods

This study was a single-group, repeated-measures reliability study. The testing was conducted in a university physical therapy center. The study protocol was approved by the Ethic Committee of Mahidol University (MU-CIRB); protocol No. MU-IRB, COA. No. 2014. 033.2103, and Protocol No. MU-IRB 2014/006.0901. The data collection was undertaken from January to December 2015. The examination according to the criteria reported in the study of Wilde[17] included the interview of symptom behaviors (unilateral local pain, referred pain not beyond knee, and no nerve root pain); pain response in movement tests (less pain in flexion, pain in extension, and pain in extension with lateral flexion and rotation to the same side of facet); and manual tests (activate pain with unilateral pressure on the joint or transverse process, pressure on the joint found decreased range or increased resistance on the painful side, and muscle spasm same side to the facet).

The participants were consecutive patients aged 18–60 years old with LBP longer than three months. On the day of examination, the patient had pain measured by VAS 21–79 from 100 mm. Subjects were excluded if they presented with history of suspected spinal fracture or severe trauma; cauda equine syndrome with sensory impairment, leg weakness and incontinence; medical diagnosis of spondylolisthesis, foraminal or central stenosis, scoliosis or other spinal deformities; extended period of steroid use; taking pain medication within 24 h; pregnancy or menstruation.

The examiners were two physical therapists with clinical experiences in the musculoskeletal area of 20 and 9 years. They reviewed and practiced the testing procedure together before

beginning of subject recruitment. After history taking, the first examiner assessed the participants. After a brief rest, the second examiner, who did not see the examination process or knew the results, performed the assessment again. Both examiners also gave their impressions that the LBP was from facet joint origin or not. The subjects were asked to conceal the information of the first testing session. Each session took about 10 min. The examination order was randomized. The results of examination were analyzed for the agreements of each criteria and conclusion of facet joint diagnosis.

The sample size of subjects was estimated to minimize the standard error associated with the percent agreement between two arbitrary raters.[20] Setting the error margin at 15%, at least 44 patients were needed in the study. The percentage of agreement and generalized Kappa statistics were used to determine the agreement between therapists for each of the dichotomous scale items of the examination. The levels of Kappa statistics agreement were as follows: $0.00 < K < 0.20$ poor or slight agreement; $0.21 < K < 0.40$ fair; $0.41 < K < 0.60$ moderate; $0.61 < K < 0.80$ substantial or good; $0.81 < K < 1.00$ very good or almost perfect.[21]

Results

There were 45 patients with LBP, 32 females and 13 males, participating in this study. The average age of subjects was 33.64 ± 10.12 years. All subjects had chronic LBP with duration of symptom ranged from 3 months to 10 years. The pain intensity on the day of assessment was 4.47 ± 1.57.

The results of examination are presented in Table 1. The numbers of yes and no determined by each examiner were different for all items. However, the number of cases concluded to have or not to have facet lesion determined by two examiners was the same.

Table 2 shows the agreement levels of each examination item and the conclusion. All items had percent agreement greater than 70%. The highest level of percent agreement was the referred pain not beyond knee. The lowest percent agreements were reporting of pain location as unilateral and the palpation of muscle spasm on the same side of facet.

The Kappa coefficients showed statistically significant agreements between examiners for all examination items. There were different levels of agreement. Three items had fair agreement level, four had moderate agreement, and two had substantial agreement. The agreement of diagnosis if the patients had facet joint lesion was fair.

Discussion

To date, there are no specific clinical and radiographic indications of pain originated from facet joint.[22] A systematic review showed that the controlled anesthetic block of the facet or its nerve supply, the medial branch had good psychometric properties for diagnosis of LBP.[22] However, the test is invasive and not specific, therefore it is not

Table 1. Results of the examination of each criterion.

Examination items	Examiner 1		Examiner 2	
	Yes (%)	No (%)	Yes (%)	No (%)
(1) Unilateral local pain	23 (51.1)	22 (48.9)	29 (64.4)	16 (35.6)
(2) Referred pain not beyond knee	37 (82.2)	8 (17.81)	41 (91.1)	4 (8.9)
(3) No nerve root pain	36 (80.0)	9 (20.0)	39 (86.7)	6 (13.3)
(4) Less pain in flexion	20 (44.4)	25 (55.6)	23 (51.1)	22 (48.9)
(5) Pain in extension	34 (75.6)	11 (24.4)	38 (84.4)	7 (15.6)
(6) Pain in extension with lateral flexion and rotation to the same side of facet	33 (73.3)	12 (26.7)	36 (86.7)	6 (13.3)
(7) Muscle spasm same side to the facet	30 (66.7)	15 (33.3)	42 (93.3)	3 (6.7)
(8) Activated pain with unilateral pressure on the joint or transverse process	37 (82.2)	8 (17.8)	41 (91.1)	4 (8.9)
(9) Pressure on the joint found decreased range or increased resistance on the painful side	40 (88.9)	5 (11.1)	39 (86.7)	6 (13.3)
Conclusion of facet joint origin	38 (84.4)	7 (15.6)	38 (84.4)	7 (15.6)

Table 2. Percentage of agreement and Kappa Coefficient of criteria and diagnosis.

Examination items	% Agreement	Kappa coefficient	*p*-value[a]
(1) Unilateral local pain	73.3	0.436	0.001**
(2) Referred pain not beyond knee	91.1	0.622	< 0.001**
(3) No nerve root pain	80.0	0.286	0.048*
(4) Less pain in flexion	84.4	0.690	< 0.001**
(5) Pain in extension	86.7	0.588	< 0.001**
(6) Pain in extension with lateral flexion and rotation to the same side of facet	77.8	0.324	0.017*
(7) Muscle spasm same side to the facet	73.3	0.250	0.011*
(8) Activated pain with unilateral pressure on the joint or transverse process	86.7	0.433	0.002**
(9) Pressure on the joint found decreased range or increased resistance on the painful side	88.9	0.483	0.001**
Conclusion of facet joint origin	86.7	0.492	0.001**

[a]*p*-values for Kappa Statistics; *statistical significance at $p < 0.05$; **statistical significance at $p < 0.01$.

commonly performed even in orthopedics clinic. The criteria for diagnosing by symptom response and physical examination still needed for clinical use had face validity from the expert consensus in a Delphi study.[17]

For the clinical examination tested in this study, three criteria of symptom behaviors, including unilateral local pain, referred pain not beyond knee, and no nerve root pain, had various levels of agreements with Kappa coefficients ranged from 0.286 to 0.622 (% agreement from 73.3–91.1). There were three tests determining pain responses to movements, including less pain in flexion, pain in extension, and pain in extension with lateral flexion and rotation to the same side of facet. This examination domain also showed the same trend as the symptom behaviors with Kappa coefficients from 0.324 to 0.690 (% agreement from 77.8–86.7). The evidences of reliability of pain response to repeated movements were reportedly controversial in a systematic review.[23] However, based on the movement impairment classification system examined, Van Dillen *et al.* reported the very good level of agreements ($K > 0.75$) of physical examination items of symptom behaviors in patients with LBP.[24] In their study, the researchers established and defined the wording of questions and response choices before testing. However in our study, the interview and movement tests were performed in the same manner as routinely done in the physical therapy clinic which did not control the way to ask or test patients. With this manner, the answers and responses from each patient might result in different interpretations by different examiners. The interview and movement tests might be lacking of definition and training consideration, since they are usually supposed as basic clinical skill of therapists.

There were three examination items which needed tactile determination in this study including muscle spasm same side to the facet, activated pain with unilateral pressure on the joint or transverse process, and pressure on the joint found decreased range or increased resistance on the painful side. These items had low to fair agreements with Kappa coefficients of 0.250 to 0.483 (% agreement from 73.3–88.9). Systematic reviews reported generally low reliability of palpation-based assessment.[25–27] There were conflicting evidences about the reliability of evaluating muscle tension or spasm as well as the intersegmental stiffness of vertebral disc.[24,27–29] Strong evidence indicating low reliability pain on palpation and trigger points was also reported. Consistent results showed that the judgments based on visual and tactile information were usually difficult to perform reliably.[25–29]

The agreement for the pressure on the joint found decreased range or increased resistance on the painful side was moderate in this study. However, multiple reviews have suggested low agreement of this manual assessment.[27–29] The better agreement shown in this study might be due to both the subject and examiner characteristics. The patients recruited in this study had chronic symptom duration longer than three months with

moderate pain scale which might be related to the changes of spinal resistance. The examiners might have similar manual technique and interpretation of pressure on joint since they graduated from the same physical therapy school, have been working in the same clinical setting, and have several chances to discuss and practice in terms of manual therapeutic procedure together. They also reviewed and practiced the testing procedure together before beginning of subject recruitment in this study. The three-point grading (hypomobility, normal, and hypermobility) which used in this study was also recommended in review by Wong *et al.*[27]

The agreement level of conclusion was moderate in this study. The review of agreement levels of different diagnosis systems used in physical therapy clinics reported variability of agreements depended on the methodology and definitions used in the study.[24] Since the physical examination in physical therapy clinic depends largely on communication, manual skills and judgments of examiners, the explicitly defined techniques, operational definition, and consistent training would be necessary for improving reliability.[25]

The major limitation of result interpretation of this study is due to the statistics used. The results of low Kappa coefficient values in this study were partly due to the small number of some response category results from characteristics of the study sample. This would result in skewed response distribution and effect on Kappa statistics.[30] More studies which used greater variety of symptom and examination responses would be needed to confirm the agreement of therapists.

In addition, the study to test validity of the criteria set is warrant. The construct validity examination using factor analysis would result in the known redundant items which guide to more concrete set of examination instrument.

Conclusion

There was adequate reliability between two examiners showed by the percent agreements greater than 70% for all items used for confirming facet joint lesion. The test protocol of all assessment items was reviewed and practiced together by both examiners. However, the words of questions in history taking part and the manual techniques used were assumed to be routine practice in clinic, therefore these issues were not standardized which might be the source of disagreement in patient responses. However, due to skewness of symptom response, the Kappa coefficients were only low to moderate in this study. The items of "muscle spasm same side to the facet", "no nerve root pain", and "pain in extension with lateral flexion and rotation to the same side of facet" which had Kappa coefficients less than 0.40 (fair level of agreement) might need special cautions when performed and interpreted for facet diagnosis. Further study with variety of sign and symptoms of LBP would be beneficial to confirm the reliability of physical examination of facet joint in physical therapy clinic. The validity study compared with standard tests, i.e., nerve block and intra-articular injection would also verify these criteria of diagnosis. The reliability study assessing therapists with different clinical experiences would be valuable to prove clinical practicality of the testing protocol. Also, the study using reliable instrumental spinal stiffness measurements[27,31] might add clinical insight to this specific lesion condition.

Acknowledgments

The authors would like to thank all participants for their cooperation. The physical therapy staff of the Physical Therapy Center also provided a great help in the subject recruitment process.

Conflict of Interest

The authors have no conflicts of interest relevant to this paper.

Funding/Support

No financial or material support of any kind was received for the work described in this paper.

Author Contributions

Vongsirinavarat M took part in study conception and design, data collection and analysis, as well as writing and revision of the manuscript. Wahyuddin W was involved in study design and manuscript revision. Adisaiphaopun R performed subject screening and data collection. All authors have given final approval of the version to be published.

References

1. Mohseni MA, Stephenson BR, Richardson B. Spinal manipulation in the treatment of low back pain: A review of the literature with particular emphasis on randomised controlled clinical trials. Phys Ther Rev 1998;3:185–94.
2. Kuslich SD, Ulstrom CL, Michael CJ. The tissue origin of low back pain and sciatica: A report of pain response to tissue stimulation during operation on the lumbar spine using local anesthesia. Orthop Clin North Am 1991;22:181–7.
3. Bogduk N. Psychology and low back pain. Int J Osteopath Med 2006;9:49–53.
4. Waddell G. The Back Pain Revolution. Edinburgh: Churchill Livingstone, 2004.
5. Manchikanti L, Pampati V, Fellows B, Ghafoor A. The inability of the clinical picture to characterize pain from facet joints. Pain Physician 2000; 3:158–66.
6. Manchikanti L, Kaye AD, Boswell MV, et al. A systematic review and best evidence synthesis of effectiveness of therapeutic facet joint interventions in managing chronic spinal pain. Pain Physician 2015;18:E535–82.
7. Suri P, Hunter DJ, Rainville J, Guermazi A, Katz JN. Presence and extent of severe facet joint osteoarthritis are associated with back pain in older adults. Osteoarthritis Cartilage 2013;21 (9):1199–206.
8. Flavell CA, Gordon S, Marshman L. Classification characteristics of a chronic low back pain population using a combined McKenzie and patho-anatomical assessment. Man Ther 2016; 26:201–7.
9. Manchikanti L, Hirsch JA, Falco FJE, Boswell MV. Management of lumbar zygapophysial (facet) joint pain. World J Orthop 2016;7(5): 315–37.
10. Manchikanti L, Pampati V, Fellows B, et al. Influence of psychological factors on the ability to diagnose chronic low back pain of facet joint origin. Pain Physician 2001;4(4):349–57.
11. Sahrmann S. Diagnosis and Treatment of Movement Impairment Syndromes. 1st ed. St. Louis, MO: Mosby, 2001.
12. Eubanks JD, Lee MJ, Cassinelli E. Prevalence of lumbar facet arthrosis and its relationship to age, sex, and race: An anatomic study of cadaveric specimens. Spine 2007;32:2058–62.
13. Kalichman L, Li L, Kim DH, et al. Facet joint osteoarthritis and low back pain in the community-based population. Spine 2008;33:2560–5.
14. Kirkadly-Willis WH. The relationship of structural pathology to nerve root. Spine 1984; 9:49–52.
15. Starkey C, Brown SD, Ryan JL. Examination of Orthopedic and Athletic Injuries. 3rd ed. Philadelphia, PA: F.A. Davis Company, 2010.
16. Hresko MT. Thoracic and lumbosacral spine. In: Steinberg G, ed. Orthopaedics in Primary Care. 2nd ed. Baltimore: Williams & Wilkins, 1992.
17. Wilde VE, Ford JJ, McMeeken JM. Indicators of lumbar zygapophyseal joint pain: Survey of an expert panel with the Delphi technique. Phys Ther 2007;87(10):1348–61.
18. Schwarzer AC, Aprill CN, Derby R, Fortin J, Kine G, Bogduk N. Clinical features of patients with pain stemming from the lumbar zygapophysial joints. Is the lumbar facet syndrome a clinical entity? Spine 1994;19:1132–7.
19. Pang WW, Mok MS, Lin ML, Chang DP, Hwang MH. Application of spinal pain mapping in the diagnosis of low back pain — analysis of 104 cases. Acta Anaesthesiol Sin 1998;36:71–4.
20. Gwet K. Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters, 3rd ed. Maryland, USA: Advanced Analytics, LLC, 2012.
21. Landis JR and Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
22. Sehgal N, Dunbar EE, Shah RV, Colson J. Systematic review of diagnostic utility of facet (zygapophysial) joint injections in chronic spinal pain: An update. Pain Physician 2007;10(1): 213–28.
23. May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: A systematic review. Aust J Physiother 2006;52:91–102.
24. Van Dillen LR, Sahrmann SA, Norton BJ, et al. Reliability of physical examination items used for classification of patients with low back pain. Phys Ther 1998;78:979–88.
25. Maher CG, Adams R. Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. Phys Ther 1994;74(9):801–9.
26. Strender LE, Sjöblom A, Sundell K, Ludwig R, Taube A. Interexaminer reliability in physical examination of patients with low back pain. Spine 1997;22(7):814–20.
27. Wong A, Kawchuk G. The clinical value of lumbar posterioanterior segmental stiffness: A narrative review. 2017 PM&R 2017.
28. Seffinger MA, Najm WI, Mishra SI, et al. Reliability of spinal palpation for diagnosis of back and neck pain: A systematic review of the literature. Spine 2004;29:E413–25.
29. Snodgrass SJ, Haskins R, Rivett DA. A structured review of spinal stiffness as a kinesiological outcome

