Scand J Rehab Med 31: 139–152, 1999

A CRITICAL ANALYSIS OF RANDOMISED CLINICAL TRIALS ON NECK PAIN AND TREATMENT EFFICACY. A REVIEW OF THE LITERATURE

Görel V. Kjellman, RPT, Elisabeth I. Skargren, RPT, PhD and Birgitta E. Öberg, RPT, Dr Med Sc

From the Department of Neuroscience and Locomotion, Physiotherapy, Faculty of Health Sciences, Linköping University, Linköping, Sweden

ABSTRACT. The efficacy of physiotherapy or chiropractic treatment for patients with neck pain was analysed by reviewing 27 randomised clinical trials published 1966–1995. Three different methods were employed: systematic analyses of; methodological quality; comparison of effect size; analysis of inclusion criteria, intervention and outcome according to The Disablement Process model. The quality of most of the studies was low; only one-third scored 50 or more of a possible 100 points. Positive outcomes were noted for 18 of the investigations, and the methodological quality was high in studies using electromagnetic therapy, manipulation, or active physiotherapy. High methodological quality was also noted in studies with traction and acupuncture, however, the interventions had either no effect or a negative effect on outcome. Pooling data and calculation of effect size showed that treatments used in the studies were effective for pain, range of motion, and activities of daily living. Inclusion criteria, intervention, and outcome were based on impairment in most of the analysed investigations. Broader outcome assessments probably would have revealed relationships between treatment effect and impairment, functional limitation and disability.

Key words: neck pain; physiotherapy; chiropractic; randomised clinical trials; review; meta-analysis; methodology; effect size; the "Disablement Process".

INTRODUCTION

Neck problems are extremely common, especially among women (4, 16, 17, 28). This is clearly illustrated by two different follow-up studies showing that more than 80% of patients with neck or shoulder complaints continued to have symptoms two years after participating in a rehabilitation programme (10), and that 57% still experienced neck pain ten years after the onset of symptoms (15). Even if neck problems are prevalent, the underlying pathology and aetiology remain unclear, hence treatments are often based on symptomatic signs. Consequently, it is important that research be focused on neck pain and that therapeutic outcomes be evaluated.

Although randomised controlled trials are highly reliable, few have focused on neck pain. In two separate systematic analyses, Koes et al. (23) found only five such trials on neck pain, and van der Heijden et al. (18) found only three concerning the efficacy of traction for back and neck pain. Due to the limited number of trials examined in those studies, it is difficult to draw any conclusions about treatment efficacy. Aker et al. (1) reviewed 24 studies that dealt solely with the outcome of neck pain, and, by calculating effect size, observed a positive effect of manual treatment. When ascertaining therapeutic effect, it is important to assess the outcome in relation to the purpose of the intervention, but this is seldom considered in traditional meta-analyses. We examined outcomes in relation to the main neck problems or conditions reported in the studies we reviewed.

Our methodology included use of a model called The Disablement Process, which has been developed to clarify concepts related to disablement and to aid communication between researchers and clinical workers (47). The first disablement scheme was developed in the 1960s by Nagi (33). In the 1990s, Verbrugge & Jette (47) extended and revised the scheme, which they called The Disablement Process, adding "risk factors", "extra-individual factors" and "intra-individual factors" to the Nagi scheme. The Disablement Process describes how chronic and acute conditions affect the function of specific body systems, fundamental physical and mental actions, and activities of daily life. It also elucidates personal and environmental factors that speed up or slow down the disablement process. Neck problems can seldom be described by defining a disease or an organic cause *per se,* and therefore we used the scheme in our investigation, as it reflects an illness perspective by classifying the consequences of disease. One can expect that in clinical studies that deal with neck problems, an illness rather than a disease perspective is used. Our

objective was to critically review randomised studies of neck pain in regard to methodological quality and treatment effect size, as well as types of assessment, inclusion criteria, and interventions.

METHODS

MEDLINE and Cinahl were searched for relevant articles published during the period 1966–1995. The key word *randomised* was used in combination with free-text terms, such as neck, cervical, pain, physiotherapy, physical therapy, chiropractic, exercise, rehabilitation, studies, outcome, and evaluation. We also checked the reference lists of the papers identified in the database searches. We excluded abstracts, unpublished studies, and studies that involved both neck and low back pain, as well as articles in which it was impossible to distinguish between the results obtained for different groups of disorders. To be included in our analyses, the trials had to meet the following criteria: a randomised procedure was used for treatment allocation; treatment modality was either physiotherapy or chiropractic; patients had ongoing neck pain at the time of the study; the paper was written in English or a Scandinavian language.

When scrutinising studies that used more than one treatment modality, we analysed the method focused upon by the authors or pointed out by reviewers. We conducted our work with three different methods: first, we examined methodological quality by using a procedure described by Ter Riet et al. (43) and modified by Koes et al. (23, 24); second, treatment efficacy (effect size) was determined by comparing values obtained with Cohen's D (38); third, The Disablement Process scheme (21, 47) was used to analyse inclusion of patient groups, interventions, and outcomes.

Analysis of methodological quality (MQ)

When analysing methodological quality, we minimised bias by acting as separate reviewers. One reviewer (B.Ö.) determined whether a published trial fulfilled the inclusion criteria for our investigation; this reviewer was blinded to the author(s), to the journal of publication, and to treatment outcome. The same paper was then independently analysed by the other two reviewers (G.K. and E.S.), who were blinded only to the author(s) and the journal; if divergent results were obtained, we tried to reach consensus through discussion.

The analysis was based on methodological criteria developed by Ter Riet et al. (43). These criteria are well accepted for use in intervention studies and, after some modifications, have been employed in several systematic analyses and meta-analyses (18, 23–25, 39). The criteria cover four main categories: study population, intervention, measurement of effect, and data presentation and analysis. An individual category comprises two to six items, each with a given weight; together, the 4 categories include 16 items. In our investigation, the reviewed studies were scored on the basis of how well the different items were described. The maximum score was 100 points, and, according to Koes et al. (23, 24), a score of 50 or greater indicates good methodological quality; further details have been presented in the cited investigations. We hereafter refer to our methodological quality as MQ scores.

