The clinical course and prognostic factors of non-specific neck pain: a systematic review

Borghouts, J.A.J.; Koes, B.W.; Bouter, L.M.

published in
Pain
1998

DOI (link to publisher)
10.1016/S0304-3959(98)00058-X

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

PAIN

Pain 77 (1998) 1–13

Review Article

The clinical course and prognostic factors of non-specific neck pain: a systematic review

Jeroen A.J. Borghouts*, Bart W. Koes, Lex M. Bouter

Faculty of Medicine, Institute for Research in Extramural Medicine, Vrije Universiteit, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

Received 24 October 1997; received in revised form 18 February 1998; accepted 18 February 1998

Abstract

Neck pain occurs frequently in western societies. In the majority of cases, no specific cause can be identified. In order to gain insight into the clinical course and prognostic factors of non-specific neck pain, a systematic review was conducted. A computerized literature search was carried out to identify observational studies on non-specific neck pain and randomized clinical trials (RCTs) on conservative treatment of non-specific neck pain. Two reviewers scored independently, the methodological quality of all identified publications, using a standardized set of 13 criteria which were divided into five categories according to: study population, study design, follow-up, outcome measures and analysis/data presentation. To determine prognosis per study, an overall percentage of recovery for the most important outcome measures (pain, general improvement, functional status, health care utilization and lost days of work) was calculated. In total 23 eligible publications were identified (six observational studies and 17 RCTs). Only seven of 23 studies scored 50% or more of the 13 items, indicating a generally poor quality of methods. The most prevalent methodological shortcomings appeared to be selection of the study population, the sample size and analysis techniques. Most information regarding the clinical course is available for the group of patients with complaints for more than 6 months, who are treated in a secondary care or an occupational setting. In this group of patients, 46% (median) had less pain, with a range of 22–79% and a general improvement that ranged between 37 and 95% (47% median). The reduction in the use of analgesics ranged between 32 and 80% (37% median). Six studies reported on prognostic factors. Bearing in mind the limited number of studies and the low methodological quality, there are some indications that the localization (radiation to the arms/neurologic signs) and radiologic findings (degenerative changes in the discs and joints) are not associated with a worse prognosis. A higher severity of pain and a history of previous attacks however, seems to be associated with a worse prognosis.

Keywords: Neck Pain; Systematic review; Prognostic factors; Clinical course

1. Introduction

Neck pain occurs frequently in western societies (Andersson, 1997). A large epidemiologic study in The Netherlands reported a life-time prevalence of neck pain in 30% of the male and in 43% of the female participants. At the moment of questioning in this cross-sectional study, about 10% of the males and 18% of the females reported to have neck complaints (Valkenburg et al., 1980).

The pain may arise from any of the structures in the neck. These include the intervertebral discs, ligaments, muscles, facet joints, dura and nerve roots (Bogduk, 1988). There are a large number of potential causes of neck pain. These vary from tumours, trauma (e.g. fractures, whiplash), infection, inflammatory disorders (e.g. rheumatoid arthritis) and congenital disorders. In most cases, however, no systemic disease can be detected as underlying cause of the complaints. This group consists of patients with mainly mechanical disorders including degenerative changes and could be labelled as non-specific neck pain (Bogduk, 1984).

To inform patients, doctors and policy makers about the outcome of non-specific neck pain, information is needed

* Corresponding author. Tel.: +31 20 4448187; fax: +31 20 4448181; e-mail: jhj.borghouts.emgo@med.vu.nl

0304-3959/98/$19.00
PII S0304-3959(98)00058-X

about the prognosis. This paper is therefore focusing on the clinical course of non-specific neck pain and the factors predicting recovery of non-specific neck pain.

Studies on clinical course assess the course of a disease subsequent to diagnosis and initiation of treatment (Von Korff, 1994). Prognostic factors can potentially predict the future course subsequent to disease onset, while etiologic factors (or risk factors) are associated with the onset of the disease. Prognostic factors need not necessarily be causally related to the outcomes; the presence of a strong association is all that matters (Laupacis et al., 1994).

In order to gain insight in the clinical course and prognostic factors, we systematically reviewed the available studies on this topic.

2. Methods

2.1. Study selection

For our systematic review a computerized literature search was carried out using two search strategies. One strategy focusing on the identification of observational studies and the other on the identification of RCTs.

To identify the observational studies the following keywords were used: MEDLINE (1966–1996) [MesH] neck, neck muscles, cervical vertebrae, occupational diseases, musculoskeletal diseases, prognosis, cohort studies, longitudinal studies, follow-up studies, prospective studies, case-control studies, retrospective studies; EMBASE (1988–1996) [MesH] neck, neck pain, neck injury, neck muscle, cervical spine injury, prognosis, disease course. To identify the RCTs we have used the search strategy as described by Dickersin et al. (1994). This search strategy was combined with keywords for neck (as described above) and keywords for conservative treatment as described in a review by Gross et al. (1996). In addition for both observational studies and RCTs, the references given in relevant publications were further examined. Abstracts and unpublished studies were not included.