of manipulation: Its measurement and utility in diagnosis, prognosis and treatment decision-making. J Electromyogr Kinesiol 2012;22:708–23.
30. Lantz CA. Application and evaluation of the kappa statistic in the design and interpretation of chiropractic clinical research. J Manipulative Physiol Ther 1977;20:521–8.
31. Wong AY, Kawchuk G, Parent E, Prasad N. Within- and between-day reliability of spinal stiffness measurements obtained using a computer-controlled mechanical indenter in individuals with and without low back pain. Man Ther 2013;18 (5):395–402.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) 133–139
DOI: 10.1142/S1013702518500117

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

Side-to-side elbow range of movement variability in an ulnar neurodynamic test sequence variant in asymptomatic people

Michelle Meng Yim Tong, Vincent Cheng-Hsin Liu and Toby Hall*

School of Physiotherapy and Exercise Science, Curtin University of Technology, Perth, Western Australia 6102, Australia

**halltm@netspace.net.au*

Received 25 March 2017; Accepted 28 July 2017; Published 14 August 2018

Background: Range of motion (ROM) asymmetry between sides is one indicator of a positive neurodynamic test, but this has been less well studied for the ulnar nerve.
Objective: The purpose of this study was to investigate side-to-side variation in elbow ROM during an ulnar neurodynamic test sequence, including contralateral cervical side flexion, in 40 asymptomatic subjects.
Methods: A traditional goniometer was used to measure elbow flexion ROM at two end points, onset of resistance ($R1$) and symptom onset ($P1$). Two repeated measures of $R1$ and $P1$ were taken on each side.
Results: Reliability for $R1$ and $P1$ was found to be good (ICC ≥ 0.83, SEM ≤ 5.37) with no significant difference in mean ROM between sides. A significant relationship between sides was seen (r values ≥ 0.48) and R^2 values > 0.23; this indicates at least 23% of the variance observed in one limb was accounted for by range in the opposite limb. This relationship was slightly stronger for $R1$ than $P1$. Lower bound scores indicate that intra-individual ROM difference $> 23^\circ$ for $R1$ and 22° for $P1$ would exceed normal ROM asymmetry.
Conclusion: These findings provide clinicians with background information of ROM asymmetry during the ulnar neurodynamic test.

Keywords: Ulnar Neurodynamic test; upper limb; variability; reliability.

*Corresponding author.

Introduction

Increased neural tissue mechanosensitivity evaluated by neurodynamic tests[1,2] is frequently reported during the examination of patients with musculoskeletal disorders.[3,4] In particular, the ulnar nerve neurodynamic test is recommended in the examination of cubital tunnel syndrome,[5,6] thoracic outlet syndrome,[7] and C8 nerve root radiculopathy.[8]

To define a positive neurodynamic test, the following have been recommended. Firstly, the patient's symptoms must be reproduced and sensitizing maneuvres must increase or decrease symptoms.[2,3,8–10] Secondly, there should be a discrepancy in joint range between sides.[2] Finally, increased resistance is perceived by the examiner on the side of symptoms.[2,3,11] Side-to-side discrepancy in ROM and reproduction of symptoms are considered the most essential criteria for interpretation of neurodynamic tests,[2,12] and also useful comparable signs to evaluate treatment.[13]

Elbow ROM is frequently used as an outcome measure in studies investigating upper limb neurodynamic tests, due to the ease of side-to-side comparison.[1,14,15] Various studies show that asymmetry in elbow ROM between sides is common during neurodynamic testing in asymptomatic people.[16–18] Within-person side-to-side variability for the median nerve was reported between 15.5° and 27°.[16,18] Within-person side-to-side variability for the radial nerve was reported between 11.2° and 20°.[16,18] Similarly, between limb values of elbow ROM for the ulnar nerve was 21°.[16] The difference in values for elbow ROM reported by Covill and Petersen[16] and Stalioraitis *et al.*[18] may be due to different testing sequences. Specifically, the addition of contralateral cervical side flexion in the latter study reduced variability ROM between sides, potentially increasing the probability that ROM can be used to determine a positive neurodynamic test.[18] Structural differentiation to determine if symptom provocation is neurogenic in origin for the case of upper limb neurodynamic tests is determined by assessing the effect of adding contralateral cervical side flexion. Hence, it is important to know side-to-side variation in ROM during neurodynamic tests during different neurodynamic test variants, including cervical side flexion, as this can provide the clinician with an expectation of what could potentially be normal variance.

The purpose of this study was to investigate side-to-side variation in elbow ROM during the ulnar neurodynamic test sequence, using the sequence described by Hall and Elvey[4] in asymptomatic people. Two end points were investigated: $P1$, perceived by the subject at onset of the discomfort, and $R1$, determined by examiner, onset of resistance. The results of this study should provide clinician with background information regarding elbow ROM variability for the ulnar neurodynamic test in normal subjects, which may enable the determination of a positive test in symptomatic people.

Methods

Study design

A within-subject comparative measurement design was used to identify differences between sides during the ulnar neurodynamic test in asymptomatic people. The objective of this study was to determine the minimum side-to-side elbow ROM asymmetry required to classify an abnormal response to this specific test.

Participants

Forty asymptomatic subjects (19 females and 21 males, mean age 30.14 years) were included in this study. Participants were excluded if they had a current or previous history of trauma to the cervical spine, thoracic spine, shoulder, elbow, wrist, or hand. They were also excluded if they had any limitation of ROM in the upper quadrant. All participants underwent pre-test screening to ensure that they had pain-free and normal range of upper limb joint movement. This study received approval from Curtin University Human Research Ethics Committee. All participants were provided with information and gave informed consent. Using the two-tailed paired t-test, with an alpha level 0.05, power of 0.8, and a medium effect size of 0.5, 34 subjects were calculated to be needed for this study.

Equipment and measurements

The independent variable was side (left or right). The dependent variable was range of elbow flexion. Extraneous variables include body mass index, age, gender, and hand-dominant side. A traditional

goniometer was used to record elbow ROM. Acceptable validity for measurement of elbow ROM using a traditional goniometer when compared with radiograph measurement has been reported.[19] In that study, the intra-class correlation coefficient (ICC) ranged from 0.94 to 0.97 for the goniometric measurements and from 0.98 to 0.99 for the radiographic measurements.[19] The two methods correlated and the maximum error of the goniometric measurement was 7.0° for flexion, 95% of the time.[19]

Procedure

Participants were tested according to a standard clinical testing protocol, without fixation devices. The untested limb was placed in a relaxed position with the hand resting on the abdomen. The cervical spine was placed in contralateral side flexion to the side tested, without rotation. The shoulder girdle on the tested side was held in neutral elevation and/depression position manually by the examining therapist to mimic the clinical situation.

The ulnar neurodynamic test sequence was tested on each side in random order. The participants underwent one familiarization trial. Two measurements of range of elbow movement were recorded after the familiarization trial using the traditional goniometer by a separate independent researcher while the main researcher maintained the arm position during the measurement process. The goniometer was not visible to the main researcher to avoid bias. The goniometer axis was aligned with the medial epicondyle, with the proximal arm aligned with the midline of humerus and the distal arm aligned with the line formed by the medial epicondyle and radial styloid process.

Good intra-tester reliability of goniometric measurement has been shown when the mean of two or three measurements is taken,[20–22] hence only two measurements were taken for each end point and each test. The end points were $R1$ and $P1$, as these have been shown to have excellent inter- and intra-rater reliability.[21,23,24] Participants were instructed to say "now" upon the onset of any sensation change during the neurodynamic test and the movement paused for measurement purposes. The examiner said "R" when the onset of resistance was felt and again the movement was paused for measurement.

Neurodynamic test sequence for the Ulnar nerve

The tested arm was positioned in 90° shoulder abduction, 90° shoulder external rotation with the shoulder girdle maintained in neutral. The cervical spine was placed in maximal lateral flexion to the contralateral side.[4] The elbow was fully extended, with the forearm in maximum pronation, and wrist/fingers maximally extended. Elbow flexion was initiated, and at $P1$ and $R1$, elbow movement was paused while ROM recorded. The neurodynamic test continued until both end points had been achieved. The subjects were given one familiarization trial on each side followed by two trials where measurements were recorded in between a 10-second rest interval.

Data Analysis

Data analysis was carried out using SPSS v19. (SPSS Inc., 444 N. Michigan Avenue, Chicago, Illinois, 60611). All data were normally distributed. Intra-tester reliability for repeated measures on each arm was calculated using ICC (2,1), standard error of the measurement (SEM), and minimal detectable change (MDC). Mean elbow ROM and standard deviation was determined for the Ulnar neurodynamic test sequence for both arms. Dependent t-tests were used to compare within-subject range of motion (ROM) between the right and left arms for each test. Relationship in ROM between limbs was calculated using the Pearson correlation coefficient and coefficient of determination (r^2). The mean absolute values (MAVs) were calculated to determine differences between limbs while ensuring that all values remain positive. A lower bound score was used to determine the cut-off point at which the degree of difference between limbs could be considered greater than that accounted for by measurement error and variability. This was carried out according to the method reported in another study by multiplying the standard deviation of the MAV by the z-score (1.65) of a one-tailed t-test ($\alpha = 0.05$) and adding the MAV (lower bound score = (SD) (1.65) + MAV).[16]

Results

All data were checked and found to be normally distributed. The results for intra-therapist reliability

Table 1. Reliability statistics for elbow ROM for ulnar neurodynamic test ($n = 40$).

Measurement	ICC [2,1] (95% CI)	SEM°	MDC°
Right Ulnar $R1$	0.83 (0.67, 0.91)	5.4	14.9
Right Ulnar $P1$	0.84 (0.69, 0.91)	4.7	13.0
Left Ulnar $R1$	0.90 (0.81, 0.95)	3.7	10.1
Left Ulnar $P1$	0.90 (0.81, 0.95)	4.2	11.6

are shown in Table 1. For both $R1$ and $P1$, the range recorded during Ulnar neurodynamic test, ICC (2,1) values was greater than 0.83 indicating good reliability.[16] In addition, the SEM and MDC for each assessment point were also relatively small.

Means and standard deviations for elbow ROM during the ulnar neurodynamic test are presented in Table 2. The mean difference between the left and right sides, for both $R1$ and $P1$, was very small, with at most 1.5° between sides. At any assessment point, there was no significant difference between the left and right sides as reflected by the 95% confidence intervals (Table 2).

A Pearson correlation analysis revealed a significant relationship between the limbs, with r values greater than 0.48. Furthermore, the R^2 values were greater than 0.23, indicating that at least 23% of the variance observed in one limb was accounted for by range in the opposite limb. This relationship was slightly stronger for $R1$ than $P1$. These data point to a relationship for elbow ROM between limbs, indicating that elbow ROM of one side can be used to some degree to predict elbow ROM of the opposite limb.

The MAV and lower bound scores shown in Table 3 revealed some degree of variability between the right and left limbs for any assessment point. Elbow ranges recorded at $R1$ had slightly more variability between limbs than $P1$ during the ulnar neurodynamic test.

Discussion

This study investigated side-to-side variation in elbow ROM at $P1$ and $R1$ for the ulnar neurodynamic test sequence as described by Hall and Elvey.[4] Small mean differences were detected between sides for $R1$ (1.6°) and $P1$ (1.1°). However, despite these small mean differences, this did not equate to a strong correlation between sides as seen in Table 2. This might be explained by the relatively large MAVs for ROM differences between limbs, which indicate large intra-individual differences in ROM between limbs as shown in Table 3.

The MAVs for discrepancy between sides were similar in order of magnitude for both $R1$ and $P1$. The lower bound scores were calculated from the MAV, and indicate that elbow ROM difference between limbs for the ulnar neurodynamic test must be greater than 23° for $R1$ and 22° for $P1$ for the ROM findings to be considered relevant beyond normal variation and measurement error. These findings were similar to those reported by Covill and Petersen,[16] who had also investigated the Ulnar neurodynamic test. In that study, the MAV was 6.1° and lower bound score was 20.9°. Small differences in MAVs for ROM between Covill and Petersen[16] and the current study were likely to be attributed to differences in end-point measurement, type of goniometer used, and variation in neurodynamic test sequence. Also, $P2$ and $R2$ were the end points measured by Covill and Petersen,[16] while in the present study, $P1$ and $R1$ were recorded instead. To the best of our knowledge, no other study has reported lower bound scores for the ulnar neurodynamic test.