Based on the conclusions drawn by the publishing authors, we also designated the treatment outcomes of the analysed trials as "positive", meaning better for the experimental group than for the reference group, or as "equivalent/negative", indicating that there was no difference in outcome between the treatments studied or that the reference treatment gave better results. Moreover, the papers were assigned to one of the following two groups on the basis of the duration of complaints described in the inclusion criteria: "acute" = <12 weeks; "chronic" ≥ 12 weeks. Investigations in which no specific criteria had been used for acute and chronic problems were instead assigned to a group designated "mixed". The trials were further categorised on the basis of the types of treatment used: specific individual manual interventions, active multiple and group interventions, and different kinds of electrotherapy.

Calculation of effect size

To further examine the methodological quality of the published investigations, we calculated effect size to determine the possible influence of therapies on neck problems and to summarise the results presented in the papers, regardless of the intervention used. Effect size analysis does not take into account conclusions drawn by the authors of the articles. The last follow-up assessment in each paper was included in our computations, if the standard deviation had been reported or could be calculated. Effect size was determined separately for the outcome variables mentioned in the papers. If a trial comprised more than two treatment modalities, effect size was ascertained solely for the experimental group and the first reference group. Effect size is a standard pooling statistic used in meta-analysis (2, 38), and it is computed by using Cohen's D (38), which is written as follows:

$$Effect\ size = [Xe - Xc]/Sc$$

where Xe is the mean of the experimental group, Xc is the mean of the control group, and Sc is the standard deviation of the control group. The value obtained with Cohen's D represents the difference between the outcomes in the experimental and the control groups; a positive value indicates a better outcome for the experimental group, and a negative value a better outcome for the control group(s). According to Thomas & Nelson (44) an effect size greater than 0.8 is large, around 0.5 is moderate and an effect size less than 0.2 is small.

Classification by use of the Disablement Process scheme

The inclusion criteria, interventions, and outcomes reported in the analysed studies were classified according to the Disablement Process (21, 47). The scheme has four components, i.e. active pathology, impairment, functional limitation, and disability. The central goal of the scheme is to delineate the major pathway from disease or active pathology to various types of functional consequences. The model was used as a framework for this part of the analysis.

Components of the Disablement Process

Pathology. Refers to biochemical and physiological abnormalities: e.g. diagnoses of disease, injury, and congenital/developmental conditions.

Impairment. Refers to a dysfunction and significant structural abnormalities in specific body systems (such as musculoskeletal, cardiovascular and neurological): e.g. muscle weakness and joint restriction.

Functional limitations. Refers to restrictions in performing fundamental physical and mental actions in activities and daily

life: e.g. ambulating, reaching, stooping, climbing stairs, producing intelligible speech, and seeing standard print.

Disability. Refers to the difficulty experienced in carrying out activities in any domain of life due to a health or physical problem: e.g. job, household management, personal care, hobbies, active recreation, clubs, socialising with friends and kin, child care, errands, sleep, and trips.

Besides the above-mentioned components, there are three other aspects that influence the Disablement Process: *extra-individual factors*, such as medical care and rehabilitation, medications, external support, and the physical and social environment; *intra-individual factors*, such as lifestyle and behavioural changes, psychosocial attributes and coping; and *risk factors* of predisposing characteristics such as lifestyle and social, behavioural, and biological factors.

An intervention, whether physiotherapy or chiropractic, is an extra-individual factor, and we classified this component according to where, on the main pathway of the disablement process, it is meant to exert an influence. A reference treatment was classified if we considered it to represent active therapy but not if it represented placebo treatment, a control group, or prescribed medication.

RESULTS

Twenty-eight papers (3, 5–9, 11–14, 19, 20, 22, 26, 27, 29–32, 34–37, 41, 42, 45, 46, 48) were found to satisfy our inclusion criteria. However, two of these ((30) a two-month follow-up, and (31) a two-year follow-up) were analysed as a single publication, because they emanated from the same investigation. Accordingly, a total of 27 studies were considered in our analysis.

Methodological quality

Only 9 of the 27 trials (3, 5, 8, 12, 13, 22, 26, 36, 45) had MQ scores $\geq$50 (Table I); the total scores varied from 24 to 62. The greatest deficiency was seen in the category *measurement of effect*, for which scores were very low for three of the four criteria: patient blinded or time restriction, relevant outcome measures and blinded outcome assessments. The studies also scored very low in two of the four criteria in the category *intervention:* avoiding other simultaneous treatments and comparison with a placebo therapy. Small sample size was a general problem; the smallest group consisted of 50 or more subjects in only three studies and never included more than 100 subjects. Moreover, MQ scores were low for the following categories: adequate randomisation procedures, comparability of relevant baseline characteristics, and comparison with an existing treatment modality.

Of the 27 studies we examined, based on conclusions of the publishing authors, 18 had a positive outcome and nine showed an equivalent/negative outcome between the treatments used (Table I). In all, 33% of the studies (9/27) had MQ scores $\geq$50. Furthermore, one-third of the investigations with positive outcomes (6/18) and one-third of those with equivalent/negative outcomes (3/9) had scores $\geq$50.

As described in the "Methods" section, we used the inclusion criteria for duration of complaints to assign the studies to groups comprising acute, chronic, and mixed neck problems. Four studies were placed in the acute group (Table II); author-based outcomes were positive for three of these (13, 32, 34) and equivalent/negative for one (30). Only one of the positive-outcomes studies (13), in which patients had been given electromagnetic therapy, had an MQ score $\geq$50. Twelve of the studies were assigned to the chronic group (Table III); outcomes were positive for nine of these (3, 7, 9, 11, 20, 29, 35, 37, 45) and equivalent/negative for three (22, 36, 46). Two of the positive-outcome investigations (manipulation [3] and electromagnetic therapy [45]) and two of the equivalent/negative studies (traction [22] and acupuncture treatment [36]) had MQ scores $\geq$50. Eleven studies were included in the group designated mixed (Table IV); outcomes were positive for six of these (6, 8, 12, 19, 26, 48) and equivalent/negative for five (5, 14, 27, 41, 42). Four of the investigations in the mixed group—three with positive outcomes (electromagnetic therapy [12], manipulation [8] and active physiotherapy treatment [26]) and one with an equivalent/negative outcome (traction [5])—had MQ scores $\geq$50.