A study was included if: (1) the study population consisted of patients suffering from non-specific neck pain or musculoskeletal pain of which a subgroup of patients with neck pain was presented separately. Non-specific pain was defined as pain (with or without radiation) without a specific systemic disease being detected as underlying cause of the complaints. Neck was defined as: the cervical spine, occiput region, cervico thoracic junction and muscles originating from the cervical region acting on the head and shoulders; (2) the article was published in English, Dutch or German; (3) it concerned an observational study (prospective or retrospective) or a RCT.

A study was excluded if: (1) the study population concerned patients with specific underlying pathology such as: tumours, trauma (fractures), infection, inflammatory disorders (rheumatoid arthritis), osteoporosis etc. Whiplash also could be regarded as a non-specific diagnosis, because the term refers to the putative cause of complaints without specifying the patho-anatomical mechanism involved. However, due to its separate place in literature, the need for specific outcome measures in this group of patients and problems with interpretation of outcome due to litigation etc, we excluded whiplash studies as well; (2) it was a cross-sectional study (without follow-up); (3) the total duration of the study was less than 3 weeks (including the intervention period).

2.2. Quality assessment

Two reviewers (JAJB and BWK) scored independently, the quality of each study, according to a standardized set of predefined criteria (Table 1). The criteria were adapted from Von Korff (1994), Sackett et al. (1991) and Cole and Hudak (1996) and modified to cover the topic of our review.

The quality of a study can be described in terms of internal and external validity. Studies on the clinical course and prognostic factors, should not only be of high methodological quality (internal validity), but also be informative (external validity). Contrary to cause-effect research in which internal validity is the most important, representativeness and generalization are also of great importance in descriptive epidemiologic studies (Bouter and van Dongen, 1991). Consequently, our criteria list covers both types of validity.

Thirteen criteria were divided into five categories according to: study population, study design, follow-up, outcome measures and analysis/data presentation. The criteria are

Table 1

Criteria list for assessment of the quality of studies on the clinical course and prognostic indicators of non-specific neck pain[a]

Criteria	Score
Study population	
(A) Selection of study population	+/–
(B) Description of inclusion and exclusion criteria	+/–
(C) Description of potential prognostic factors	+/–
Study design	
(D) Prospective study design	+/–
(E) Study size	
(a) Course cohort ≥100 patient-years	+/–
(b) Prognostic factors sub-groups ≥200 patients-years	+/–
Follow-up	
(F) Follow-up ≥12 months	+/–
(G) Follow-up	
(a) Drop-outs/loss to follow-up ≤20%	+/–
(b) Drop-outs/loss to follow-up ≤10%	+/–
(c) Information completers versus loss to follow-up/drop-outs	+/–
Outcome measures	
(H) Relevant outcome measures	+/–
Analysis and data presentation	
(I) Frequencies of most important outcome measures	+/–
(J) Appropriate analysis techniques	+/–

[a]A copy of the more detailed criteria list can be obtained from the first author on request.

described in more detail in a separate appendix (not included)[1]. Each item of a selected study which met our criteria, was assigned a '+' (positive). If the item did not meet our criteria or was insufficiently or not described at all, a '−' was assigned. The highest attainable overall score was 13 '+'. The overall score was used to assess the hierarchical order of the studies. Studies scoring 50% or more of the maximum attainable score were, arbitrarily, considered to be of 'high quality'. All studies scoring less then 50% were rated as 'low quality'.

2.3. Outcome of the studies

We divided the results of each study into two main categories: course of the complaints and prognostic factors. To determine recovery per study an overall percentage for the most important outcome measures (pain, general improvement, functional status, health care utilization and lost days of work) was calculated. Recovery was identified using reported categories like: 'less or no more pain', '(slight) improvement', 'no symptoms', etc. In studies with different (treatment) groups all groups were analyzed together as one cohort, using the longest follow-up. For example, from a group treated with medication, 50% of the patients had a decrease in pain. From the group receiving the placebo treatment, 30% of the patients had a decrease in pain. The overall outcome for the total number of patients was 50 + 30% divided by 2, which is 40%. Next, a range and median (using the overall percentages) was calculated for the number of studies reporting on the most important outcome measures. Prognostic factors were considered to be all factors influencing the clinical course as reported by the author.