Table 2. Mean range, mean differences between left and right sides (SD) with 95% confidence interval (CI), Pearson correlation coefficient (r), and coefficient of determination (R^2) ($n = 40$).

Measurement	Mean range (SD)		Mean difference scores (95% CI)°	r	R^2
	Left	Right			
Ulnar $R1$	110.8 (1.6)	112.4 (13.0)	1.6 (−2.5, 5.6)	0.53 $p < 0.001$	0.28
Ulnar $P1$	110.8 (13.2)	111.9 (11.8)	1.1 (−2.8, 5.0)	0.48 $p < 0.001$	0.23

Table 3. Mean absolute differences (MAV) in elbow ROM between right and left sides together with lower bound scores for neurodynamic testing ($n = 40$).

Measurement	MAV (SD)°	Lower bound scores°
Ulnar *R*1	9.7 (8.1)	23.1
Ulnar *P*1	9.0 (8.2)	22.5

MAVs of 10.1° and 6.7° for the median and radial neurodynamic tests were also reported by Covill and Petersen,[16] which are higher than that reported by Stalioraitis *et al.*[18] where ROM values were 5.5° for *R*1 and 5.8° for *P*1 for the median nerve and 4.2° for *R*1 and 4.8° for *P*1 for the radial nerve. As such, it might be expected that the MAV for the ulnar nerve reported by Covill and Petersen[16] would be greater than those in the current study, but the reverse was seen. One explanation for these differences between the three studies might be in the type of measurement device which in the current study was a traditional goniometer. In contrast, an electrogoniometer has often been used in the previous research.[16,18] Additionally, the type of neurodynamic sequence might also affect the difference in MAV. A different test sequence that did not include the contralateral cervical side flexion was used by Covill and Petersen.[16] Cervical side flexion is an important component of neurodynamic testing used in structural differentiation which increases strain on the nervous system without differing the mechanical load on the musculoskeletal system.[23] The effect of including contralateral cervical side flexion in the median neurodynamic test sequence was to reduce mean elbow ROM to 132.8° in asymptomatic subjects whereas the same sequence without contralateral cervical side flexion had achieved elbow ROM of 149°.[23]

Mean ROM difference between sides for *R*1 and *P*1 was noted to be similar to Stalioraitis *et al.*[18] with the exception of *R*1 for the median nerve which had a mean difference of only 0.9°. This similarity may be attributed to the consistency in neurodynamic test sequence used in both studies.

The mean elbow ROM values were smaller in this study than those reported by Covill and Petersen[12] for the ulnar neurodynamic test. This could be explained by the various differences in methodology used by the two studies, most notably, the different end-point measures. As observed by Vanti *et al.*,[21] for the median neurodynamic test, mean elbow ROM was 155° for *P*1 and 164° for *P*2. As such, the use of *P*2 as an end-point measure might explain the greater mean elbow ROM reported by Covill and Petersen.[16] There are differences of opinion as to the use of end-point measure with *P*1 being more clinically relevant as it is included in the definition of a positive neurodynamic test to reproduce the patient's symptoms[3,8] without bringing on greater pain.

The correlation (R^2) in elbow ROM between sides was higher being 0.23 for *R*1 and *P*1 in this study compared to 0.13 reported by Covill and Petersen.[16] Although values varied slightly, the clinical interpretation of these small R^2 values was that it may not be possible to use elbow ROM comparison between sides as one of the criteria to support a positive ulnar neurodynamic test. In other words, the range in one limb accounts for only a small proportion of the predicted range in the opposite limb during this neurodynamic test. This would indicate that other criteria should be used to identify a positive neurodynamic test, with ROM difference between limbs being of minor importance.

Intra-tester reliability was found to be good for the left and right sides of ulnar neurodynamic test (Table 1). These findings are consistent with other studies reporting reliability of other neurodynamic tests when measuring elbow ROM with an electrogoniometer.[18,21,24] Neither the left nor right side showed any indication of substantially better intra-tester reliability. Hence, a possible conclusion would be the degree of intra-tester reliability which is not related to the side measured and it would suggest that future neurodynamic test studies should include intra-tester reliability within every study.

The use of traditional goniometer has been recommended to depict a realistic clinical situation.[16] Despite this, a traditional goniometer is clumsy to use, as the therapist must control many components during the ulnar neurodynamic test. Trying to correctly place the goniometer during the neurodynamic testing process greatly adds to this complexity and makes it difficult to accurately measure ROM. Future studies should look for simpler methods of measuring ROM that can be easily applied and read by the clinician.

Conclusion

The results of this study provide clinicians with baseline knowledge of normal ROM variation

between sides during the ulnar neurodynamic test using a commonly used measurement device. The lower bound score of 23° for measurement at $R1$ and 22° for $P1$ would suggest that side-to-side variation of more than 23° would exceed normal variability and would likely not be due to measurement errors and is therefore clinically relevant. Large variation in ROM between sides indicates that ROM is less helpful in determining a positive neurodynamic test than other test criteria. One explanation for large side-to-side variation in ROM is the cumbersome nature of using a traditional goniometer during neurodynamic testing.

Conflict of Interest

The authors declare that there is no conflict of interest relevant to the study.

Funding/Support

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors of this study would like to acknowledge Karen Lam for her contributions to this study.

Author Contributions

Study design was initiated by Dr. Toby Hall. Data Collection was collected by Ms. Michelle Tong with the assistance of Mr. Vincent Liu. Michelle and Vincent had written the majority of this paper under the supervision of Dr. Toby Hall. Data Analysis was done by Dr. Toby Hall. Subsequent revisions of the drafts made for publication were done by Michelle under the supervision of Dr. Toby Hall. Each author has reviewed and approved the submission of this updated draft of the manuscript and takes full responsibility for the manuscript.

References

1. Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner RS, Whitman JM. Neck pain: Clinical practice guidelines linked to the International Classification of Functioning, Disability, and Health from the Orthopedic Section of the American Physical Therapy Association. J Orthop Sports Phys Ther 2008;38(9):A1–34. doi:10.2519/jospt.2008.0303.
2. Nee RJ, Jull GA, Vicenzino B, Coppieters M. The validity of upper-limb neurodynamic tests for detecting peripheral neuropathic pain. J Orthop Sports Phys Ther 2012;42(5):413–24. doi:10.2519/jospt.2012.3988.
3. Butler D. The Sensitive Nervous System. 1st ed. Adelaide: Noigroup Publications, 2000.
4. Hall T, Elvey RL. Neural tissue evaluation and treatment. In: Donatelli RA, ed. Physical Therapy of the Shoulder. 5the ed. Chap. 6. Saint Louis: Churchill Livingstone, 2012:131–45.
5. Coppieters M, Bartholomeeusen K, Stappaerts K. Incorporating nerve-gliding techniques in the conservative treatment of cubital tunnel syndrome. J Manipulative Physiol Ther 2004;27(9):560–68. doi:10.1016/j.jmpt.2004.10.006.
6. Goyal M, Mehta S, Rana N, et al. Motor nerve conduction velocity and function in carpal tunnel syndrome following neural mobilization: A randomized clinical trial. Int J Health Allied Sci 2016;5(2):104–10. doi:10.4103/2278-344x.180434.
7. Sanders RJ, Hammond SL, Rao NM. Diagnosis of thoracic outlet syndrome. J Vasc Surg 2007;46(3):601–4. doi:10.1016/j.jvs.2007.04.050.
8. Shacklock MO. Clinical Neurodynamics: A New System of Musculoskeletal Treatment. Edinburgh: Elsevier Butterworth-Heinemann.
9. Manvell JJ, Manvell N, Snodgrass SJ, Reid SA. Improving the radial nerve neurodynamic test: An observation of tension of the radial, median and ulnar nerves during upper limb positioning. Man Ther 2015;20(6):790–6. doi:10.1016/j.math.2015.03.007.
10. Mintken P, Puentedura E, Louw A. Neurodynamic interventions and physiological effects: Clinical neurodynamics in neck and upper extremity pain. In: de las Peñas CF, Cleland JA, Huijbregts PA, eds. Neck and Arm Pain Syndrome. Edinburgh: Churchill Livingstone, 2011:496–515.
11. Hall T, Elvey RL, Davies N, Dutton L, Moog M. Efficacy of manipulative physiotherapy for the treatment of cervicobrachial pain. In: Tenth Biennial Conf MPAA. Melbourne: Manipulative Physiotherapists Association of Australia, 1997.
12. Butler D, Gifford L. The concept of adverse mechanical tension in the nervous system Part 1: Testing for "Dural tension". Physiotherapy 1989;75(11):622–9. doi:10.1016/s0031-9406(10)62374-7.
13. Vicenzino B, Collins D, Wright A. The initial effects of a cervical spine manipulative physiotherapy treatment on the pain and dysfunction of lateral epicondylalgia. Pain 1996;68(1):69–74.
14. Nee RJ, Vicenzino B, Jull GA, Cleland JA, Coppieters MW. Neural tissue management provides

immediate clinically relevant benefits without harmful effects for patients with nerve-related neck and arm pain: A randomised trial. J Physiother 2012;58(1):23–31. doi:10.1016/s1836-9553(12)70069-3.
15. Wainner RS, Fritz JM, Irrgang JJ, Delitto A, Allison S, Boninger ML. Development of a clinical prediction rule for the diagnosis of carpal tunnel syndrome. Arch Phys Med Rehabil 2005;86(4):609–18. doi:10.1016/j.apmr.2004.11.008.
16. Covill LG, Petersen SM. Upper extremity neurodynamic tests: Range of motion asymmetry may not indicate impairment. Physiother Theory Pract 2012;28(7):535–41. doi:10.3109/09593985.2011.641198.
17. Martinez MD, Cubas CL, Girbes EL. Ulnar nerve neurodynamic test: Study of the normal sensory response in asymptomatic individuals. J Orthop Sports Phys Ther 2014;44(6):450–6. doi:10.2519/jospt.2014.5207.
18. Stalioraitis V, Robinson K, Hall T. Side-to-side range of movement variability in variants of the median and radial neurodynamic test sequences in asymptomatic people. Man Ther 2014;19(4): 338–42. doi:http://dx.doi.org/10.1016/j.math. 2014.03.005.
19. Chapleau J, Canet F, Petit Y, Laflamme GY, Rouleau DM. Validity of goniometric elbow measurements: Comparative study with a radiographic method. Clin Orthop Relat Res 2011;469(11):3134–40. doi:10.1007/s11999-011-1986-8.
20. Lohkamp M, Small K. Normal response to upper limb neurodynamic test 1 and 2A. Man Ther 2011;16(2):125–30. doi:10.1016/j.math.2010.07.008.
21. Vanti C, Bonfiglioli R, Calabrese M, Marinelli F, Violante FS, Pillastrini P. Relationship between interpretation and accuracy of the upper limb neurodynamic test 1 in carpal tunnel syndrome. J Manipulative Physiol Ther 2012;35(1):54–63. doi:10.1016/j.jmpt.2011.09.008.
22. Yaxley GA, Jull GA. A modified upper limb tension test: An investigation of responses in normal subjects. Aust J Physiother 1991;37(3):143–52. doi:10.1016/S0004-9514(14)60536-5.
23. Coppieters M, Stappaerts K, Janssens K, Jull G. Reliability of detecting 'onset of pain' and 'submaximal pain' during neural provocation testing of the upper quadrant. Physiother Res Int 2002;7(3):146–56.
24. Oliver GS, Rushton A. A study to explore the reliability and precision of intra and inter-rater measures of ULNT1 on an asymptomatic population. Man Ther 2011;16(2):203–6. doi:10.1016/j.math.2010.05.009.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) 141–147
DOI: 10.1142/S1013702518500129

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

Stratification of stroke rehabilitation: Five-year profiles of functional outcomes

Bryan Ping Ho Chung

Physiotherapy Department, Tai Po Hospital
Tai Po, New Territories, Hong Kong

taipobryan@yahoo.com

Received 6 June 2017; Accepted 13 September 2017; Published 14 August 2018

Background: Stroke rehabilitation in inpatient setting requires high intensity of manpower and resources. Early stratification of patients with stroke could facilitate early discharge plan and reduce avoidable length of stay (LOS) in hospital. Stratification of patients with stroke in clinical setting is usually based on functional scores which are quite time-consuming and require a special training to complete the full score.
Objective: The objective of the study was to explore whether Modified Functional Ambulation Category (MFAC) can serve as a stratification tool of patients with stroke in inpatient rehabilitation.
Methods: This was a retrospective, descriptive study of the demographic, functional outcomes of patients with stroke in an inpatient rehabilitation center. A total of 2,722 patients completed a stroke rehabilitation program from 2011 to 2015 were recruited. The patients were divided into seven groups according to their admission MFAC. The between-group difference in LOS, functional outcomes at admission and discharge including Modified Rivermead Mobility Index (MRMI) and Modified Barthel Index (MBI) as well as MRMI gain, MRMI efficiency, MBI gain, and MBI efficiency were analyzed.
Results: Subjects with admission categories of MFAC 2 and 3 had a highly significant ($p < 0.001$) MRMI gain (6.2 and 6.6, respectively) and subjects with admission categories of MFAC 3 to 5 had highly significant ($P < 0.001$) MRMI efficiency (0.34, 0.40, and 0.39, respectively). The subjects with admission categories of MFAC 2 to 5 had a highly significant ($p < 0.001$) MBI gain (9.7, 10.2, 9.3, and 7.0, respectively) and the subjects with admission categories of MFAC 4 to 5 had a highly significant ($p < 0.001$) MBI efficiency (0.70 and 0.72, respectively). The subjects with admission categories of MFAC 1 and 2 had a highly significant ($p < 0.001$) LOS (27.7 and 26.6, respectively). MFAC profile was also established to represent the distribution of discharge MFAC of subjects according to their admission MFAC. The chance of subjects with admission categories of MFAC 1 and MFAC 2 progress to any kind of walker (MFAC > 2) is 12.7% and 58.2%, respectively. The chance of subjects with admission MFAC 3, MFAC 4 and MFAC 5 progress to independent walker (MFAC > 5) is 6.7%, 14.8%, and 50.3%, respectively. Both admission MFAC and

admission MBI had strong correlations with discharge MFAC ($r = 0.84$, $P < 0.0001$ and $r = 0.78$, $P < 0.0001$, respectively), discharge MRMI ($r = 0.82$, $P < 0.0001$ and $r = 0.78$, $P < 0.0001$, respectively) and discharge MBI ($r = 0.78$, $P < 0.0001$ and $r = 0.94$, $P < 0.0001$, respectively).
Conclusion: This study showed that patients on admission with moderate disability in term of MFAC had the greatest mobility gain and basic activities of daily living (ADL) gain from inpatient stroke rehabilitation. Admission MFAC could be a stratification tool of patients with stroke in inpatient rehabilitation.