The 27 studies were also grouped according to the type of intervention used: four employed acupuncture, four manipulation, three mobilisation, three traction, eight active physiotherapy and group intervention, and five electrostimulation/local heat. (Details regarding main and reference treatments are given in Tables II–IV). Six of the nine studies with MQ scores $\geq$50 had positive author-based outcomes, and these investigations comprised three different kinds of interventions: active physiotherapy (26), electromagnetic therapy (12, 13, 45), and manipulation (3, 8). Equivalent/negative outcomes were noted for studies concerning acupuncture (36) and cervical traction (5, 22). None of the studies using mobilisation had MQ scores $\geq$50.

Follow-up periods were generally short (Table I). In eleven of the studies (41%) treatment effect was assessed only directly after interventions (6, 8, 11–13, 27, 29, 32, 35, 37, 48). Only five studies (19%) included long-term follow-ups, which were performed at 6 months (22, 34, 41), at 1 year (26), and at 2 years (31).

Table I. *Randomised clinical trials on neck pain patients receiving physiotherapy or chiropractic treatment, in order of the methodological quality scores*

First author	Year of publication	Methods criteria [max possible points] A [2]	B [5]	C [4]	D [3]	E [4]	F [17]	G [10]	H [5]	I [5]	J [5]	K [5]	L [10]	M [10]	N [5]	O [5]	P [5]	Total MQ score [100]	Conclusions of author(s)	Indication	Follow-up period from start	Total study population on entry
Boline	1995	2	4	4	3	2	8	10	5	5			6		3	5	5	62	pos	chronic headache	10 wk	150
Foley-Nolan	1992	2	4		3	4		10	5		5	5	8	8	3	5		62	pos	acute whiplash	12 wk	40
Trock	1994	2	3	4	3	4		10			5	5	6	6	3	5	5	61	pos	chronic neck pain	8–10 wk	81
Foley-Nolan	1990	2	3	4	3	4		10			5	5	8	8	3	5		60	pos	mixed	6 wk	20
Klaber Moffett	1990	2	4			2	8	10			5	5	8	8	3		5	60	equiv/neg	chronic neck pain	12 wk	100
Petrie	1986	2	3		3	4		10			5	3	10	2	3	5	5	55	equiv/neg	chronic neck pain	8 wk	27
Cassidy	1992	2	4		3	4		10	5	5			4	4	3	5	5	54	pos	mixed	1 wk	100
BAPM	1966	1	3				8	10	5		5	3	6	6	5			52	equiv/neg	mixed	24 wk	466
Levoska	1993	2	2		3	4		10	5	5			4		5	5	5	50	pos	mixed	52 wk	47
Goldie	1970	2	3	4	3	4		10	5				4	4	5	5		49	equiv/neg	mixed	24 wk	73
Sloop	1982	1	2	2	3	4		5			5	5	4	4	3	5	5	48	equiv/neg	mixed	12 wk	39
Petrie	1983	2	3		3	4		10			5	5	2		3	5	5	47	pos	chronic neck pain	4 wk	13
Revel	1994	1	3		3	4		10		5			8		3	5	5	47	pos	chronic neck pain	10 wk	60
Zylbergold	1985	1	3		3	4		10	5				4	4	3	5	5	47	pos	mixed	6 wk	100
Vasseljen	1995	2	3		3	4		10	5				4		5	5	5	46	equiv/neg	chronic neck pain	24 wk	24
Fitz-Ritson	1995	1	1	4	3	4		10	5				2	2	3	5	5	45	pos	chronic whiplash	8 wk	31
McKinney	1989	2	3	4				10	5			3	4	4	5		5	45	equiv/neg	acute whiplash	2 year	247
Jensen	1990	2	4	4		2		10	5				4	4	3		5	43	pos	chronic headache	8 wk	23
Nordemar	1981	1	3		3	4		10	5				4		3	5	5	43	pos	acute neck pain	12 wk	30
Lewith	1981	2	1	4	3	4		10			5		4		3	5		41	equiv/neg	mixed	2 wk	26
Loy	1983	2	2		3	2		10	5	5		5	4		3			41	pos	chronic neck pain	6 wk	60
Mealy	1986	2	3	4		2		5	5			3	4	4	3		5	40	pos	acute whiplash	8 wk	61
Carlsson	1990	2	2		3	2		10	5				6		3		5	38	pos	chronic headache	4wk/12wk	62
Coan	1982	1	2	4	3	4		5					6		3	5	5	38	pos	chronic neck pain	12 wk	30
Brodin	1983		2	4		2		5	5				4	4	3		5	34	pos	mixed	4 wk	71
Howe	1983	1	3	4				5					4	4	3		5	29	pos	mixed	3 wk	52
Takala	1994		3					5					4	4	3		5	24	equiv/neg	mixed	10 wk	45

Study population: (A) Homogeneity, (B) comparability of relevant baseline characteristics, (C) randomisation procedure, (D) drop-outs described for each study group separately, (E) percentage of loss to follow-up, <20% = 2 points, <10% = 4 points, (F) sample size: >50 subjects in smallest group = 8 p, >100 subjects in smallest group = 17 p, *Interventions:* (G) Interventions included and described, (H) pragmatic study, (I) co-intervention avoided, (J) placebo controlled. *Measurement of effect:* (K) Patients blinded, (L) relevant outcome measures, (M) blinded assessments, (N) adequate follow-up period. *Data presentation and analysis:* (O) Intention to treat analysis, (P) frequencies of most important outcomes presented for each treatment group.