3. Results

Twenty-three studies did meet our selection criteria. Table 2A represents detailed information regarding the observational studies (Gore et al., 1987; Abenhaim et al., 1988; Berg et al., 1988; Rossignol et al., 1988; Tellnes, 1989; Abbot et al., 1990; Takala et al., 1992) ($n = 6$). Two papers (Abenhaim et al., 1988; Rossignol et al., 1988) reported on the same observational study. Table 2B represents detailed information regarding the RCTs (Anonymous, 1966; Horvath and Fellmann, 1969; Goldie and Landquist, 1970; Nordemar and Thörner, 1981; Sloop et al., 1982; Howe et al., 1983; Loy, 1983; Petrie and Hazleman, 1986; Ceccherelli et al., 1989; Foley-Nolan et al., 1990; Coan et al., 1982; Thorsen et al., 1992; Levoska and Keinänen-Kiukaanniemi, 1993; Revel et al., 1994; Takala et al., 1994; Jensen et al., 1995; Vasseljen et al., 1995) ($n = 17$). Table 2A,B represent the studies in alphabetic order according to the first author.

[1] Available on request from the first author.

3.1. Methodological quality

The two reviewers scored 299 items. On 48 items (16%) there was disagreement, mostly due to reading and interpretation error. Forty-two percent of the disagreement was related to item G (drop-outs/loss to follow-up). The disagreements were resolved in a single consensus meeting. The methodological quality score of each study is shown in Table 2A,B and represents the percentage positive scored items (max 100%). Only seven of 23 RCTs scored 50% or more (≥7 items) of the maximum attainable score. None of the observational studies scored 50% or more. Eight studies scored 25% or less (≤3 items) out of 13 points (four RCTs, four observational studies).

3.1.1. Selection of study population

Only one study (Takala et al., 1992) identified patients at an early and uniform point (inception cohort) in the course of their disease.

3.1.2. Description of inclusion and exclusion criteria

Seven of the 24 studies formulated in- and exclusion criteria for age, duration of complaints and non-specific complaints.

3.1.3. Description of potential prognostic factors

Sixteen studies reported factors at baseline that could potentially serve as prognostic factors. Only six of these actually analyzed and reported prognostic factors and even then often no data were presented to support the conclusions.

3.1.4. Prospective study design

Only two studies (Gore et al., 1987; Abbot et al., 1990) did not use a prospective design. One study (Takala et al., 1992) used a partly prospective and partly retrospective design.

3.1.5. Study size

Four studies used a sufficient number of patients to determine course. Only Gore et al. (1987) used sufficient patients to enable subgroup comparisons.

3.1.6. Follow-up

Six times a follow-up of 12 months or more was used. Only one RCT (Anonymous, 1966) reported a follow-up of 12 months or more. Eleven studies reported a total number of drop-outs/loss to follow-up ≤10%. Several studies did not report on drop-outs/loss to follow-up.

3.1.7. Outcome measures

A great diversity of outcome measures was used. Pain and general improvement were reported most frequent as primary outcome measure (12 studies both). Most studies used a visual analogue scale (VAS) for measuring pain intensity. General improvement was usually measured as 'perceived

Table 2
Details of observational studies (A) and ranomized clinical trials (B) on the clinical course of non-specific neck pain

Author[a] (year)	Study population	Case definition	Study design	Treatment	Outcome measures	Results
A. Observational studies						
Abbot et al. (1990)[15]	Former patients of a pain clinic (Australia)	Patients with neck pain as presenting complaint, treated between 1972 and 1982 in a pain clinic. 37% had a motor vehicle accident	Retrospective (lifetime?) $n = 55$		Treatment prior to pain clinic (PT, medication, alt. treatment, surgical) Treatment at the clinic (11.4 months)	100%, with 61% no improvement 56% improvement
Berg et al. (1988)[15]	Male manual workers and office workers (Sweden)	Neck symptoms at the moment of questioning	Prospective 3 years $n = 21$ (retired men)		Retired men with fewer symptoms compared with before retirement: Manual workers Office workers Manual + office workers	 41% ($n = 15$) 40% ($n = 6$) 41%
Gore et al. (1987)[31]	Former patients of an orthopaedic surgeon (USA)	Patients with problems originating from the neck. No neck surgery, no objective neurologic deficits, no malignancies or rheumatoid arthritis	Retrospective ≥10 years $n = 205$ of which 76 had a motor vehicle accident		Received treatment % change in pain (less or no more) Change of jobs (quit or changed)	80% 79% 7%
Rossignol et al. (1988)/ Abenhaim et al. (1988)[23b]	Occupational neck injuries compensated in 1981 in Quebec (Canada)	Any musculoskeletal complaint relating to the spine (cervical reported separately for a few outcome measures), compensated for work absenteeism for at least 1 day in 1981	Prospective 3 years $n = 2342$ (of which 161 cases with cervical symptoms)		Mean cumulative duration of absence from work for 3 years Probability of being absent from work for 6 months or more during 3 years compared with lumbar. Recurrence rate during 3 years. Number of episodes among current cases. Recurrence in 3 years, compared with other back region	74.5 days (SE 12.6) RR = 0.76 38.5% 2.58 (SE 0.13) RR = 1.00
Takala et al. (1992)[38]	Female tellers of a bank (Finland)	Non-cases retrospective. Prospective: neck and shoulder symptoms	Retrospective 1 year (not analyzed) $n = 351$. Prospective 5.5 months $n = 138$		Percentage who changed to a lower category Symptoms: (0–7 days, 8–30 days, >30 days during previous 3 months)	36%
Tellnes (1989)[15]	Initial sickness certificates issued to residents of Buskerud during a period of 4 weeks (Norway)	ICHPPC − 1 codes for several diagnoses registered during a 4-week period including cervical spine syndromes (ICHPPC − 1: 7200)	Prospective 1 year ($n = 5042$) of which $n = 57$ were cervical spine syndromes		% still certified sick after 52 weeks	1.8%