Keywords: Stroke rehabilitation; physiotherapy; functional outcome; stratification.

Introduction

Stroke, also known as cerebrovascular accident (CVA), is an acute disturbance of focal or global cerebral function with signs and symptoms lasting more than 24 h or leading to death, presumably of vascular origin.[1] Stroke was the fourth leading cause of deaths in Hong Kong in 2014[2] and around 25,000 of patients with stroke are admitted to public hospitals under the Hong Kong Hospital Authority annually.[2] The most widely recognized impairment causing stroke is motor impairment which restricts functional mobility including walking.[3,4] Therefore, to improve functional mobility of patients with stroke is one of the main goals of rehabilitation.[4] However, stroke rehabilitation in inpatient setting requires high intensity of manpower and resources. Early stratification of patients with stroke is useful for clinicians to recognize patients' possible functional outcomes, level of disability and requirement for social support so as to facilitate early discharge plan and reduce avoidable length of stay (LOS) in hospital. Stratification in clinical setting is usually based on admission functional scores such as Functional Independence Measure (FIM^{TM}),[5] Barthel Index (BI),[6] Modified BI (MBI),[7] and classification system such as Inpatient Rehabilitation Facility-Case-Mix Group (IRF-CMG).[8] However, these assessment tools are quite time-consuming and require a special training to complete the full score. Using functional scores in terms of simple mobility categories to stratify patients with stroke and predict functional scores of patients with stroke at discharge is worth exploring. Modified Functional Ambulation Category (MFAC) is a 7-point Likert Scale[9] which is easy-to-use, inexpensive and commonly used to classify walking capacity of patients with stroke in Hong Kong[10,11] and Korea.[12] The objective of the study was to explore whether MFAC can serve as a stratification tool of patients with stroke in inpatient rehabilitation.

Methods

Demographic characteristics and average functional outcomes

This was a retrospective descriptive study of patients with principal diagnosis of CVA, stroke, or hemiplegia and had received stroke rehabilitation program in a hospital in Hong Kong between the periods of 1st January 2011 to 31st December 2015. Patient's demographic and hospital information including age, gender, premorbid accommodation, stroke type, days post stroke, discharge destination and LOS were retrieved from the database of physiotherapy department of the hospital and Clinical Management System of Hong Kong Hospital Authority. LOS is the total number of days spent in the rehabilitation program. The demographic characteristics and average functional outcomes of all subjects were demonstrated in number and percentage or mean and standard deviation.

Clinical outcomes

The subjects were divided into seven groups according to their admission MFAC. The between-group differences of clinical outcomes were analyzed. Clinical outcomes including LOS, admission Modified Rivermead Mobility Index (MRMI), discharge MRMI, MRMI gain, MRMI efficiency, admission MBI, discharge MBI, MBI gain, and MBI efficiency. MRMI gain is the difference between discharge MRMI and admission MRMI. MRMI efficiency is the average change in total MRMI ratings per day, which were calculated for each subject by subtracting admission MRMI from discharge MRMI ratings and then dividing by the LOS measured in days. MBI gain and MBI efficiency were measured by the same methods. The gain of a score indicates the total gain within the LOS. The efficiency of a score indicates the daily

gain of the score. Both gain and efficiency of scores were included to facilitate comparisons to other studies.

MFAC profile

MFAC profile represented the distribution of possibility of discharge MFAC by subjects' admission category of MFAC. The results were expressed in a matrix table and in a rank order for each group of subjects. The possibilities were expressed in percentage (%).

Correlation

Since previous studies showed that functional outcome at admission is positively correlated with the functional outcome at discharge,[13,14] the correlation of admission MFAC and admission MBI to discharge MFAC, discharge MRMI, and discharge MBI were measured.

Modified Functional Ambulation Category

MFAC is a 7-point Likert Scale (1–7) that is used to classify a subject's walking capacity. It was modified from Functional Ambulation Classification (FAC). FAC includes six ordinal categories (0–5) of support needed for gait, but does not evaluate whether or not an aid was used.[15] MFAC divided gait into seven categories (Category I to Categories VII), ranging from no ability to walk and requires manual assistance to sit or is unable to sit for 1 min without back or hand support (MFAC 1; Category: I, stage: Lyer) to the ability to walk independently on level and non-level surfaces, stairs, and inclines (MFAC 7; Category: VII, stage: Outdoor walker).[9] The inter-rater reliability of the MFAC (intraclass correlation coefficient (ICC)) was 0.982 (0.971–0.989), with a kappa coefficient of 0.923 and a consistency ratio of 94% for stroke patient[12] and the ICC of the MFAC in patients with hip fractures is 0.96, with a construct validity of $r = 0.81$ on the Elderly Mobility Scale (EMS).[9]

Modified Rivermead Mobility Index

MRMI was used to assess subjects' mobility in this study. The MRMI is highly reliable between raters (ICC = 0.98) and has high internal consistency (Cronbach's alpha = 0.93) to early stage patients with stroke. The MRMI consists of eight test items, including turning over, changing from lying to sitting, maintaining sitting balance, going from sitting to standing, standing, transferring, walking indoors, and climbing stairs. The score of MRMI ranges from 0 to 40. One main characteristic of the MRMI is that participants are scored by observation of their performance on the items directly.[16]

Modified Barthel Index

MBI was used to assess subjects' basic activities of daily living (ADL) in this study. MBI measures the subject's performance on 10 functional items including self-care, continence, and locomotion. The values assigned to each item are based on the amount of physical assistance required to perform the task and added to give a total score ranging from 0 to 100 (0 = fully dependent, 100 = fully independent) with higher scores indicating higher levels of physical function.[11] There are no subtotal score because there is no subscale.[11] The internal consistency reliability coefficient for MBI is 0.90.[17]

Statistical Analysis

In order to stratify subjects with greatest functional gain from the stroke rehabilitation program, the subjects were divided into seven groups according to the admission categories of MFAC. The between-group differences of the characteristics including LOS, admission MRMI, discharge MRMI, admission MBI and discharge MBI, MRMI gain, MRMI efficiency, MBI gain, and MBI efficiency were analyzed by one way analysis of variance (ANOVA), post-hoc Bonferroni's test was administered to identify subsets of each group. For each characteristic, the groups with relative high score and without within-group post-hoc difference were highlighted.

In order to compare the correlation of admission MFAC and admission MBI to functional outcomes of subjects, Spearman's rank correlation coefficient (r_s) was used to test the relationship between admission MFAC and admission MBI to discharge MFAC, discharge MRMI and discharge MBI. Results were considered statistically significant when $p < 0.05$. Data were analyzed with the use of the SPSS − V20 statistical package (SPSS Inc., Chicago, LL).

Ethics Statement

Ethics approval was granted by the Joint Chinese University of Hong Kong — New Territories East Cluster Clinical Research Ethics Committee.

Results

Demographic characteristics and average functional outcomes

Medical records of 2,722 of 3,085 subjects admitted to a rehabilitation hospital were analyzed for this study. A total of 363 subjects were excluded, in which, 187 subjects were transferred back to acute hospital due to unstable medical conditions, 18 subjects were discharged against medical advice (DAMA) and 158 subjects died. Of the 2,722 subjects, the mean age was 74.6 in which 1,433 (52.7%) subjects were male and 1,289 (47.3%) subjects were female. There were 2,333 (85.7%) subjects lived at home whereas 389 (14.3%) subjects were institutionalized. There were 2,312 (84.9%) subjects suffered from cerebral infarction and 410 (15.1%) subjects suffered from cerebral hemorrhage. A total of 1,955 (71.8%) subjects were first-time stroke patients and 767 (28.2%) subjects had recurrent stroke. The mean days post stroke was 12.5 days and the mean LOS in the rehabilitation program was 22.3 days (Table 1). Before the admission, 139 (5.1%) subjects were lyers (MFAC = 1), 168 (6.2%) subjects were sitters (MFAC = 2), 97 (3.6%) subjects were dependent walkers (MFAC = 3), 193 (7.0%) subjects were assisted walkers (MFAC = 4), 70 (2.6%) subjects were supervised walkers (MFAC = 5), 396 (14.3%) subjects were independent indoor walkers (MFAC = 6) and 1,665 (61.2%) subjects were independent outdoor walkers (MFAC = 7) (Table 1). The average MRMI was increased from 13.7 at admission to 18.1 at discharge. The average MRMI gain was 4.4 and the average MRMI efficiency was 0.26 (Table 1). The average MBI was increased from 32.5 at admission and 39.7 at discharge. The average MBI gain was 7.2 and the average MBI efficiency was 0.4 (Table 1).

Table 1. Demographic characteristic and average functional outcomes of patients ($N = 2,722$).

Demographic characteristic	N (%)	Mean (SD)
Age	—	74.8 (12.24)
Gender	—	
Male	1,433 (52.7)	—
Female	1,289 (47.3)	—
Premorbid MFAC		
Category 1	139 (5.1)	—
Category 2	168 (6.2)	—
Category 3	97 (3.6)	—
Category 4	193 (7.0)	—
Category 5	70 (2.6)	—
Category 6	390 (14.3)	—
Category 7	1,665 (61.2)	—
Premorbid accommodation		
Home	2,333 (85.7)	—
Institution	389 (14.3)	—
Stroke type (%)		
Cerebral infraction	2,312 (84.9)	—
Cerebral hemorrhage	410 (15.1)	—
First stroke	1,955 (71.8)	—
Recurrent stroke	767 (28.2)	
Days post stroke (day)	—	12.5 (7.65)
LOS in rehabilitation	—	22.3 (18.24)
Discharge destination from rehab	—	
Home	1,690 (62.1)	—
Institution	1,032 (37.9)	—
Admission MRMI	—	13.7 (11.07)
Discharge MRMI	—	18.1 (12.83)
MRMI gain	—	4.4 (6.14)
MRMI efficiency	—	0.26 (0.43)
Admission MBI	—	32.5 (29.19)
Discharge MBI	—	39.7 (32.33)
MBI gain	—	7.2 (12.00)
MBI efficiency	—	0.40 (0.77)

Clinical outcomes

The outcomes among the seven groups of subjects with different admission MFAC were shown in Table 2. On admission, 889 (32.7%) subjects were lyers (MFAC = 1), 546 (20.6%) subjects were sitters (MFAC = 2), 439 (16.1%) subjects were dependent walkers (MFAC = 3), 570 (20.9%) subjects were assisted walkers (MFAC = 4), 190 (7.0%) subjects were supervised walkers (MFAC = 5), 74 (2.7%) subjects were independent indoor walkers (MFAC = 6) and 14 (0.5%) subjects were independent outdoor walkers (MFAC = 7) (Table 2). The LOS of subjects with admission categories of MFAC 1 and 2 (27.7 and 26.6, respectively) had significant differences ($p < 0.05$) when compared to subjects with admission categories of MFAC 3 to 7 (21.8, 16.1, 10.9, 8.3, and 6.1, respectively). The MRMI gain of subjects with admission categories of MFAC 2 and 3 (6.2 and 6.6, respectively) had significant differences ($p < 0.05$) when compared to subjects with admission categories of MFAC 1 and 4 to 7

Table 2. Clinical outcomes of subjects.

Outcome	Admission MFAC							
	1 ($n = 889$)	2 ($n = 546$)	3 ($n = 439$)	4 ($n = 570$)	5 ($n = 190$)	6 ($n = 74$)	7 ($n = 14$)	*P*-value
LOS	27.7 (22.74)	26.6 (17.52)	21.8 (14.36)	16.1 (11.47)	10.9 (7.08)	8.3 (7.54)	6.1 (4.94)	$P < 0.001$*
Admission MRMI	2.0 (3.12)	10.2 (4.26)	15.7 (4.28)	23.4 (5.55)	30.5 (4.54)	37.2 (2.59)	37.1 (8.74)	$P < 0.001$*
Discharge MRMI	4.7 (7.05)	16.4 (8.37)	22.3 (7.96)	28.1 (6.71)	33.9 (4.91)	38.2 (2.46)	37.5 (7.42)	$P < 0.001$*
Admission MBI	5.6 (10.98)	25.9 (18.91)	41.3 (20.14)	54.3 (22.03)	71.6 (18.78)	80.6 (21.36)	82.3 (23.51)	$P < 0.001$*
Discharge MBI	8.6 (15.60)	35.6 (23.80)	51.5 (22.50)	63.6 (22.56)	78.6 (18.07)	84.3 (20.67)	85.5 (24.21)	$P < 0.001$*
MRMI gain	2.7 (5.91)	6.2 (6.97)	6.6 (6.86)	4.8 (4.91)	3.4 (3.90)	1.0 (1.69)	0.4 (1.34)	$P < 0.001$*
MRMI efficiency	0.11 (0.35)	0.26 (0.34)	0.34 (0.41)	0.40 (0.53)	0.39 (0.51)	0.15 (0.26)	0.05 (0.17)	$P < 0.001$*
MBI gain	3.0 (8.94)	9.7 (14.84)	10.2 (12.81)	9.3 (12.10)	7.0 (8.51)	3.7 (6.11)	3.2 (4.76)	$P < 0.001$*
MBI efficiency	0.10 (0.35)	0.38 (0.70)	0.51 (0.74)	0.70 (1.04)	0.72 (0.96)	0.51 (0.88)	0.52 (0.65)	$P < 0.001$*

Notes: All scores are reported as a mean (SD). *For each characteristic, one-way ANOVA with admission MFAC as a fixed factor, post-hoc differences exist between the shaded and non-shaded scores, but no post-hoc difference among the shaded scores.