Table II. *Presentation of inclusion criteria, main intervention, reference treatment and outcome according to the Disablement Process and the Methodological quality for each separate trials on acute neck pain*

Author	Inclusion criteria	Main intervention	Reference treatment	Outcomes	Score
Acute neck problem, positive outcome, n = 3					
Foley-Nolan et al. 1992 (13)	**Pathology** acute whiplash	**Impairment** collar with pulsed electromagnetic therapy	**Not classified** collar with placebo units	**Impairment** pain, range of motion, analgesic consumption **Unclassified** subjective assessment of progress	**62**
Nordemar & Thörner 1981 (34)	**Impairment** acute neck pain <3 days	**Impairment** TNS	**Impairment** (1) manual treatment **Not classified** (2) neck collar	**Impairment** range of motion, pain at rest and in motion	**43**
Mealy et al. 1986 (32)	**Pathology** acute whiplash	**Impairment** active treatment; mobilisation, exercises	**Not classified** collar, advice to rest	**Impairment** range of motion, pain intensity	40
Acute neck problem, equivalent/negative outcome, n = 1					
McKinney et al. 1989 (30)	**Pathology** acute whiplash	**Impairment** active treatment: active and passive movements, traction, hydrotherapy, posture	**Impairment** (1) mobilization advice **Not classified** (2) rest and analgesics	**Impairment** range of motion, intensity of neck pain	**45**

Table III. *Presentation of inclusion criteria, main intervention, reference treatment and outcome according to the Disablement Process and the Methodological quality for each separate trial on chronic neck pain*

Author	Inclusion criteria	Main intervention	Reference treatment	Outcomes	Score
Chronic neck problem, positive outcome, n = 10					
Boline et al. 1995 (3)	**Impairment** tension-type headache >3 months	**Impairment** manipulation	**Not classified** medication: amitriptyline	**Impairment** headache frequency, over-the-counter medication usage **Imp + Funct + Disab** headache pain intensity, SF-36	**62**
Trock et al. 1994 (45)	**Pathology** cervical pain and stiffness ≥1 year, osteoarthritis, radiographic evidence of disk space narrowing with osteophyte formation	**Impairment** pulsed electromagnetic fields	**Not classified** placebo electromagnetic therapy	**Impairment** pain intensity, pain on motion, tenderness **Disability** activity of daily living **Unclassified** subjective improvement, objective global improvement	**61**
Petrie & Langley 1983 (35)	**Impairment** chronic neck pain >2 years	**Impairment** acupuncture (no electrostim)	**Not classified** placebo TNS	**Impairment** pain severity	**47**
Revel et al. 1994 (37)	**Impairment** chronic neck pain >3 months	**Impairment** rehabilitation program; improve neck-proprioception	**Not classified** symptomatic treatment (NSAID, analgesic)	**Impairment** pain intensity, range of motion, analgesic consumption, cervicocephalic kinesthesia **Functional limitation** pat assessment of functional improvement	**47**
Fitz-Ritson 1995 (11)	**Pathology** whiplash, pain >12 wk	**Impairment + Functional limitation** phasic neck exercises	**Impairment** rehabilitation exercises	**Imp + Funct + Disab** neck pain disability index	**45**
Jensen et al. 1990 (20)	**Impairment** headache 9–12 months post-traumatic symptoms	**Impairment** manual therapy	**Not classified** cold packs	**Impairment** intensity of headache, analgesics, associated symptoms such as dizziness, tinitus, etc	**43**
Loy 1983 (29)	**Pathology** cervical spondylosis confirmed by X-ray	**Impairment** electro-acupuncture	**Impairment** short-wave and traction	**Impairment** range of motion **Unclassified** subjective symptomatic relief	**41**

Coan et al. 1982 (9)	**Impairment** neck pain and/or radicular pain >6 months	**Impairment** acupuncture with or without electrostim	**Not classified** control group	**Impairment** pain score, analgesic consumption, pain hours per day **Disability** limitation of activity	**38**
Carlsson et al. 1990 (7)	**Impairment** chronic tension headache	**Functional limitation** active physiotherapy; the goal was to have control of the physical tension	**Impairment** acupuncture, electrostim	**Impairment** range of motion, muscle tenderness due to levels of pain, headache intensity	**38**
Chronic neck problem, equivalent/negative outcome, n = 3					
Klaber Moffett et al. 1990 (22)	**Impairment** neck and arm pain >3 months	**Impairment** traction, static	**Not classified** placebo-traction	**Impairment** pain intensity, range of motion **Disability** sleep disturbance, social dysfunction, activity of daily living	**60**
Petrie & Hazleman 1986 (36)	**Impairment** chronic neck pain >6 months	**Impairment** acupuncture (no electrostim)	**Not classified** sham TNS	**Impairment** pain intensity, McGill Pain Questionnaire, range of motion, patient pain rating scale, analgesic consumption **Disability** disability scores	**55**
Vasseljen et al. 1995 (46)	**Impairment** neck and shoulder pain >6 months	**Impairment + Functional limitation** indiv. based physiotheraphy treatment	**Impairment + Functional limitation** group exercise	**Impairment** pain intensity, perceived general tension, pressure pain sensitivity, palpated trigger points, EMG; m. trapetzius	**46**

Table IV. *Presentation of inclusion criteria, main intervention, reference treatment and outcome according to the Disablement Process and the Methodological quality for each separate trials on mixed neck pain.*

Author	Inclusion criteria	Main intervention	Reference treatment	Outcomes	Score
Mixed, no specific criteria for acute or chronic conditions, positive outcome, n = 6					
Foley-Nolan et al. 1990 (12)	**Impairment** neck pain >8 weeks	**Impairment** collar with pulsed electromagnetic therapy	**Not classified** collar with a placebo facsimile unit	**Impairment** pain, range of motion, analgesic consumption **Unclassified** subj assessment of progress	**60**
Cassidy et al. 1992 (8)	**Impairment** mechanical neck pain with radiation	**Impairment** manipulation	**Impairment** mobilization	**Impairment** pain intensity, range of motion	**54**
Levoska & Keinänen-Kiukaanniemi 1993 (26)	**Impairment** neck and shoulder symptoms once a week or more, and feeling of disturbance of neck, shoulder symptoms and muscle spasm and tenderness	**Impairment** active physiotherapy	**Impairment** (1) passive physiotherapy **Not classified** (2) no treatment	**Impairment** max isom neck muscle strength, max isometric grip test, endurance forces in shoulder muscles, muscle tonus in neck and shoulder, tender points by palpation, tender points pressure threshold, cephalagy or neck-shoulder pain	**50**
Zylbergold & Piper 1985 (48)	**Impairment** disorders of the cervical spine	**Impairment** traction: static, intermittent, manual	**Impairment** neck care instruction, exercises, moist heat	**Impairment** pain intensity: McGill Pain-Questionnaire, range of motion	**47**
Brodin 1983 (6)	**Impairment** cervical pain and restricted movement	**Impairment** manual mobilization	**Not classified** (1) mock therapy (2) medication	**Impairment** pain level, mobility of cervical spine	**34**
Howe et al. 1983 (19)	**Impairment** pain in neck, arm or hand and reduced movement	**Impairment** manipulation	**Not classified** control group	**Impairment** range of motion, affected symptoms as; neck pain, stiffness, headache, pain in shoulder or arm	**29**