B. Randomized clinical trials

Anonymous (1966)[38]	Patients attending the department of physical medicine (UK)	Pain in the neck and arm, with or without paraesthesia, root distribution Or: pain in the neck and arm of full root-distribution with paraesthesia (without clinical evidence) Or: pain or paraesthesia in the neck and arm of partial root-distribution with abnormality in the neck	RCT 6 months $n = 493$	(I) Traction, instruction in posture, collar, aspirin (II) Positioning (position with maximum relief of pain), instruction in posture, collar, aspirin (III) Collar, instruction in posture, aspirin (IV) Placebo heat (short-wave diathermy), aspirin (V) Placebo tablets, aspirin	After 4 weeks: Proportion showing improvement: Physicians' assessment Patients' assessment Patients' + physicians' assessment Increase in joint range (patients with initially restricted movements) Interference with work or stopping work (reduction) Sleep disturbed or seriously disturbed (reduction) After 6 months: Proportion of patients with pain decrease since follow-up 6 weeks	 21%(I), 23%(II), 24%(III), 21%(IV) 16%(V) 71%(I), 81%(II), 74%(III), 70%(IV) 56%(V) 46%(I), 52%(II), 49%(III), 46%(IV) 36%(V) Flexion: 11°, extension: 9°, rotation: 7° lat. flexion: 7° (all patients, groups not specified) 30%(I), 23%(II), 18%(III), 24%(IV) 20%(V) 35%(I), 34%(II), 27%(III), 34%(IV) 15%(V) 31% (all patients, groups not specified)
Coan et al. (1982)[62]	Public service announcements in newspapers (MD, USA)	Neck pain and/or radicular arm and hand pain present for at least 6 months	RCT 12 weeks $n = 30$	(I) Acupuncture (II) Control (no therapy)	Decrease in: Mean hours of pain Mean pain score Mean pain (hours + pain score) Mean pain pills Mean limitation of activity Proportion showing improvement	 68%(I), 0%(II) 40%(I), 2%(II) 54%(I), 1%(II) 54%(I), 10%(II) 32%(I), 12%(II) 80%(I), 13%(II)
Ceccherelli et al. (1989)[46]	Female (Italy)	Painful myofascial syndromes in the cervical region as a result of mild cervical arthrosis or poor posture	RCT 3 months $n = 27$	(I) Laser (II) Placebo laser	Mean pain decrease (McGill pain questionnaire): No. of words Total score No. of words + total score	 49% decrease (I), 2% increase (II) 55% decrease (I), 7% decrease (II) 52% decrease (I), 3% decrease (II)

Table 2 (*continued*)