(2.7, 4.8, 3.4, 1.0, and 0.4, respectively). The MRMI efficiency of patients with admission categories of MFAC 3 to 5 (0.34, 0.40, and 0.39) had significant differences ($p < 0.05$) when compared to subjects with admission categories of MFAC 1, 2, 6, and 7 (0.11, 0.26, 0.15, and 0.05, respectively). The MBI gain of subjects with admission categories of MFAC 2 to 5 (9.7, 10.2, 9.3, and 7.0) had significant differences ($p < 0.05$) when compared to subjects with admission categories of MFAC 1, 6, and 7 (3.0, 3.7, and 3.2, respectively). The MBI efficiency of subjects with admission categories of MFAC 4 and 5 (0.70 and 0.72) had significant differences ($p < 0.05$) when compared to subjects with admission categories of MFAC 1 to 3 and 6 to 7 (0.10, 0.38, 0.51, 0.51, and 0.52, respectively) (Table 2).

MFAC profile

FAC profile is a matrix table representing the distribution of discharge MFAC of subjects according to admission category of MFAC (Table 3). The left-hand column of this profile lists the admission categories of MFAC and the top row lists the discharge categories of MFAC. At the intersection of the rows and columns, the possibility in percentage (%) of discharge categories of MFAC of subjects in a rank order for each admission categories of MFAC of subjects could be identified. For example, the change of subjects with admission category of MFAC 2 (sitter) progress to dependent walkers, assisted walkers, supervised walkers, indoor walkers and outdoor walkers upon discharge is 23.9%, 22.2%, 10.1%, 2%, and 0%, respectively (Table 3). The chance of subjects with admission categories of MFAC 1 and 2 progress to any kind of walker (MFAC 3 to 7) is 12.7% and 58.2%, respectively. The chance of subjects with admission categories of MFAC 3 to 5 progress to independent walker (MFAC 5 to 7) are 6.7%, 14.8%, and 50.3%, respectively.

Correlation

Both admission MFAC and admission MBI had strong correlations with discharge MFAC ($r = 0.84$, $P < 0.0001$ and $r = 0.78$, $P < 0.0001$,

Table 3. MFAC profile of subjects ($n = 2{,}722$).

	Discharge MFAC (%)						
	1 Lyer	2 Sitter	3 Dependent walker	4 Assisted walker	5 Supervised walker	6 Indoor walker	7 Outdoor walker
Admission MFAC							
1 Lyer	74.8	12.5	6.7	5.1	0.8	0.1	0
2 Sitter	2.8	39.0	23.9	22.2	10.1	2.0	0
3 Dependent Walker	0.4	4.2	34.4	34.9	19.4	6.5	0.2
4 Assisted Walker	0.2	0.7	1.4	43.4	39.5	13.0	1.8
5 Supervised Walker	0	0.5	0	0.5	48.7	42.3	8.0
6 Indoor Walker	0	0	0	0	0	77.5	22.5
7 Outdoor Walker	0	0	0	0	0	0	100

Table 4. Spearman's rank correlation coefficients (r_s) of functional score at admission and discharge.

	Discharge		
	MFAC	MRMI	MBI
Admission			
MFAC	0.84*	0.82*	0.78*
MBI	0.78*	0.78*	0.94*

Notes: *$p < 0.0001$.

respectively), discharge MRMI ($r = 0.82$, $P < 0.0001$ and $r = 0.78$, $P < 0.0001$, respectively) and discharge MBI ($r = 0.78$, $P < 0.0001$ and $r = 0.94$, $P < 0.0001$, respectively) (Table 4).

Discussion

MFAC is easy to be scored by clinicians and to be understood by patients and their families. To our best knowledge, this is the first study to demonstrate that admission MFAC could be a stratification tool for patients with stroke in inpatient rehabilitation. The present study showed that the groups of patients who had the greatest benefit from stroke rehabilitation program were those with admission categories of MFAC 3 to 5 in terms of MRMI efficiency and those with admission categories of MFAC 4 to 5 in terms of MBI efficiency. The present study echoes with Louie's study[7] that patients with stroke with moderate disability on admission presented the best rate of improvement in functional outcomes on discharge. Under resource constraints, stroke rehabilitation teams have to stratify patients with stroke as soon as possible so as to allocate suitable resources, such as therapy sessions or LOS in hospital, to patients with stroke. Therefore, according to the result of present study, the rehabilitation team could allocate more resources to patients with moderate disability on admission, and formulate early discharge plan for severe and mild disability. However, our finding was a little different from previous studies by Gagnon and colleagues[5] and Lin.[8] Gagnon and colleagues[5] showed that the best total (41.6) and motor-FIM (38.9) gains were observed in most severely disabled patients with stroke (IRF-CMG classification system: 114). In addition, Lin[8] showed that patients with stroke admission FIM total scores of ≥ 73 were scored lower functional gain (16.6+/−11.7) than those who scored of ≤ 36 'and functional gain of (27.6+/−23.3). When considering the functional gain, present study still showed that the best MRMI gains (6.6) and best MBI gain belonged to the patients with moderate disability, i.e., admission category of MFAC 3, but not the most severely disabled patients, i.e., MFAC 1 and 2.

The present study also showed that MFAC and MBI at admission had similar correlation to MFAC, MRMI, and MBI of patients with stroke upon discharging from inpatient rehabilitation program. The admission MFAC had a strong correlation ($r_s = 0.84$, $P < 0.0001$) with discharge MFAC. This finding echoed with previous studies which found that functional score at admission is positively correlated with the functional score at discharge.[13,14] However, how to apply that strong correlation between admission and discharge MFAC and in clinical situation is a challenge. Hence, the present study developed the MFAC profile which makes use of admission mobility level to estimate various possibility of discharge mobility level of patients with stroke. For instance, the chance of patients with admission categories of MFAC 1 and 2 progress to any kind of walker (MFAC 3 to 7) is 12.7% and 58.2%, respectively. Therefore, the sitter has 4.5 times of chance to walk again upon discharge than the lyer. The chance of patients with admission categories of MFAC 3 to 5 progress to independent walker (MFAC 5 to 7) is 6.7%, 14.8% and 50.3%, respectively. The possible mobility level of patient upon discharge (discharge MFAC) from stroke rehabilitation program is useful information for discharge planning with patients and their family.

The present study had its limitations. First, sampling bias may exist due to all subjects which were recruited from only one local rehabilitation hospital. Further studies are suggested to verify the pattern of MFAC profile and functional outcomes stratified by admission MFAC of patients with stroke in different setting, different phases, and different countries.

Conclusion

This study showed that patients on admission with moderate disability in terms of MFAC had the greatest mobility gain and basic ADL gain from inpatient stroke rehabilitation. Admission

MFAC could be a stratification tool of patients with stroke in inpatient rehabilitation.

Conflict of Interest

The author declares that he has no financial affiliations (including research funding) or involvement with any commercial organization that has a direct financial interest in any matter described in this manuscript. The author has no other financial or nonfinancial conflicts of interest related to any matter in this study.

Funding/Support

There were no sources of funding or grant support for this study. The author would like to thank the colleagues of the Physiotherapy Department of Tai Po Hospital for their support and contributions.

Author Contributions

Bryan Ping Ho Chung contributed to the design, data collection, data analysis and manuscript writing of the study.

References

1. World Health Organization. Recommendations on stroke prevention, diagnosis, and therapy: Report of the WHO task force on stroke and other cerebrovascular disorders. Stroke 1989;20:1407–31.
2. Hospital Authority. Hospital Authority Statistical Report 2014;2012–13.
3. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: A systematic review. Lancet Neurol 2009;8:741–54.
4. Bohannon RW, Horton MG, Wikholm JB. Importance of four variables of walking to stroke patients. Int J Rehabil Res 1991;14(3):246–50.
5. Gagnon D, Nadeau S, Tam V. Clinical and administrative outcomes during publicly-funded inpatient stroke rehabilitation based on a case-mix group classification model. J Rehabil Med 2005;37 (1):45–52.
6. Nakao S, Takata S, Uemura H, et al. Relationship between Barthel Index scores during the acute phase of rehabilitation and subsequent ADL in stroke patients. J Med Invest 2010;57(1–2):81–8.
7. Louie SW, Wong SK, Wong CM. Profiles of functional outcomes in stroke rehabilitation for Chinese population: A cluster analysis. NeuroRehabilitation 2009;25(2):129–35.
8. Lin JH. Influence of admission functional status on functional gain and efficiency of rehabilitation in first time stroke patients. Kaohsiung J Med Sci 2001;17(6):312–8.
9. Chau MWR, Chan SP, Wong YW, et al. Reliability and validity of the modified functional ambulation classification in patients with hip fracture. Hong Kong Physiother J 2013;31(1):41–4.
10. Chung BPH. The effectiveness of robotic-assisted gait training in stroke rehabilitation: A retrospective matched control study. Hong Kong Physiother J 2017;36:10–6.
11. Chung BPH. Effect of different combination of physiotherapy treatment approaches on functional outcomes of patients with stroke: A retrospective analysis. Hong Kong Physiother J 2014;32 (1):21–7.
12. Park CS, An SH. Reliability and validity of the modified functional ambulation category scale in patients with hemiparalysis. J Phys Ther Sci 2016;28(8):2264–7.
13. Joa KL, Han TR, Pyun SB, et al. Inpatient stroke rehabilitation outcomes in Korea derived from the Korean brain rehabilitation centers' online database system for the years 2007 to 2011. J Korean Med Sci 2015;30(5):644–50.
14. Inouye M, Kishi K, Ikeda Y. Prediction of functional outcome after stroke rehabilitation. Am J Phys Med Rehabil 2000;79(6):513–8.
15. Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L. Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. Phys Ther 1984;64(1):35–40.
16. Lennon S, Johnson L. The Modified Rivermead Mobility Index: Validity and reliability. Disabil Rehabil 2000;22:833–9.
17. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42:703–9.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 38, No. 2 (2018) 149–160
DOI: 10.1142/S1013702518500130

Hong Kong Physiotherapy Journal
http://www.worldscientific.com/worldscinet/hkpj

Comparative effectiveness of transverse oscillatory pressure and cervical traction in the management of cervical radiculopathy: A randomized controlled study

Adesola Ojo Ojoawo* and Ayodele Damilare Olabode

Department of Medical Rehabilitation, Faculty of Basic Medical Sciences
College of Health Sciences, Obafemi Awolowo University, Ile Ife, Nigeria

**aoojoawo@yahoo.com*

Received 14 September 2016; Accepted 24 October 2017; Published 14 August 2018

Background: Radiating neck pain is one of the major symptoms of cervical radiculopathy (CR).
Objective: This study compared the effects of cervical traction (CT) and transverse oscillatory pressure (TOP) in management of CR.
Methods: Seventy-five participants with unilateral radiating neck pain were randomly allocated into three groups, 25 (14 males, 11 females) for CT, 25 (15 males and 10 females) for TOP and 25 (11 males and 14 females) control (Cnt) group. All participants received massage, cryotherapy and active exercises three times in a week for six weeks. CT was administered to CT group, TOP to TOP group while the third group served as control. Pain intensity (PI) and neck functional disability (NFD) were assessed pretreatment, 3rd and 6th week of intervention. Data were analyzed using descriptive and inferential statistics.
Results: There was a significant reduction in PI and NFD between pretreatment and 6th week in all the groups ($p < 0.05$). The effect size of PI ($F = 7.533$, $p < 0.001$) and disability index ($F = 37.888$, $p < 0.001$) in CT group were significantly lower than that of TOP group at 3rd week. PI of TOP was significantly ($p < 0.05$) lower than that of CT and Cnt groups at the 6th week.
Conclusion: TOP reduces the PI and disability of patients with CR faster compared to CT.

Keywords: Cervical traction; transverse oscillatory pressure (TOP); cryotherapy; neck disability index; visual analogue scale.

*Corresponding author.

Introduction

Radiating neck pain, one of the major symptoms of cervical radiculopathy (CR), though less common than non-radiating neck pain has constituted an important cause of disability; therefore, it is imperative to discover the best way to manage it.[1–6] In addition to neck pain, other most common complaint of individuals with CR are paresthesia and radicular pain and while sensory manifestation can be dermatomal, the expression of pain may be myotomal.[7] Patterns of dematomal pain is common at C4 level followed by C6 and then C7 and scapular pain may occur in 51.6%, pain at periscapular region and in the upper limb, as well as neurological signs such as numbness, weakness and loss of reflexes in the affected nerve root distribution.[8] Painful range of motion and reduced tendon reflex are typically found on the course of examination with more than 10% having upper limb weakness and one-third may present with decreased sensation, and muscular atrophy may be present in less than 2%.[9] The major causes of CR are discogenic and spondylitis combined, which form 68%, while 22% of cases were from intervertebral disc.[9]

The annual incidence report of CR was 83.2 per 100 000 and an increased prevalence in the fifth decade of life among the general population. Approximately 14–71% of adults experience neck pain at some points in their lifetime and the one-year prevalence rate for neck pain in adults ranges from 16% to 75%.[10] Study has found a prevalence of neck pain of 53.6% among learners in the Gauteng Province, South Africa.[11] Also, researchers have documented that neck pain was found to be common among Nigerian university undergraduate and affects females than males.[12] In the south western part of Nigeria, it has been documented that the leading work-related musculoskeletal disorder was low back pain, followed by neck pain among nurses and physiotherapist which is an indication that neck pain is very prominent among musculoskeletal pain in Nigeria.[13,14]

The etiology of neck pain though multifactorial and poorly understood has been linked to factors like poor posture, depression, anxiety, aging, acute injury and occupational or sporting activities.[15,16] This leads to altered joint mechanics, muscle structure or function resulting into mechanical neck pain. Researchers also reported that the most common cause of mechanical neck pain is zygophyseal joint locking and muscle strain.[16]

With respect to the management of CR, Costello[17] observed that conservative treatment is more effective than surgical options. Conservative treatment for CR typically includes therapeutic exercise (range of motion, strengthening), manual therapy (muscle energy techniques, non-thrust mobilization, manipulation), modalities (cryotherapy, traction), massage therapy, medication and cervical collar.[18–21] From the empirical observations of Maitland,[22] transverse oscillatory pressure (TOP) which is one of the manipulative techniques was recommended for unilaterally distributed symptoms of cervical origin. TOP, originated by Nwuga,[23] although one of the frequently used manipulative techniques by physiotherapists, has been claimed to be effective in amelioration of pain intensity (PI) especially radiating pain in cervical, thoracic and lumbar regions.[23,24] It involves mobilization of the spinous process of the vertebrae in the region of the spine that had mechanical pain.[24] This technique was reported to be useful when pain has a unilateral distribution, whether localized to the neck or referred to the upper limb.[24]

Cervical traction (CT) consists of administering a distracting force to the neck in order to separate the cervical segments and relieve compression of nerve roots by intervertebral disks. Various techniques (supine versus sitting; intermittent versus sustained; motorized or hydraulic versus an over-the-door pulley with weights) and durations (minutes versus up to an hour) have been recommended for the management of CR.[25] According to Shirai,[26] CT increases blood flow to neck muscles 2 min after it is applied. A systematic review by Graham *et al.*[27] also reported that there was moderate evidence to support the use of mechanical intermittent CT in the management of cervical disorder. Ojoawo *et al*[28] in their study reported that CT is effective in the management of CR. CT in addition to other exercises is the major treatment technique in many facilities in Nigeria physiotherapy clinic to manage CR but there are paucity of data on its efficacy in Nigerian environment. More so, TOP though effective was not a common practice in Nigerian physiotherapy clinic because of the skill required. The question is, that does TOP and CT have the same result in the management of CR? CT requires kit which in some facilities may not be readily available and TOP demands special skills which were not known by all physiotherapists. If TOP cannot be applied, will CT give the desired result? The purpose of the

study was to compare the effects of TOP and CT in the management of CR.