Mixed, no specific criteria for acute or chronic conditions, equivalent/negative outcome, n = 5					
BAPM 1966 (5)	**Impairment** pain in neck and arm, root distribution or pain in neck and arm, full root distribution, or pain or paraesthesiae in neck and arm, evidence of clinical abnormality	**Impairment** traction; static	**Impairment** (1) positioning **Not classified** (2) collar (3) placebo tablets (4) placebo heat	**Impairment** pain, range of motion **Disability** ability to work, sleep disturbance **Not classified** obj assessm of severity	**52**
Goldie & Landquist 1970 (14)	**Impairment** cervical pain radiating into the arm	**Impairment** isometric exercise	**Impairment** (1) traction, intermittent **Not classified** (2) control	**Impairment** mobility of the neck **Unclassified** obj overall improvement, subj overall improvement	**49**
Sloop et al. 1982 (41)	**Impairment** neck pain >1 month	**Impairment** manipulation	**Not classified** diazepam	**Impairment** pain, pain level, range of motion **Disability** activities: sleep, work	**48**
Mixed, no specific criteria for acute or chronic conditions, equivalent/negative outcome					
Lewith & Machin 1981 (27)	**Pathology** cervical osteoarthrosis and neck pain >1 month	**Impairment** infrared light (local heat)	**Not classified** mock TNS	**Impairment** pain score, analgesic consumption **Disability** sleep disturbance	**41**
Takala et al. 1994 (42)	**Impairment** frequent neck symptoms	**Functional limitation** group gymnastics	**Not classified** control group	**Impairment** neck pain, pressure pain threshold	**24**

Effect size

Effect size was calculated for 17 studies (3, 6–9, 20, 22, 26, 30, 32, 34, 36, 37, 41, 45, 46, 48), and MQ scores were ≥50 for six of these (3, 8, 22, 26, 36, 45). The formula we used to determine effect size was based on the standard deviation (SD), which was either not included in or could not be calculated for ten of the studies.

The most frequently used outcomes were pain, $n = 16$ (3, 6–9, 20, 22, 26, 32, 34, 36, 37, 41, 45, 46, 48) and range of motion, $n = 8$ (8, 22, 30, 32, 34, 36, 37, 48). Assessment of the activities of daily life was only used in four studies (9, 22, 36, 45). These three outcome measurements were included in a regression analysis of methodological quality, study population and follow-up period. Various other types of outcome measurements were also used, but only in one study each.

In all of the studies, outcome was measured both before and after the treatment period. Follow-ups were conducted in eight of the 17 studies for which we calculated effect size. This was done within 1 month in four of the eight investigations (3, 30, 36, 45); at least 1 month but not later than 3 months after treatment in three of the studies (20, 22, 34); and after 6 months in one of the investigations (46).

Information regarding median effect size, including the interquartile range for pain, range of motion, and activity of daily living, is presented in Table V. To be able to compare the assessment of neck mobility, a summary of the total range of motion was calculated for each study that had presented the results in separate movements. Effect size values for the experimental groups were positive for pain, range of motion, and daily activities; however, only the effect size for pain and activity of daily living was moderate, whereas the effect size for range of motion was small.

There was no relation between pain and the size of the study population or the time point of the assessment, whereas there was a negative relation ($p < 0.05$) between pain and the methodological quality of the studies (Fig. 1), i.e. the higher the MQ scores, the lower the effect size ($R^2 = 0.25$). No relation was found between range of motion and either the study population or the methodological quality. However, we did observe a tendency towards a negative relation between range of motion and time for follow-up, i.e. the later the follow-up, the lower the effect size for range of motion ($R^2 = 0.14$).

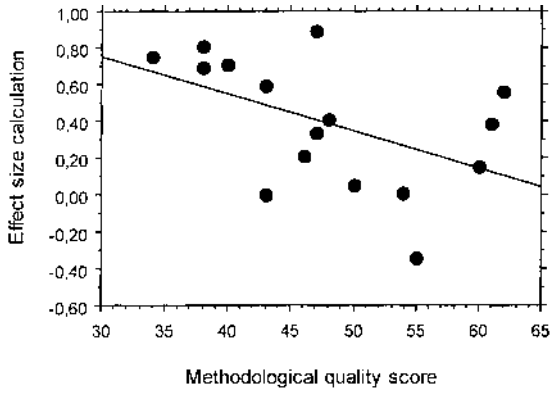

Fig. 1. Effect size for pain outcome correlated to methodological quality scores for 16 of the analysed randomised clinical trials. $R^2 = 0.25$.

Table V. *Randomised neck studies in the order of their effect size calculation*

Outcome	Median	IQR	*n*
Pain	0.4	0.59	16
ROM	0.1	0.37	8
ADL	0.3	0.37	4

Calculation of the medians with interquartile range (IQR). ROM = range of motion, ADL = activity of daily living.

Disablement process

Inclusion criteria. Inclusion criteria were defined in terms of impairment in 20 of the studies and were based on pathology in seven. No study used functional limitation or disability to describe the study population (Tables II–IV).

Intervention. The main intervention was never described in terms of pathology or disability but was focused on impairment in 23 of the studies and on functional limitation or a combination of impairment and functional limitation in the other four.

Outcome measurements. All but two studies used more than one measure of outcome.