Author[a] (year)	Study population	Case definition	Study design	Treatment	Outcome measures	Results
Foley-Nolan et al. (1990)[54]	Rheumatology outpatients or physiotherapy department (Ireland)	>18 years, >8 weeks neck pain, unresponsive to at least one course of NSAIDs	RCT Partial cross-over 6 weeks $n = 20$	(I) Neck collar connected with a unit of pulsed high frequency electromagnetic energy, continue NSAID use (II) Placebo neck collar (3 weeks) and 3 weeks like (I), continue NSAID use	Mean decrease: Pain (VAS) ROM Pill count Proportion showing improvement (patients' assessment on pain and ROM)	 71%(I), 55%(II) 33%(I), 39%(II) 80%(I), 80%(II) 100% (I), 90%(II)
Goldie and Landquist (1970)[31]	Patients at a department of orthopaedic surgery (Sweden)	Cervical pain radiating down either of the upper extremities following a segmental pattern	RCT 3 weeks $n = 73$	(I) Isometric exercises + analgesics + muscle relaxant (II) Traction + analg + muscle relaxant (III) No treatment + analg + muscle relaxant	Proportion showing improvement: Physicians' assessment Patients' assessment Patients' + physicians' assessment Mean increase in ROM (patients with initially restricted movements): Rotation Lat. flexion	 46%(I), 69%(II), 69%(III) 70%(I), 65%(II), 30%(III) 58%(I), 67%(II), 50%(III) 5°(I), 7°(II), 5°(III) 5°(I), 6°(II), 5°(III)
Horvath and Fellmann (1969)[8]	Patients of a rheumatology clinic (Switzerland)	Cervical syndrome	RCT 3 weeks $n = 40$	(I) Nifluril (II) Placebo Both groups received sollux therapy, massage and a thermal bath in addition	Mean decrease (3-point scale): Pain (spontaneously) Pain (provoked) Pain (spontaneously + provoked) Tension Mobility increase of the neck	 54%(I), 35%(II) 37%(I), 32%(II) 45%(I), 34%(II) 33%(I), 33%(II) 35%(I), 18%(II)
Howe et al. (1983)[15]	Patients attending the surgery of one two-man practice (UK)	15–65 years, pain in neck/arm or hand due to a lesion in cervical spine and evidence of reduced movement in one or more cervical intervertebral joints or palpable asymmetry of the transverse processes of the atlas	RCT 3 weeks $n = 52$	(I) Manipulation and/or injection + azapropazone (II) Azapropazone	Proportion showing improvement: Pain in neck Stiff neck Headache Pain/paraesthesia of shoulder Pain/paraesthesia of arm/hand Increase in ROM: Rotation Lat. flexion	 76%(I), 58%(II) 73%(I), 64%(II) 92%(I), 100%(II) 75%(I), 69%(II) 82%(I), 80%(II) 5°(I), not reported(II) 0°(I), not reported(II)

Jensen et al. (1995)[31]	Patients referred to an in-patient orthopaedic department (Sweden)	Neck and shoulder pain without objective neurological signs, 20–55 years, fluent in Swedish language Excluded comorbidity that could impair participation (e.g. heart condition, alcoholism)	RCT 6 months $n = 66$	(I) Physical fitness, health behaviour, plan for return to work (II) Like (I) + cognitive behavioural intervention	Mean: Pain decrease (VAS) Disability decrease Anxiety decrease Depression decrease % patients noticed: Pain decrease Disability decrease Anxiety decrease Depression decrease	 6%(I), 13%(II) 6%(I), 3%(II) 3%(I), 44%(II) 8%(I), 13%(II) 47%(I), 44%(II) 39%(I), 32%(II) 31%(I), 40%(II) 30%(I), 36%(II)
Levoska and Keinänen-Kiukaanniemi (1993)[46]	Female office workers employed by a bank or social insurance institution (Finland)	Neck or shoulder symptoms once a week or more, feeling of disturbance, muscle spasm and tenderness in neck/shoulder on palpation	RCT 1 year $n = 47$	(I) Passive physiotherapy (surface heat, massage, stretching, exercises) (II) Active physiotherapy (stretching, dynamic muscle training, home exercises)	Proportion showing improvement: Neck-shoulder pain Cephalagy or neck-shoulder pain Neck-shoulder + cephalgy or neck-shoulder pain	 14%(I), 18%(II) 18%(I), 36%(II) 16%(I), 27%(II)
Loy (1983)[15]	Consecutive patients of a general orthopaedic clinic (Hong Kong)	Cervical spondylosis Excluded patients with acute symptoms for a few days	RCT 6 weeks $n = 60$	(I) Traction + SWD (II) Electro-acupuncture	ROM % patients noticed benefit	Slight decrease (I), increase (II) 54%(I), 87%(II)
Nordemar and Thörner (1981)[46]	Consecutive patients of a department of physical medicine and medical rehabilitation (Sweden)	Patients with acute cervical pain, <3 days, without neurological symptoms	RCT 3 months $n = 30$	(I) Collar + rest + analg (II) TENS + collar + rest + analg (III) Collar + rest + analg + MT	Mean pain decrease (VAS) Mean decrease in ROM	100%(I), 100%(II), 100%(III) 100%(I), 100%(II), 100%(III)
Petrie and Hazleman (1986)[62]	Out-patients at a rheumatology clinic (UK)	Chronic neck pain, arising from the neck with or without radiation to the shoulders and/or occiput, present on a daily basis for at least 6 months	RCT 2 months $n = 26$	(I) Acupuncture (II) Placebo TENS	Mean decrease: Daily pain (VAS) Daily disability (VAS) Daily pill count MPQ % patients noticed benefit	 33%(I), 22%(II) 25%(I), 8% increase(II) 23%(I), 9% increase(II) 32%(I), 20%(II) 77%(I), 50%(II)
Revel et al. (1994)[62]	Out-patients at a department of rheumatology (France)	Chronic neck pain >3 months, considered for a medical program, ≥16 years	RCT 10 weeks $n = 60$	(I) Rehabilitation (improve neck proprioception by active and passive motion) + medication (NSAID, analgesic drugs) (II) Medication (NSAID, analgesic drugs)	Mean pain decrease (VAS) Mean decrease in daily medication (NSAID, analgesic drugs) Noticed benefit by patient Mean decrease in ROM Flexion/extension Rotation	43%(I), 9%(II) 79%(I), 37%(II) 60%(I), 27%(II) 2%(I), 1%(II) 0%(I), 0%(II)
Sloop et al. (1982)[54]	Patients referred to a department of medical rehabilitation (UK)	Cervical spondylosis or non-specific neck pain as described in standard texts, pain (duration ≥1 month)	RCT, pseudo cross-over (not improved patients crossed-over, not reported here) 3 weeks $n = 39$	(I) MT + diazepam (II) Diazepam	Percentage patients noticed benefit (pain, selected daily activities)	57%(I), 28%(II)