Methods

The participants for this study were 75 (40 males, 35 females) individuals referred for physiotherapy treatment at the Obafemi Awolowo University Teaching Hospital Complex, Ile-Ife, Nigeria, with CR in either right or left upper limb. They were recently diagnosed patients from the Orthopaedic Clinic of the same hospital. All patients' reported neck pain that radiated distally down the right or left arm to the elbow.

Sample Size Determination

Sample size equation $22S^2/d^2 + 1$ for calculation of study with comparison of three groups according to Dallal[29] is adopted for the study.

$S =$ within group standard deviation, $d =$ expected difference between means within the group. With respect to the study of Ojoawo *et al.*,[30] $S = 2.98$ in one of the groups and d was 2.01. The equation sample size is 22 $(2.98)^2/(2.01)^2 + 1 = 39$ for each group. The total number for the three groups should be 116. The study lasted for a period of two years and three months, the total number of patients with CR fulfilled the requirement for that period was 87 and all of them were considered for the study. However, only 75 (86.2%) patients were able to participate in the study.

Fifty patients had pain radiating to right while 25 patients reported radiation toward the left upper limb. Their pain started 6–8 weeks before the commencement of the study. None of the patients could remember any pathology that precipitated his/her complaint. There were various descriptions of pains with a greater percentage which described the pain as a deep ache in the neck and a peppery sensation into either of the arms

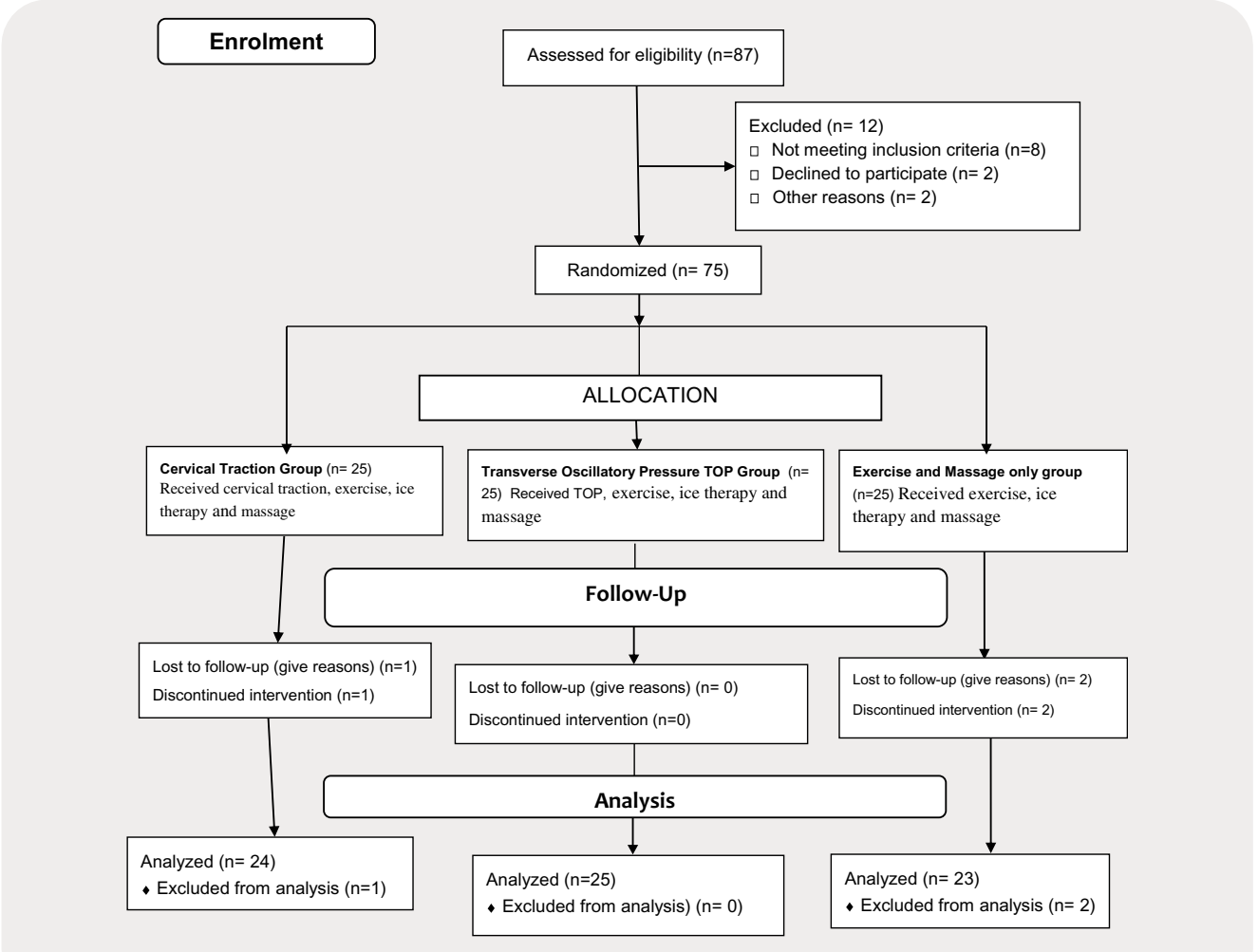

Fig. 1. Consort diagram of random allocation of subjects in to three groups.

that was aggravated by activities at work or even at rest. Fifteen patients reported numbness of some fingers in addition to the radiating pain. Patients with unilateral radiating neck pain that was not of mechanical origin and patients with recent major trauma or fracture of the cervical spine. Patients whose primary complaint was that of headaches or facial pain associated with unilateral radiating neck pain, and any patient who had received manual therapy of the cervical region in the past three months were excluded from the study. Each participant's blood pressure (mmHg), height (m) and weight (kg) were measured. Active range of motion of the neck-elicited pain, especially lateral flexion toward the side of radiculopathy. Skin rolling test according to Bansevicius and Pareja[31] and posterior–anterior pressure according to Egwu[32] to the cervical region provoked pain between the fourth cervical and seventh cervical vertebrae in all patients. Spurling's distraction and Valsalva tests were carried out according to Konin *et al.*[33] and were found positive. Individuals found suitable for the study were randomly allocated into three groups. Seventy five pieces of paper with inscription of CT, TOP and control group (Cnt) on 25 each were put in an opaque envelope. Each participant was asked to pick one and such patient was allocated as inscribed into the paper. Each participant was randomly allocated into CT, TOP and Cnt group. CT group had 24 participants (13 males 11 females) TOP group with 25 participants (15 males and 10 females) while Cnt group had 23 (11male and 12 females) as shown in Fig. 1.

Each participant in the three Groups was treated two times per week. The maximum experimental treatment period for a participant was six weeks, after which the treatment time was estimated from the patient's record.

This amounted to 12 treatment sessions for each participant in the groups. Present PI and neck functional disability (NFD) were measured using visual analogue scale (VAS) and neck disability index (NDI), respectively.

Outcome Measures

Patients were introduced to a 10-point VAS with instructions not to over or underestimate the pain. VAS is a continuous scale comprised of a horizontal or vertical line, usually 10 cm (100 mm) in length, anchored by two verbal descriptors, one for each symptom extreme. The scale is most commonly anchored by "no pain" (score of 0) and "pain as bad as it could be" or "worst imaginable pain" (score of 10; 100-mm scale).[34] Each participant was asked to point to the number corresponding to the PI, which was recorded.

The NDI utilized in the study was in English Language. NDI is a commonly utilized outcome measure to capture perceived disability in patients with neck pain.[35] The NDI contains 10 items: seven related to activities of daily living, two related to pain and one related to concentration.[36] Each item is scored from 0 to 5, and the total score is expressed as a percentage, with higher scores corresponding to greater disability.[36] The NDI has demonstrated moderate test–retest reliability and has been shown to be a valid health outcome measure in a patient population with CR. Westaway *et al.*[37] identified the minimum detectable change as 5 (10% points) in a group of 31 patients with neck pain. Stratford and colleagues[38] identified the minimal detectable change also to be 5 (10% points) in a group of 48 patients with neck pain and arm pain.

Interventions

Exercise therapy

During each appointment, participants in the three groups underwent exercises. All participants performed cervical spine retraction, rotation in each direction especially away from the direction of pain, extension and side-bending stretching exercises on the side that is not painful. Stretching exercises were applied with the aim of releasing any contracture that might set in by virtue of unilateral flexion or side rotation away from the pain. The goal of exercises was to improve endurance of the cervical deep neck muscles to cause the muscles to stop exhibiting painful response and to improve strength according to Ylinen *et al.*[39] Although Ylinen *et al.*[39] examined the general musculature of the neck, deep neck muscles are part of the general neck muscles. Any cervical range of motion that aggravated pain was avoided during the exercise. In a case of high PI, cryotherapy was applied first to relieve the pain which was followed by exercise. The exercises also included passive stretching and isometric exercises to the posterior neck muscles. Isometric exercise was administered according to Kisner and Colby[40] to the posterior neck muscles for 10 s in 10 rounds, in which the contraction was

against resistance of the physiotherapist's hand. Ice chips were packed in a towel and applied to the cervical region for 7 min. This was done to suppress muscle spasm caused by the pain. The isometric exercise, cryotherapy and stretching served as baseline intervention for all the groups.

Cervical traction

The CT group was given CT using the "over the door" CT for 15 min in addition to exercise, ice therapy and massage. A strap was affixed under the chin of the patient. This chinstrap was then connected to a water bag that was held aloft over a doorway via pulleys that were utilized. The water bag was loaded in kilogramme to 10% of the patient's total body weight according to Akinbo *et al.*[41] The 10% weight administered was the minimum weight; this was increased gradually according to the tolerance of each patient to the extent to which there was a desired pull at the cervical region. Treatment was administered twice per week for 6 weeks making 12 treatment sessions. Each session was followed with isometric exercise, cryotherapy and stretching as baseline treatment to improve the strength release contracture and suppress muscle spasm. Patient response was assessed after each second treatment session using VAS and NDI.

Transverse oscillatory pressure

TOP was administered with the patient lying prone on a couch with the forehead placed on the backs of her fingers. Standing on the side of the patient, the therapist placed the pad of the thumbs against the left side (or the right side depending on the location of the pain) of the spinous process of the vertebrae to be moved. The fingers are spread out on the neck and the upper thoracic region. Pressure is directed horizontally through the thumbs to the side of the spinous process. TOP is executed by a pressure-relaxed sequence on the spinous process. Movement is initiated from the trunk and transmitted down the arm to the thumbs.[24] Treatment is affected by a push–relax sequence on the spinous process using the thumbs to produce an oscillatory movement. Transverse pressure was directed toward the side of pain on the cervical vertebrae. The rationale of performing transverse pressure toward painful side is in order to rotate the body of vertebrae away from the side of pain when pressure is applied on the spinous process toward the side of pain. This will restore the joint play and mobilize the articular surfaces on the painful side.[24] The oscillation is done rhythmically with low amplitudes for a period of 20 s. This was repeated three times with a rest period of 2 min for a session per day.[24]

Treatment was administered twice per week for 6 weeks and cumulated to 12 treatment sessions. Each session was followed with isometric exercise in order to improve the strength of the muscles; cryotherapy was applied to suppress the muscle spasm and stretching of the neck muscles with the aim of preventing and releasing any muscular contracture that may be setting in. Patient response was assessed after each second treatment session using VAS and NDI. Patients were advised not to involve in any other intervention without consulting the corresponding author of this paper.

Data Analysis

SPSS version 17.0 (SPSS Inc., Chicago, IL, USA) was used to analyze the data. Descriptive and inferential statistics were used to summarize the result. Analysis of variance (ANOVA) was used to compare each of the physical characteristics of the participants in the three groups. Repeated measures ANOVA was used to compare the pretreatment, 3rd week and 6th week values of PI and NDI of participants in CT, TOP and Cnt within the groups. Repeated measures ANOVA was also used to compare pre-treatment, 3rd week and 6th week values of the outcome measures among the groups. *Post hoc* analysis using Tukey's highest significant difference was carried out to examine which variables were significantly different from each other. An alpha level of 0.05 was set as level of significant. With respect to the subjects that could not complete the six weeks, the principle of intention to treat was applied and the last observation carried forward (LOCF) method was used for the analysis.

Results

Figure 1 is the consort flowchart of the participants. Eighty-seven were recruited for the study but twelve participants did not meet the inclusion criterial. Seventy-five were randomly allocated to three groups but a candidate dropped out from CT group and two participants were not able to complete the study in the Cnt group. Based on the

premise that those who did not complete the study had report of their PI and disability, and were able to complete three weeks of treatment, they were also included in the study. Indicating that 75 participants were analyzed for the study.