Impairment. The outcome of impairment was ascertained in 22 of the studies by using the following pain assessment tools: the Visual Analogue Scale (VAS), (12, 13, 20, 22, 30, 32, 34, 36, 37, 41, 42, 45, 46), the Numerical Rating Scale (NRS-101), (8), the McGill Pain Questionnaire (36, 48), and different pain scores (6, 9, 27, 35). A pain diary was used in two investiga-

tions, combined with or including pain intensity, disability, and pill count (3, 36); headache frequency was also considered in one of the studies (3). Pain was even assessed subjectively by determining headache intensity (3, 7), consumption of analgesics (3, 9, 12, 13, 20, 37), and pain during passive motion (45). Moreover, outcome measurement of impairment was based on neck movement in 17 studies. In addition to the measurements mentioned above, a broad variety of single measures of impairment were used, each in only one or two studies (Tables II–IV).

Functional limitations. The outcome of functional limitations was measured in only one study (37). Functional improvement was measured by head repositioning accuracy; i.e. the ability to relocate accurately the head on the trunk after an active movement in the horizontal plane.

Disability. The following measures were used to assess disability outcomes: daily disability score (36), ability to work (5), sleep disturbance (5, 22, 27), social dysfunction (22), limitation of activity (9), and activity of daily living (22, 41, 45).

Outcome measurements referring to several components. Some of the investigators used assessments of headache intensity as an outcome (3, 7) where no difference was made between the components; impairment, functional limitation and disability of the Disablement Process due to the alternatives in the same item. The SF-36 questionnaire, which is a measurement of self-perceived health used in one study (3), and the Neck Disability Index score (11) also comprise items from three different components: impairment + functional limitation + disability.

Unclassifiable outcome measurements. The following outcome measurements could not be assigned to any of the categories in the Disablement Process: physician's assessment of the severity of the condition (no particular outcome [5]), global assessment of progress over a specific period of time (patients were asked to quantify *any* improvement [12, 13]), and overall improvement (14, 29, 45).

DISCUSSION

Our database search covering the period 1996–1995 and focusing on randomised clinical trials on neck pain patients who had undergone physiotherapy or chiropractic, yielded 27 publications. Despite our effort to be thorough, we may have overlooked some references, which could have influenced our results.

We found some evidence of positive outcomes based on the authors' conclusions in studies concerning the treatment of both acute and chronic neck pain, whereas findings were ambiguous in investigations that did not use specific inclusion criteria for acute or chronic neck problems.

It was difficult to draw any real conclusion about the efficacy of the interventions used in the analysed trials due to lack of long-term results. Only five studies conducted a follow-up at 6 months or more from start, and four of those reported equivalent/negative outcomes. Furthermore, deriving conclusions was difficult because some papers compared two treatments modalities, both supposed to have positive effects on neck pain. That might also explain the low effect size values in some studies. Considering pain, the effect size was 0.12 when calculations were based solely on trials with two active therapies (six papers) and 0.58 when examining studies that used a placebo reference treatment (10 papers). Comparison of active therapies can be of value when adding a new treatment to existing treatments (40). Conclusions about interventions can be strengthened by restricting examination to papers of high methodological quality. Of the 27 trials we analysed, nine had MQ scores $\geq$50, and seven of those had used placebo treatment in a comparison group. Considering different types of interventions in the investigations with scores $\geq$50, positive outcomes had been concluded by the authors for pulsed electromagnetic therapy (three trials), manipulation (two trials), and active physiotherapy (one trial); similar results (except regarding manipulation) were obtained in a meta-analysis performed by Aker et al. (1). We noted an author-based equivalent/negative outcome for cervical traction in two studies, and this has been confirmed in other meta-analyses (1, 18). Compared to placebo treatment, acupuncture gave equivalent/negative outcomes in one trial. In general, for most of the intervention groups, an insufficient number of papers was analysed to allow us to draw conclusions.

The methodological quality was low in most of the studies we analysed. The maximum MQ score of 100 points is probably difficult to obtain in clinical studies with this kind of analysis, therefore we have used a minimum MQ score of 50 points to indicate a good and high methodological quality (23, 24). Two-thirds of the studies we examined scored less than 50 points, indicating that the published results and our assessments of the efficacy of the interventions are not completely reliable, especially for trials with short follow-up periods. In some studies, outcome was assessed only once before

and once directly after a single treatment, hence it can be questioned whether the outcome was due to the actual treatment or simply the therapeutic contact.

The aim of our critical review was to determine the quality of research studies and the effectiveness of different kinds of treatment. Unfortunately, the systematic analysis model has many flaws, for example several criteria depend on whether or not (as well as how) special procedures such as randomisation are described. All of the authors of the papers we examined stated that their studies were randomised, but, in our evaluation, the articles that did not describe the randomisation procedure in detail were given lower MQ scores. A possible shortcoming of our quality analysis is that it illustrates better how the papers were written than the actual methodology. In some cases, the instructions for authors stipulated by the journals in question influenced the way the investigators reported materials and methods (e.g. assessment tools and study populations). We could have addressed this problem by contacting the authors, but such a task would have been time consuming. Obviously, study design, interventions, and outcome measurements must be clearly described if a reader is to be able to draw conclusions.

In our investigation, earning MQ points for measurement of effect required a thorough description of methodology and a clear presentation of the results of assessments performed. Accordingly, some of the analysed trials did not receive any points because, although they did mention that different measurements had been used, they did not report the results.

The reliability of the systematic analysis depends on the aim of the reviewers. We focused on what the authors of the analysed papers considered to be the main interventions, whereas other systematic analyses have purposely been limited to a single intervention (e.g. traction or exercise). This can influence the classification of outcomes as positive or equivalent/negative, as is evident when considering three systematic analyses (1, 18, 23) that examined some of the papers used in our review (5, 6, 14, 19, 32, 34, 41, 48). Another problem when comparing systematic analyses is the validity of the methods used, which depends on how the reviewers interpreted each item in the scheme (18, 23).

We used five outcome variables for neck patients that had been selected in previously published systematic analyses (23, 24): pain, global assessment of improvement, activity of daily living, range of motion, and consumption of analgesics. Determination of effect size and application of The Disablement Process revealed that pain, range of motion, and activity of daily living were used most often in the studies we analysed. Six trials used outcome measurements which were not scored in our analysis: muscle tenderness (7, 26, 45, 46), pressure pain threshold (42, 46), muscle strength (26), isometric grip test (26), endurance forces (26), electromyography (46), and associated symptoms (dizziness, visual disturbances, ear symptoms [20]). Even if we had included these measurements, our findings would have been essentially the same, because only one more paper (46) would have reached a score of $\geq$50, and it would have strengthened the equivalent/negative results concerning chronic neck problem and active physiotherapy.