Table 2 (*continued*)

Author[a] (year)	Study population	Case definition	Study design	Treatment	Outcome measures	Results
Takala et al. (1994)[15]	Women in a printing company (Finland)	Frequent neck symptoms without signs of cervical nerve root compression or tendinitis of the shoulder	RCT, cross-over design 11 months $n = 44$	(I) Group gymnastics (II) Controls (no treatment)	Mean pressure pain threshold (bilaterally measured on four muscles (specified) with an algometer)	First period (I,II) (spring): small increase Second period (autumn): no change
Thorsen et al. (1992)[62]	Female hospital laboratory technicians (Denmark)	Age 18–65, pain from neck and shoulder-girdle lasting at least 1 year affecting the quality of work or daily living, between 1–10 tender points	RCT, cross-over 6 weeks $n = 52$	(I) Laser/placebo (II) Placebo/laser	Mean pain decrease at function (VAS) Mean pain decrease at rest (VAS) Mean pain decrease (funct + rest) Mean decrease in analgesic consumption % patients noticed benefit	27%(I), 50%(II) 39%(I), 50%(II) 33%(I), 50%(II) 33%(I), 40%(II) 52% (all patients, groups not specified)
Vasseljen et al. (1995)[54]	Female office workers (group 1). Female patients from local physiotherapists (group 2) (Norway)	Patients ($n = 24$) with shoulder and neck pain (≥3, scale (0–6)) last 6 months and previous 2 weeks, and pain ≥3 days continuously last 2 weeks An additional ($n = 9$) group: shoulder and neck pain past 2 weeks and ≥1 trigger point upper trapezius and pain on passive stretching	RCT(partial) 6 months $n = 33$	(I) Individual physiotherapy (II) Group exercise Additional group: (III) Individual physiotherapy	Mean pain decrease (VAS) % patients noticed benefit	48%(I), 52%(II), 66%(III) 83%(I), 42%(II), not reported (III)

[a]Methodological quality score in percentages (max. 100%). A copy of the more detailed methodological quality score can be obtained from the first author on request.
[b]Two papers reporting on one study.
ICHPPC, International Classification of Health Problems in Primary Care; RR, relative risk; PT, physiotherapy; ROM, range of motion; NSAIDs, non-steroidal anti-inflammatory drugs; TENS, transcutaneous nerve stimulation; MT, manual treatment; SWD, short-wave diathermy; MPQ, McGill pain questionnaire.

benefit' by patient or doctor/therapist. The most frequently reported secondary outcome measure was cervical range of motion (ROM). Almost no studies reported on psychological factors.

3.1.8. Analysis and data presentation

Nine studies adequately reported their most important outcome measure(s). Only Takala et al. (1992) used appropriate techniques to evaluate prognostic factors. However, a part of the analysed population consisted of patients who only developed symptoms after a symptom free baseline period. Since factors related to the onset of the disease are strictly seen as risk factors, the analysis did not evaluate purely prognostic factors.

3.2. Clinical course

Twelve studies were carried out in a secondary care setting, eight in an occupational setting and only one in a primary care setting. One used patients recruited by a newspaper announcement and one study did not specify the setting.

The main outcomes of the 23 studies are summarized in Table 3. In this table, the study populations are clustered according to the duration of neck complaints. We defined the following populations: acute (complaints ≤3 months), sub-acute (complaints ≥3 months and ≤6 months), chronic (complaint ≥6 months), mixed (regardless of the duration of complaints) and unknown (duration of complaints not specified). Only two studies reported on acute patients and no studies reported on sub-acute patients.

3.2.1. Pain

Twelve studies reported on pain, with a mean decrease ranging from 9 to 100% and a median of 34%. Two of these reported on improvement, without presenting data to support the conclusions. Only two studies (Nordemar and Thörner, 1981; Takala et al., 1994) used a follow-up of more than 6 months. Eight out of 12 studies reported on chronic patients (range 26–63%; median 28%). One of these eight studies reported on improvement, without presenting data to support the conclusions.