Physical characteristics of all the participants were shown in Table 1. There were no significant difference ($p > 0.05$) among the physical characteristics of participants in all the groups. Inferring that the three groups' physical parameters were comparable and that the results obtained from the study were as a result of the intervention not based on variation in the groups' physical characteristics.

In Table 2, the repeated measures ANOVA comparing the effect of TOP, CT and exercise only on the PI and neck disability of participants at pretreatment, 3rd week and 6th week of treatment is presented. There was a significant difference ($p < 0.05$) among the pretreatment, 3rd week and 6th week of PI and neck disability of participants in all the groups.

The mean difference among the outcome measures in the three groups and the observed power is shown in Table 3. The mean difference between the pretreatment and 3rd week PI for CT group was 1.04 and between the 3rd and sixth week was 3.33 ($F = 7.355, p < 0.001$). The difference in disability between the pretreatment and 3rd week was 12.63, and between the 3rd week and 6th week was 9.50. The difference of TOP on PI and disability was greater than that of CT at the 3rd week and 6th week except the PI at 6th week which was less ($F = 23.156, p < 0.001$). The observed power using *post hoc* power analysis was 1 indicating that though the response rate of the subjects was 86.2%, the sample size was enough to give a reliable effect size.

In Table 4, the repeated measures ANOVA with *Post Hoc* Turkey Highest Significance Difference comparing the mean values of PI among the three groups' pretreatment, 3rd week and 6th week is shown. There was no significant difference ($p > 0.05$) among the mean values of pretreatment pain intensity in the three groups. It can be inferred from this that the pretreatment PI among the three groups was comparable, therefore, any result obtained from the study is due to the

Table 1. Physical characteristics of participants ($N = 75$).

Variables	CT group $n = 25$ Mean ± SD	TOP group $n = 25$ Mean ± SD	Crt group $n = 25$ Mean ± SD	F	P
Age (yrs)	51.38 ± 6.545	55.67 ± 5.35	59.50 ± 2.646	3.01	0.08
Weight (kg)	73.13 ± 13.010	73.00 ± 5.36	71.25 ± 5.377	0.05	0.95
Height (m)	1.63 ± 0.12	1.66 ± 0.10	1.65 ± 0.026	0.09	0.92
BMI (Kg/m^2)	27.99 ± 7.96	26.83 ± 4.437	26.02 ± 2.041	0.15	0.86
Sex: M	14	15	11		
F	11	10	14		

Note: BMI = Body Mass Index.

Table 2. Repeated measures ANOVA comparing the mean values of PI and neck disability of the three groups ($N = 75$).

Variables		Pretreatment Mean ± SD	WK3 Mean ± SD	WK6 Mean ± SD	F	P
CT Grp	PI	6.87 ± 0.99	5.83 ± 1.64	2.50 ± 0.53	7.533	0.001**
	ND	42.13 ± 16.86	29.50 ± 17.88	20.00 ± 17.82	37.881	0.001**
TOP Grp	PI	7.63 ± 2.98	4.83 ± 0.75	2.66 ± 0.81	23.156	0.001**
	ND	58.66 ± 8.91	39.00 ± 17.46	16.33 ± 9.75	40.352	0.001**
Cntr Grp	PI	7.00 ± 0.81	6.25 ± 0.95	3.75 ± 0.53	24.540	0.001**
	ND	55.32 ± 11.30	33.82 ± 1.67	21.50 ± 5.00	34.40	0.001**

Note: **significant at $P < 0.001$, WK = week, Grp = Group, Cntr = Control.

Table 3. Observed power and test between subjects' effects within the three groups ($N = 75$).

Groups	PnDif Wk1–3	PnDif Wk3–6	F	P	DisDif Wk1–3	DisDif Wk3–6	F	P	ObP
CT	1.04	3.33	7.533	0.001	12.63	9.50	37.881	0.001	1.000
TOP	2.80	2.17	23.156	0.001	19.66	22.67	40.352	0.001	1.000
Cntr	0.78	2.50	24.540	0.001	21.50	12.32	34.400	0.001	1.000

Note: PnDifWk = Pain difference between weeks 1 and 3, Dis = Disability, ObP = Observed Power.

Table 4. Repeated measures ANOVA with *post hoc* comparing the mean values of PI of the three groups ($N = 75$).

	CT Grp Mean ± SD	TOP Grp Mean ± SD	Cntr Grp Mean ± SD	F	P
Pretreatment	6.87 ± 0.99^{a}	7.63 ± 2.98^{a}	7.00 ± 0.81^{a}	0.688	0.518
WK3	5.83 ± 1.64^{a}	4.83 ± 0.75^{b}	6.25 ± 0.95^{a}	7.800	0.005*
WK6	$2.50 + 0.53^{c}$	1.66 ± 0.81^{d}	3.75 ± 0.53^{e}	10.121	0.002*

Notes: Superscript letters a–e in the table mean that a mean mode of the same superscript letters indicates no significant difference but mean mode with different superscript letters indicates significant difference.

intervention, and not of variation from the pretreatment PI. In the CT and Cnt groups, there was no significant difference between pretreatment PI and 3rd week ($p > 0.05$) but there exists a significant difference ($F = 7.80$, $p < 0.05$) in the pretreatment and 3rd week in TOP group inferring that PI may be ameliorated after 2nd week of intervention of TOP. Considering the time effect, TOP proves to be faster in relieving PI. Nonetheless, there was a significant difference ($p < 0.05$) in the 6th week PI among the three groups with TOP group having the minimum values, interpreting that TOP can relieve PI quicker and more than either CT or exercise only.

Table 5 revealed the magnitude of the effect size in the three groups using partial Eta square. The highest magnitude is between the Cnt group and TOP at the 6th week ($\eta_p^2 = 1.6$) followed by that of CT and Cnt at 6th week ($\eta_p^2 = 1.2$). The magnitude of effect size between TOP and CT group at the third and 6th week was 0.6 and 0.4, respectively. An indication is that TOP reduces PI faster than CT.

Table 6 shows the repeated measures ANOVA and *post hoc* comparison of the disability index among the three groups and Table 7 shows the magnitude of the effect size. There was a significant difference among ($F = 4.08, p < 0.05$) the three groups at third week, there was a significant difference between CT and TOP, and Cnt and TOP but not between CT and Cnt in the 6th week. The effect size of TOP and Cnt at the 6th week was the highest among the others ($\eta_p^2 = 0.4$).

Table 5. Magnitude of effect size using partial Eta square for PI among the three groups for 3rd and 6th weeks.

Variables	$M1$	$M2$	$M1 - M2$	SD1	SD2	SD1 + SD2	PETA (η_p^2)
CT and TOP 3rd WK	5.83	4.83	1.00	1.64	0.75	2.39	0.40
CT and Cnt 3rd WK	5.83	6.25	0.42	1.64	0.95	2.59	0.20
TOP and Cnt 3rd WK	4.83	6.25	1.42	0.75	0.95	1.70	0.80
CT and TOP 6th WK	2.50	1.66	0.84	0.53	0.81	1.34	0.60
CT and Cnt 6th WK	2.50	3.75	1.25	0.53	0.53	1.06	1.20
TOP and Cnt 6th WK	3.75	1.66	2.09	0.81	0.53	1.34	1.60

Table 6. Repeated measures ANOVA with *post hoc* comparing the mean values of neck disability index of the three groups ($N = 75$).

	CT Grp Mean ± SD	TOP Grp Mean ± SD	Cntr Grp Mean ± SD	F	P
Pretreatment	42.13 ± 16.86[a]	58.66 ± 8.91[a]	55.32 ± 11.30[a]	0.125	0.115
WK3	29.50 ± 17.88[b]	39.00 ± 17.46[d]	33.82 ± 1.67[f]	4.870	0.009
WK6	20.00 ± 17.82[c]	16.33 ± 9.75[e]	21.50 ± 5.00[c]	0.118	0.889

Note: Superscript letters a–f in the table mean that a mean mode of the same superscript letters indicates no significant difference but mean mode with different superscript letters indicates significant difference.

Table 7. Magnitude of effect size using partial Eta square for disability among the three groups for 3rd and 6th Weeks.

Variables	$M1$	$M2$	$M1 - M2$	SD1	SD2	SD1 + SD2	PETA (η_p^2)
CT and TOP 3rd WK	39.00	29.50	9.50	17.88	17.46	35.34	0.30
CT and Cnt 3rd WK	33.82	29.50	4.32	17.88	1.67	19.55	0.20
TOP and Cnt 3rd WK	39.00	33.82	5.18	17.46	1.67	19.13	0.30
CT and TOP 6th WK	20.00	16.33	3.67	17.82	9.75	27.57	0.10
CT and Cnt 6th WK	20.00	21.50	1.50	17.82	5.00	22.82	0.10
TOP and Cnt 6th WK	21.50	16.33	5.17	9.75	5.00	14.75	0.40

Notes: M = mean value of disability index; WK = Week; SD = standard deviation; PETA = partial eta.

Discussion

This study compared the therapeutic effect of CT and TOP in the management of CR. The study revealed that the pretreatment PI, neck disability assessment and other physical characteristics of participants in the CT group, TOP and Cnt group did not show any significant difference. This is an indication that the baseline parameters of the participants in all the groups were comparable and any result obtained from the study was due to the intervention. Considering the comparative effect of CT and TOP, it was observed that there was a significant reduction in PI at 3rd week in TOP group than that of CT group. Meanwhile, the mean difference between the 3rd week and 6th week of CT group PI was more than that of TOP. This may be inferred that TOP may relieve PI faster but CT may relieve the PI for longer period of time. It has to be emphasized that the technique of TOP is the application of oscillatory pressure directly to the specific region where there is pain.[24] This may alter segmental biomechanics by releasing trapped meniscoids, releasing adhesions or by diminishing distortion in the intervertebral disc[42,43] and restored joint play which immediately increases the mobility of the region.[24] Also, individual motion segments are thought to be capable of buckling, thereby producing relatively large vertebral motions that achieve a new position of stable equilibrium.[44] The manipulative impulse provides sufficient energy to restore a buckled segment to a lower energy level, thus reducing mechanical stress or strain on soft and hard spinal tissues.[45] Giles[43] proposed that spinal manipulation activates all known mechanosensitive, somatosensory receptors because they all possess mechanical thresholds lower than the peak force delivered during a manipulation and the receptor types are responsive to dynamic and/or static components of a mechanical stimulus. These may be reasons while TOP relieves the PI of patient faster.

Elnaggar *et al.*[46] in their study reported that CT methods had a significant effect on neck and arm pain reduction, a significant improvement in nerve function and a significant increase in neck mobility. Our finding was also in consistence with the work of Voltonen *et al.*[47] who concluded that traction relieves muscle spasm and significantly decreases electrical activity in the muscles and produces relaxation, which leads to systematic relief of pain.[47]

Krause *et al.*[48] found that traction has been shown to separate the vertebrae, stretch the cervical joint capsules, stretch neck muscles and open the foramina. These may be the reasons while CT relieves PI of patients with CR longer.

It was observed that there was a significant improvement between the pretreatment PI and the six-week PI and disability index in CT group. This inferred that CT was effective in the management of CR. Our findings were in line with the study of Borman *et al.*[49] and Cleland *et al.*[20] Borman *et al.*[49] and Cleland *et al.*[20] in their independent studies using intermittent CT documented that the application of CT produced a desired result in the management of CR. In addition, Rhee *et al.*[50] and Swezey *et al.*[25] reported that the application of CT at home was found to decrease radicular symptoms. Levine *et al.*[51] documented further that CT is most beneficial when acute muscular pain has subsided and should not be used in patients who have signs of myelopathy. CT was known in theory to distract the neural foramen and decompresses the affected nerve root.[52] Evidence also revealed that continuous CT decreases the pressure within the vertebral disks and stretches muscles and ligaments of the cervical region thereby unloads the structure of the spine.[53] It is probable that traction has an important role in breaking the "circle of pain" in CR caused by a herniated disk. This cycle begins when nerve roots are compressed by a herniated disk, causing entrapment within the intervertebral foramina. The irritated nerve produces a reflex response to the patient's cervical muscles, causing those muscles to contract. That contraction further narrows the foramina and the neck pain is increased. Traction helps to relieve the inflammatory reaction of nerve roots by improving the circulation and reducing the tissues swelling. Gentle alteration of stretching and relaxation of the neck soft tissue structures prevents the formation of adhesions of the dural sleeve.[54] These are additional factors why CT can reduce the PI of cervical radicular patients longer than TOP.

The study observed in addition that TOP group reported significant decrease in the outcome measures when the pretreatment mean values were compared with the posttreatment values. This is an indication that TOP has a significant therapeutic value in the management of CR. Researchers reported that manipulation may provide short-term benefit in the treatment of neck pain, cervicogenic headaches[51] and radicular symptoms.[55] Researchers have reported that TOP to the spinal region has both neurological and mechanical effects.[56,57]

Paris[57] mentioned further that mobilization technique stretches tissues by taking them into the area of plastic deformation of stress–strain curve. TOP has been found as one of the techniques that stretches cervical connective tissues and joint capsules to a reasonable point on the stress–strain curve to produce a salvo of beneficial neuro inhibitory and mechanical effects.[56,57]

Considering the participants in the cnt group of the study, it was revealed that exercise with massage and cryotherapy also reduced the PI and disability of participants. This is an indication that pain of CR and NDI can be ameliorated when treated with a combination of exercise, massage and cryotherapy. The improvement of outcome measures in cnt group is in line with the observation of Radhakrishnan *et al.*[9] Studies have examined the effect of isometric exercise on the contracting body part as well as on the contralateral and a distant body part to the contracting one and affirmed that the hypoalgesic effect of isometric exercise was multisegmental and not isolated to the contracting muscle.[58,59] Moreover, the pain-reducing effects of isometric exercise on the contralateral and distant body parts were similar in magnitude to the local body part. These results suggest that a central widespread inhibitory mechanism is activated by static muscle contractions. As discussed by Kosek and Lundberg,[60] these central mechanisms may include increased secretion of *b*-endorphins, attention mechanisms, activation of diffuse noxious inhibitory controls or an interaction of the cardiovascular and pain regulatory systems.