In our review, most of the papers published after 1990 had higher methodological quality scores than those published before 1990, which seems to indicate that authors and journals are beginning to place greater emphasis on presentation of study design, methods, and results. However, in a previous systematic analysis concerning low back pain (25), no such association was noted between year of publication and methodological quality.

Our analysis of methodological quality gave positive results for a majority of the trials examined; more precisely, treatment outcome was better for experimental groups than for reference groups. The same was true for effect size, which was positive for the outcome measurements used most often in the studies, i.e. pain, range of motion, and activity of daily living; those findings imply that physiotherapy or chiropractic (the main forms of treatment in the trials) had a positive effect on neck pain. Positive effect size in connection with manual treatment of pain has also been reported by Aker et al. (1).

Studies with high methodological quality and large population might more often be published independent the outcome is positive or negative. We can see an effect of this by the fact that the effect size for pain was lower when related to high MQ score.

The Disablement Process analysis showed that inclusion criteria, intervention, and outcome measurements concerned impairment in most of the papers we examined. Although it is important to measure the same component that the intervention is meant to influence to decide if it is effective, it might be of even greater interest to analyse whether the intervention has influenced and improved the patient's daily living activities. Thus the outcomes of several components of the Disablement Process can be used to illustrate relation-

ships between the three different components, i.e. impairment, functional limitation and disability.

CONCLUSION

Our analyses demonstrate that few randomised clinical trials on neck problems are of high methodological quality and comprise a sufficiently long follow-up time. In the studies that did show higher quality, three different interventions led to a slight tendency towards positive results, but the number of publications considered was inadequate to allow general conclusions to be drawn. The effect size calculations and the disablement process analysis indicated that the interventions in the trials had a positive influence on two impairment components, pain and range of motion. Effect size was also positive for one disability component, activities of daily living, but this finding was based on a very limited number of investigations. Further analyses are needed to determine whether physiotherapy and chiropractic treatment have positive effects on functional limitation and various aspects of disability.

ACKNOWLEDGEMENT

This study was supported by the Swedish Council for Work Life Research.

REFERENCES

1. Aker, P. D., Gross, A. R., Goldsmith, C. H. & Peloso, P.: Conservative management of mechanical neck pain: systematic overview and meta-analysis. BMJ *313:* 1291–1296, 1996.
2. Anderson, R., Meeker, W. C., Wirick, B. E., Mootz, R. D., Kirk, D. H. & Adams, A.: A meta-analysis of clinical trials of spinal manipulation. J Manipulative Physiol Ther *15:* 181–194, 1992.
3. Boline, P. D., Kassak, K., Bronfort, G., Nelson, C. & Anderson, A. V.: Spinal manipulation vs. amitriptyline for the treatment of chronic tension-type headaches: A randomised clinical trial. J Manipulative Physiol Ther *18:* 148–154, 1995.
4. Bovim, G., Schrader, H. & Sand, T.: Neck pain in the general population. Spine *19:* 1307–1309, 1994.
5. British Association of Physical Medicine: Pain in the neck and arm: a multicentre trial of the effects of physiotherapy. BMJ *1:* 253–258, 1966.
6. Brodin, H.: Cervical pain and mobilization. Med Phys *6:* 67–72, 1983.
7. Carlsson, J., Fahlcrantz, A. & Augustinsson, L.-E.: Muscle tenderness in tension headache treated with acupuncture or physiotherapy. Cephalalgia *10:* 131–141, 1990.
8. Cassidy, J. D., Lopes, A. A. & Yong-Hing, K.: The immediate effect of manipulation versus mobilization on pain and range of motion in the cervical spine: A randomised controlled trial. J Manipulative Physiol Ther *15:* 570–575, 1992.
9. Coan, R. M., Wong, G. & Coan, P. L.: The acupuncture treatment of neck pain: A randomised controlled study. Am J Chin Med *9:* 326–332, 1982.
10. Ekberg, K., Björkqvist, B., Malm, P., Bjerre-Kiely, B. & Axelson, O.: Controlled two year follow up of rehabilitation for disorders in the neck and shoulders. Occup Environ Med *51:* 833–838, 1994.
11. Fitz-Ritson, D.: Phasic exercises for cervical rehabilitation after "whiplash" trauma. J Manipulative Physiol Ther *18:* 21–24, 1995.
12. Foley-Nolan, D., Barry, C., Coughlan, R. J., O'Connor, P. & Roden, D.: Pulsed high frequency (27MHz) electromagnetic therapy for persistent neck pain. A double blind, placebo-controlled study of 20 patients. Orthopedics *13:* 445–451, 1990.
13. Foley-Nolan, D., Moore, K., Codd, M., Barry, C., O'Connor, P. & Coughlan, R. J.: Low energy high frequency pulsed electromagnetic therapy for acute whip-lash injuries. A double blind randomised controlled study. Scand J Rehab Med *24:* 51–59, 1992.
14. Goldie, I. & Landquist, A.: Evaluation of the effects of different forms of physiotherapy in cervical pain. Scand J Rehab Med *2–3:* 117–121, 1970.
15. Gore, D. R., Sepic, S. B., Gardner, G. M. & Murray, M. P.: Neck pain: A long-term follow-up of 205 patients. Spine *12:* 1–5, 1987.
16. Hagberg, M. & Wegman, D. H.: Prevalence rates and odds ratios of shoulder-neck diseases in different occupational groups. Br J Ind Med *44:* 602–610, 1987.
17. Hagberg, M. & Hogstedt, C. Stockholmsundersökningen 1 MUSIC Books 1991, Stockholm..
18. Heijden van der, G. J. M. G., Beurskens, A. J. H. M., Koes, B. W., Assendelft, W. J. J., de Vet, H. C. W. & Bouter, L. M.: The efficacy of traction for back and neck pain: A systematic, blinded review of randomised clinical trial methods. Phys Ther *75:* 93–104, 1995.
19. Howe, D. H., Newcombe, R. G. & Wade, M. T.: Manipulation of the cervical spine—a pilot study. J R Coll Gen Pract *33:* 574–579, 1983.
20. Jensen, O. K., Nielsen, F. F. & Vosmar, L.: An open study comparing manual therapy with the use of cold packs in the treatment of post-traumatic headache. Cephalalgia *10:* 241–250, 1990.
21. Jette, A. M.: Physical disablement concepts for physical therapy research and practice. Phys Ther *74:* 380–386, 1994.
22. Klaber Moffett, J. A., Hughes, G. I. & Griffiths, P.: An investigation of the effects of cervical traction. Part 1: Clinical effectiveness. Clin Rehabil *4:* 205–211, 1990.
23. Koes, B. W., Assendelft, W. J. J., Heijden, G. J. M. G., van der Bouter, L. M. & Knipschild, P. G.: Spinal manipulation and mobilization for back and neck pain. A blinded review. BMJ *303:* 1298–1303, 1991.
24. Koes, B. W., Bouter, L. M., Beckerman, H., Heijden van der, G. J. M. G. & Knipschild, P. G.: Physiotherapy exercises and back pain. A blinded review. BMJ *302:* 1572–1576, 1991.
25. Koes, B. W., Bouter, L. M. & Heijden van der, G. J. M. G.: Methodological quality of randomised clinical trials on treatment efficacy in low back pain. Spine *20:* 228–235, 1995.
26. Levoska, S. & Keinänen-Kiukaanniemi, S.: Active or passive physiotherapy for occupational cervicobrachial disorders? A comparison of two treatment methods with a