Five studies reported on the proportion of patients with pain decrease (less pain or pain free) (range 22–79%; median 46%). Only two (Gore et al., 1987; Levoska and Keinänen-Kiukaanniemi, 1993) of these studies used a follow-up period of more than 6 months.

3.2.2. General improvement

Twelve studies reported on the proportion of patients with general improvement (range 36–95%; median 47.5%). The nine studies including chronic patients showed similar results (range 37–95%; median 47%). Two studies (Berg et al., 1988; Abbot et al., 1990) reported a follow-up period of more than 6 months.

3.2.3. Functional status

Three studies used, defined and measured some sort of functional status. These studies reported on the proportion of patients who functionally improved (range 5–22%). All studies used a follow-up period of less than 6 months.

3.2.4. Health care utilization

Five studies reported a decreased intake of medication (mainly NSAIDs and analgesics) with a median of 37% (range 32–80%). All studies used a follow-up period less than 6 months. Two studies reported on treatment received. These two studies (Gore et al., 1987; Abbot et al., 1990)

Table 3
Outcome of studies on the clinical course of non-specific neck pain

Outcome	Population (*n* studies)	% (range)	% (median)	Follow-up (range, weeks)	Follow-up (median, weeks)
Mean pain decrease on VAS/numerical rating scale	Mixed (10)	9–100	34	3–52	11
	Chronic (7)	26–63	28	6–12	8
	Acute (1)	100	–	–	–
	Unknown (3)	9–49	–	–	–
% patients with pain decrease	Mixed (5)	22–79	46	24–520	38
	Chronic (1)	67	–	–	–
	Unknown (3)	22–79	–	–	–
% patients with general improvement (patients or therapists opinion)	Mixed (12)	36–95	47.5	3–156	7
	Chronic (9)	37–95	47	3–46	6
	Acute (1)	36	–	–	–
	Unknown (2)	40–71	–	–	–
Percentage patients with functional improvement (different scales)	Mixed (3)	5–22	9	8–24	12
	Chronic (2)	9–22	–	–	–
	Unknown (1)	5	–	–	–
Health care utilisation					
Mean decrease in analgesics use	Chronic (5)	7–80	37	6-12	8
% patients who received treatment in the past	Mixed (2)	80–100	90	520–∞	∞

∞ = Life-time.

reported that respectively 80 and 100% of the patients had had some form of treatment in the past. Both studies used a retrospective follow-up period of at least 10 years.

3.2.5. Lost days of work

Two papers reporting on one study (Abenhaim et al., 1988; Rossignol et al., 1988) found a mean number of 'sick listed' days of 25 per year for patients who were registered sick. Furthermore, they reported that 13% had recurrences, with a mean number of 0.86 recurrences per year. A second study (Tellnes, 1989) reported that 1.8% of the patients were still sick after 1 year of sickness certification. A third study (Anonymous, 1966) reported a 23% decrease of interference with work or stopping with work due to pain reduction. All studies used a mixed cohort of patients.

For most reported outcomes the median improvement (proportion patients improved or mean improvement) of all studies reporting, ranged between 40 and 50%. When the outcomes (median and range) of all the studies were compared with the outcomes based on just the 'high quality' studies, no important differences were apparent. The pain decrease (mixed population) changed from 34% (median), 9–100% (range) for all studies, to 35% (median), 25–63% (range) for just 'high quality' studies. For the proportion of patients with pain decrease, there were no 'high quality' studies available. The general improvement (mixed population) changed from 47.5% (median), 36–95% (range) for all studies, to 52% (median), 43–95% (range) for just 'high quality' studies. All studies reporting on medication intake were 'high quality' studies.

3.3. Prognostic factors

Only six studies reported on prognostic factors (Table 4). The most frequently reported prognostic factors were age, sex, severity of pain, localization, duration, occupation and radiologic findings. None of the studies reported the strength of the association (relative risk (RR) or odds ratio (OR)) between a prognostic factor and the outcome. In some cases the direction of the association between a prognostic factor and the outcome, was not specified. In three studies, it was not clear whether a statistical test was used.