The effect of kneading massage in this study is explained by a researcher who documented that massage has traditionally been used to relieve pain in producing short-lived analgesia by activating the "pain gate" mechanism.[61] Cutaneous mechanoreceptors are stimulated by touch and transmit information within large nerve fibers to the spinal cord.[62] These impulses block the passage of painful stimuli entering the same spinal segment along small, slowly conducting neurons.[63] Massage is a potent mechanical stimulus and a particularly effective trigger for the pain gate process which can reinforce a naturally occurring discomfort, cause much greater release of opiates and achieve more profound pain suppression.[64] The contribution of

cryotherapy in the relief of pain has been reported in studies noting that cryotherapy may be most effective when combined with exercise.[65] Adequate cooling can reduce pain, spasm and neural inhibition, thereby allowing for earlier and more aggressive exercises. Cryotherapy can increase pain tolerance and pain threshold and decrease nerve conduction velocity.[66]

Limitations

There are some limitations of this study, one of them is that the physiotherapist who treated the patients was not blinded to the group allocation and the assessors were not totally independent of the intervention. The reason is that the hospital administration where the study was carried out did not permit blinding. The study also did not assess the range of motion and strength of the cervical muscles. The researchers in the proposal did not consider the variables as part of the objectives of the study, though it may be an omission which has been noted for subsequent study, but the opinion is that ones the pain and disability have been addressed, other variables will fall in line. Researchers were aware of the results of intermittent traction using traction machine in the management of CR, but lack of funds was a major constrain why this could not be used. More so, facilities in Nigeria with CT machine are very limited; it is then imperative that a research is carried out on what is commonly available in the environment of practice. The less favorable results of CT could be due to continuous traction adopted in the present study. The results may be different if intermittent CT was given to the patients. Graham *et al.*[21] in the systematic review concluded that intermittent traction is better than continuous traction for mechanical neck disorders. The short- or long-time follow-up could not be concluded because of some logistic problems.

Conclusion

It can be concluded from the study that combination of exercise, massage and cryotherapy reduces PI and disability of patient with CR in the 6th week of intervention but addition of TOP proved better and may be better than inclusion of CT.

Conflict of Interest

Authors did not have any conflict of interest on the study.

Funding/Support

The study did not enjoy funding from any organization or government.

Authors Contributions

Ojoawo A.O conceived the idea of the study, analyzed the data, interpreted and did the final write-up including the proof reading to make it suitable for publication. Olabode A.D collected the data, supplied the literature and did the skeletal write-up.

References

1. Rubinstein SM, Pool JJ, van Tulder MW, Riphagen II, de Vet HC. A systematic review of the diagnostic accuracy of provocative tests of the neck for diagnosing cervical radiculopathy. Eur Spine J 2007;16:307–19.
2. Childs JD, Cleland JA, Elliott JM, et al. Neck pain: Clinical practice guidelines linked to the international classification of functioning, disability, and health from the orthopedic section of the American Physical Therapy Association. J Orthop Sports Phys Ther 2008;38:A1–34.
3. Tsakitzidis G, Remmen R, Peremans L, et al. Non-specific neck pain: Diagnosis and treatment. KCE reports. 2009:119C.
4. Manchikanti L, Boswell MV, Singh V, et al. Comprehensive evidence-based guidelines for interventional techniques in the management of chronic spinal pain. Pain Physician 2009;12:699–802.
5. Daffner SD, Hilibrand AS, Hanscom BS, Brislin BT, Vaccaro AR, Albert TJ. Impact of neck and arm pain on overall health status. Spine 2003;28:2030–5.
6. Haldeman S, Carroll L, Cassidy JD, Schubert J, Nygren A. The bone and joint decade 2000–2010. Task force on neck pain and its associated disorders: Executive summary. Spine 2008;33:S5–7.
7. Slipman CW, Plastaras CT, Palmitier RA. Symptom provocation of fluroscopically guided cervical nerve root stimulation. Are dynatomal maps identical to dermatomal maps? Spine 1998;23 (20):2235–42.
8. Murphy D, Hurwitz E, Gregory A. A nonsurgical approach to the management of patients with

cervical radiculopathy: A prospective observational cohort study. J Manipulative Physiol Ther 2006;29 (4):279–87.
9. Radhakrishnan K, Litchy WJ, O'Fallon WM, Kurland LT. Epidemiology of cervical radiculopathy: A population-based study from Rochester, Minnesota, 1976 through 1990. Brain 1994;117: 325–35.
10. Fejer R, Kyvik KO, Hartvigsen J. The prevalence of neck pain in the world population: A systematic critical review of the literature. Eur Spine J 2006;15:834–48.
11. Mafanya C, Rhoda A. Predictors of neck pain among South African youth. Afr J Phys Health Educ Recreat Dance 2011;18(3):237–42.
12. Ayanniyi O, Mbada CE, Iroko OP. Neck pain occurrence and characteristics in nigerian University Undergraduates. TAF Prev Med Bull 2010;9 (3):167–74.
13. Tinubu B, Mbada CE, Oyeyemi AL, Fabunmi AA. Work-related musculoskeletal disorders among nurses in Ibadan, South-west Nigeria: A cross-sectional survey. BMC Musculoskelet Disord 2010;11:12.
14. Adegoke BO, Akodu Ak, Oyeyemi AL. Work-related musculoskeletal disorders among Nigerian physiotherapists. BMC Musculoskelet Disord. 2008;9:112. doi: 10.1186/1471-2474-9-112.
15. Binder AI. Cervical spondylosis and neck pain. Br Med J 2007;334(7592):527–31.
16. Peterson DH, Bergman TF. Chiropractic Technique: Principles and Procedures. United States of America: Mosby, 2002.
17. Costello M. Treatment of a patient with cervical radiculopathy using thoracic spine thrust manipulation, soft tissue mobilization, and exercise. J Man Manip Ther 2008;16:129–35.
18. Wainner RS, Gill H. Diagnosis and nonoperative management of cervical radiculopathy. J Orthop Sports Phys Ther 2000;30:728–44.
19. Cleland JA, Fritz JM, Whitman JM, Heath R. Predictors of short-term outcome in people with a clinical diagnosis of cervical radiculopathy. Phys Ther 2007;87:1619–32.
20. Cleland JA, Whitman JM, Fritz JM, Palmer JA. Manual physical therapy, cervical traction, and strengthening exercises in patients with cervical radiculopathy: A case series. J Orthop Sports Phys Ther 2005;35:802–11.
21. Waldrop MA. Diagnosis and treatment of cervical radiculopathy using a clinical prediction rule and a multimodal intervention approach: A case series. J Orthop Sports Phys Ther 2006;36:152–9.
22. Maitland GD. Vertebral Manipulation. 5th ed. Boston: Butterworth-Hienemann, 2003:201–12.
23. Nwuga VCB. Techniques of spinal manipulation. In: Nwuga VC, ed. Manipulation of the Spine. Baltimore, MD: William and Wilkins, 1975:53.
24. Nwuga VCB. Techniques of spinal manual therapy. In: Manual Treatment of Back Pain. 2nd ed. Baltimore, MD: William and Wilkins, 2007:115.
25. Swezey RL, Swezey AM, Warner K. Efficacy of home cervical traction therapy. Am J Phys Med Rehabil 1999;78(1):30–2.
26. Shirai Y. Intermittent cervical traction in subjects with neck and shoulder pain — Analysis of a blood flow volume and EMG signals. Sogo-Riha 1995;23:25–30.
27. Graham N, Gross AR, Goldsmith C. Mechanical traction for mechanical neck disorders: A systemic review. J Rehab Med 2006;38(3):145–52.
28. Ojoawo AO, Olabode A, Esan O, Badru A, Odejide S, Arilewola B. Therapeutic efficacy of cervical traction in the management of cervical radiculopathy: A control trial. Rwanda J Health Sci 2013;2 (2):25–29.
29. Dallal GE. Sample size calculations simplified. In: The Little Handbook of Statistical Practice. Boston, MA: Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University, 2012. Available at: The UCLA Department of Statistics website. Downloaded on June 25, 2016.
30. Ojoawo AO, Olabode A, Esan O, Badru A, Odejide S, Arilewola B. Transverse oscillatory pressure in management of cervical radiculopathy: A randomised controlled study. Hong Kong Physiother J 2016;34:19–26.
31. Bansevicius D, Pareja JA. The "skin roll" test: A diagnostic test for cervicogenic headache? Funct Neurol 1998;13:125–33.
32. Egwu MO. Relative therapeutic efficacy of some vertebral mobilization techniques in the management of unilateral cervical spondylosis: A comparative study. J Phys Ther Sci 2008;20:103–8.
33. Konin JG, Wiksten DL, Isear Jr JA, Brader H. Cervical spine. Special test for orthopedic examination. 3rd ed. Thorofare, NJ: Slack, 2006:13–7.
34. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain. Arthritis Care Res 2011;63: s240–52.
35. Pietrobon R, Coeytaux RR, Carey TS, Richardson WJ, DeVellis RF. Standard scales for measurement of functional outcome for cervical pain or dysfunction: A systematic review. Spine 2002;5: 515–22.
36. Vernon H, Mior S. The neck disability index: A study of reliability and validity. J Manipul Physiol Ther 1991;14:409–515.
37. 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;27:331–38.
38. Stratford PW, Riddle DL, Binkley JM, Spadoni G, Westaway MD, Padfield B. Using the neck disability index to make decisions concerning individual patients. Physiother Can 1999;51:107–12.
39. Ylinen J, Takala EP, Nykanen M. Active neck muscle training in the treatment of chronic neck pain in women: A randomized controlled trial. JAMA 2003;289:2509–16.
40. Kisner C, Colby LA. Resisted exercise for impaired muscle performance. In: Therapeutic Exercise Foundation and Techniques. 5th ed. Philadelphia: F. A. Davis Company, 2007:169.
41. Akinbo SR, Noronha CC, Okanlawon AO, Danesi MA. Effects of different cervical traction weights on neck pain and mobility. Niger Postgrad Med J 2006;13(3):230–5.
42. Farfan HF. The scientific basis of manipulation procedures. In: Buchanan WW, Kahn MF, Laine V, et al. eds. Clinics in Rheumatic Diseases. London: W. B. Saunders Company, 1980:159–77.
43. Giles LGF. Anatomical Basis of Low Back Pain. Baltimore: Williams & Wilkins, 1989.
44. Wilder DG, Pope MH, Frymoyer JW. The biomechanics of lumbar disc herniation and the effect of overload and instability. J Spinal Disord 1988;1:16–32.
45. Triano J. The mechanics of spinal manipulation. In: Herzog W, ed. Clinical Biomechanics of Spinal Manipulation. New York: Churchill Livingstone, 2001:92–190.
46. Elnaggar IM, Elhabashy HR, Abd El-Menam EM. Influence of spinal traction in treatment of cervical radiculopathy Egypt. J Neurol Psychiat Neurosurg 2009;46(2):455–60.
47. Voltonen EJ, Moller K, Wiljasob M, Arate B. Comparative radiographic study of intermittent and continuous traction on elongation of cervical spine. J Ann Med Intern 1996;57:143–6.
48. Krause M, Refshauge KM, Desen M, Boland R. Lumbar spine traction: Evaluation of effects and recommended application for treatment. Man Ther 2000;5:72–81.
49. 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):1249–53. doi: 10.1007/s10067-008-0895-z.
50. Rhee JM, Yoon T, Riew KD. Cervical radiculopathy. J Am Acad Orthop Surg 2007;15(8):486–94.
51. Levine MJ, Albert TJ, Smith MD. Cervical radiculopathy: Diagnosis and nonoperative management. J Am Acad Orthop Surg 1996;4(6):305–16.
52. Eubanks JD. Cervical radiculopathy: Nonoperative management of neck pain and radicular symptoms. Am Fam Physician 2010;81(1):33–40.
53. Saunders DH. Use of spinal traction in the treatment of neck and back conditions. Clin Orthop 1983;179:31–8.
54. Bland JH. Disorders of the cervical spine: Diagnosis and medical management. Philadelphia, PA: WB Saunders, 1994; Bleakley C, McDonough S, MacAuley D, The use of ice in the treatment of acute soft-tissue injury. Am J Sports Med 2004;32 (1):251–61. doi: 10.1177/0363546503260757.
55. Haneline M. Chiropractic manipulation in the presence of acute cervical intervertebral disc herniation. Dyn Chiropract 1999;17(25).
56. Egwu MO, Alabi M, Nwuga VCB. Effect of vertical oscillatory pressure on neck pain and some cardiovascular variables. Physiotherapy 2003;89:666–74.
57. Paris SV. Mobilization of spine. Phys, Ther 1979;58:988–95.
58. Lannersten L, Kosek E. Dysfunction of endogenous pain inhibition during exercise with painful muscles in patients with shoulder myalgia and fibromyalgia. Pain 2010;151:77–86.
59. Staud R, Robinson M, Price D. Isometric exercise has opposite effects on central pain mechanisms in fibromyalgia patients compared to normal controls. Pain 2005;118:176–84.
60. Kosek E, Lundberg L. Segmental and plurisegmental modulation of pressure pain thresholds during static muscle contractions in healthy individuals. Eur J Pain 2003;7:251–8.
61. Jacobs M. Massage for the relief of pain: Anatomical and physiological considerations. Phys Ther Rev 1960;40:93–8.
62. Watson J. Pain mechanisms: A review: 1. Characteristics of the peripheral receptors. Aust J Physiother 1981;27:135–43.
63. Goats GC. Massage — the scientific basis of an ancient art: Part 2. Physiological and therapeutic effects. Br J Sports Med 1994;28(3):153–6.
64. Nordschow M, Bierman W. Influence of manual massage on muscle relaxation: Effect on trunk flexion. Phys Ther 1962;42:653.
65. Knight KL, Brucker JB, Stoneman PD. Muscle injury management with cryotherapy. Athletic Ther Today 2000;5:26–30.
66. Algafly AA, George KP. The effect of cryotherapy on nerve conduction velocity, pain threshold and pain tolerance. Br J Sports Med 2007;41:365–9. Available at: http://dx.doi.org/10.1136/bjsm.2006.031237.