1-year follow-up. Arch Phys Med Rehabil *74:* 425–430, 1993.
27. Lewith, G. T. & Machin, D.: A randomised trial to evaluate the effect of infra-red stimulation of local trigger points, versus placebo, on the pain caused by cervical osteoarthrosis. Acupunct Electrother Res *6:* 277–284, 1981.
28. Linton, S. J.: Risk factors for neck and back pain in a working population in Sweden. Work Stress *4:* 41–49, 1990.
29. Loy, T. T.: Treatment of cervical spondylosis. Electroacupunture versus physiotherapy. Med J Aust *2:* 32–34, 1983.
30. McKinney, L. A., Dornan, J. O. & Ryan, M.: The role of physiotherapy in the management of acute neck sprains following road-traffic accidents. Arch Emerg Med *6:* 27–33, 1989.
31. McKinney, L. A.: Early mobilisation and outcome in acute sprains of the neck. BMJ *299:* 1006–1008, 1989.
32. Mealy, K., Brennan, H. & Fenelon, G. C. C.: Early mobilisation of acute whiplash injuries. BMJ *292:* 656–657, 1986.
33. Nagi, S.: Disability concepts revisited: implications for prevention. *In* Disability in America: Toward a national agenda for prevention (ed. A. M. Pope & A. R. Tarlov), pp. 309–327. National Academy Press, Washington, D.C. 1991.
34. Nordemar, R. & Thörner, C.: Treatment of acute cervical pain—a comparative group study. Pain *10:* 93–101, 1981.
35. Petrie, J. P. & Langley, G. B.: Acupuncture in the treatment of chronic cervical pain. A pilot study. Clin Exp Rheumatol *1:* 333–336, 1983.
36. Petrie, J. P. & Hazleman, B. L.: A controlled study of acupuncture in neck pain. Br J Rheumatol *25:* 271–275, 1986.
37. Revel, M., Minguet, M., Gergoy, P., Vaillant, J. V. & Manuel, J. L.: Changes in cervicocephalic kinesthesia after a proprioceptive rehabilitation program in patients with neck pain: A randomised controlled study. Arch Phys Med Rehabil *75:* 895–899, 1994.
38. Rosenthal, R. Meta-analytic procedures for social research. Beverly Hills, CA: Sage publication, 1984: (applied social research methods series, vol 6).
39. Shekelle, P. G., Adams, A. H., Chassin, M. R., Hurwitz, E. L. & Brook, R. H.: Spinal manipulation for low back pain. Ann Inter Med *117:* 118–126, 1992.
40. Skargren, E. I., Öberg, B. E., Carlsson, P. G. & Gade, M.: Cost and effectiveness analysis of chiropractic and physiotherapy treatment for low back pain and neck pain. Six months follow-up. Spine *22:* 2167–2177, 1997.
41. Sloop, P. R., Smith, D. S., Goldenberg, E. & Doré, C.: Manipulation for chronic neck pain, a double-blind controlled study. Spine *7:* 532–535, 1982.
42. Takala, E.-P., Viikari-Juntura, E. & Tynkkynen, E.-M.: Does group gymnastics at the workplace help in neck pain? A controlled study. Scand J Rehab Med *26:* 17–20, 1994.
43. Ter Riet, G., Kleiknen, J. & Knipschild, P.: Acupuncture and chronic pain: a criteria-based meta-analysis. J Clin Epidemiol *43:* 1191–1199, 1990.
44. Thomas, J. & Nelson, J. Research methods in physical activity. 2nd ed. Human Kinetics, Illinois, 1985.
45. Trock, D. H., Bollet, A. J. & Markoll, R.: The effect of pulsed electromagnetic fields in the treatments of osteoarthritis of the knee and cervical spine. Report of randomised, double blind, placebo-controlled trials. J Rheumatol *21:* 1903–1911, 1994.
46. Vasseljen, O., Mörk Johansen, B. & Westgaard, R. H.: The effect of pain reduction on perceived tension and EMG-recorded trapezius muscle activity in workers with shoulder and neck pain. Scand J Rehab Med *27:* 243–252, 1995.
47. Verbrugge, L. M. & Jette, A. M.: The Disablement Process. Soc Sci Med *38:* 1–14, 1994.
48. Zylbergold, R. S. & Piper, M. C.: Cervical spine disorders, a comparison of three types of traction. Spine *10:* 867–871, 1985.

Accepted November 24, 1998

Address for offprints:

Görel Kjellman
Faculty of Health Sciences
Department of Neuroscience and Locomotion; Physiotherapy
SE-581 85 Linköping
Sweden