3.3.1. Age

Three studies reported on age as a prognostic factor. Two studies (Sloop et al., 1982; Loy, 1983) reported no association between age and a worse prognosis. However, in one of

Table 4
Prognostic factors for non-specific neck pain

Author (year)	Prognostic factors	Outcome	Association
Abbot et al. (1990)	Financial compensation, change in occupation	Improvement	No[a]
Berg et al. (1988)	Manual work (vs. office work)	More symptoms	Yes
	After retirement (both manual/office workers)	Fewer symptoms	Yes
	Type of work (manual, office) before retirement	Fewer symptoms	No[a]
Gore et al. (1987)	Sex	Worse outcome	No
	Initially severe pain		
	Injured subgroup	Worse outcome	Yes
	Not injured subgroup	Worse outcome	No
	Combined subgroups	Worse outcome	Yes
	Roentgenographic findings	Level of pain	No
	Localization of pain	Worse outcome	No
Sloop et al. (1982)	Middlesex Hospital Questionnaire (emotional and situational factors), Social Readjustment Rating Scale (life changes), age, sex, history of trauma, tablet count, presence of arm pain, radiographic grate, central nervous system symptoms	Improvement (VAS)	No
	Local tenderness on initial examination	Worse prognosis	Yes
Loy (1983)	Duration of symptoms, age, sex, occupation, severity of radiological changes	Worse outcome	No[a]
Anonymous (1966)	Age, severity of attack, number of previous attacks, average duration of symptoms in previous attack whether symptoms were getting better or worse when the patient was first seen	Improvement (after 4 weeks)	Yes[b]
	Range of neck movement, abnormal neurological signs, x-ray changes	Improvement (after 4 weeks)	No
	History of attacks for more than 5 years, >3 previous attacks, bilateral paraesthesia, women >50 years, symptoms that were getting better or worse when the patient was first seen	Worse outcome (after 6 months)	Yes

[a]Not clear whether a statistical test was used.
[b]Direction of the association was not specified.

these studies (Loy, 1983), it was not clear whether a statistical test was used. The third study (Anonymous, 1966) reported an association between a worse prognosis for women >50 years (statistically significant).

3.3.2. Sex

Four studies reported on sex as a prognostic factor. Three studies (Sloop et al., 1982; Loy, 1983; Gore et al., 1987) reported no association between sex and a worse prognosis. However, in one of these studies (Loy, 1983), it was not clear whether a statistical test was used. A fourth study (Anonymous, 1966) found a statistically significant worse prognosis for women over 50 years.

3.3.3. Pain (severity)

Three studies reported on the severity of pain as a prognostic factor. One study (Gore et al., 1987) reported severe pain (at baseline) to have an unsatisfactory outcome (statistically significant). It was not statistically significant for the subgroup of patients without injuries. The second study (Sloop et al., 1982) reported a statistically significant worse prognosis for patients with local tenderness on the initial examination. The third study (Anonymous, 1966) reported an association (statistically significant) between the severity of the attack and the prognosis. The direction of the association was, however, not specified.

3.3.4. Localization

Three studies reported on localization. The presence of arm pain (mean duration of 6 years) and central nerve system symptoms was reported not to be associated with a worse prognosis (Sloop et al., 1982). The second study (Gore et al., 1987) reported no association between the localization of pain and a worse prognosis. The third study (Anonymous, 1966) reported a statistically significant worse prognosis for patients with bilateral paraesthesia, but no worse prognosis for abnormal neurologic signs.

3.3.5. Duration and number of attacks

Two studies reported on duration and number of attacks. One study (Loy, 1983) reported no association between the duration of symptoms at baseline and a worse outcome. In this study however, it was not clear whether a statistical test was used. The second study reported both an average duration of symptoms in the previous attack and the number of previous attacks to be (statistically significant) associated with improvement (Anonymous, 1966).The direction of the association was not specified. The same study reported a statistically significant worse prognosis for a history of attacks >5 years or >3 previous attacks.

3.3.6. Occupation

Three studies reported on occupation. One study (Berg et al., 1988) reported a statistically significant worse prognosis for manual workers compared with office workers. After retirement both office and manual workers had fewer symptoms. Type of work (before retirement) however, was not associated with improvement. In this study it was not clear whether a statistical test was used to assess this association. A second study (Abbot et al., 1990) reported change in occupation not to be associated with improvement. A third study (Loy, 1983) reported the type of occupation not to be associated with the prognosis. In the second and third studies however, it was not clear whether a statistical test was used.

3.3.7. Radiological findings

Four studies reported on radiologic findings. One study (Gore et al., 1987) reported no association between degenerative changes and the level of pain. The second study (Anonymous, 1966) reported no association between radiologic findings and improvement. Two studies (Sloop et al., 1982; Loy, 1983) reported the severity of the findings not to be associated with the prognosis. In one of these studies (Loy, 1983) however, it was not clear whether a statistical test was used.

4. Discussion

4.1. Number of studies

We originally planned to include observational studies only. However, after the initial literature search, the number of studies on this topic turned out to be very small. It was surprising to see that there were only a few studies on the clinical course of non-specific neck pain available. This is somewhat strange, since a lot of effort is put into planning and conducting RCTs, whereas at the same time little is known about the clinical course of the disease.

A number of studies reported on combined localizations for neck-shoulder pain or neck-low back pain, without reporting data for neck pain separately. Consequently, these studies could not be included in this review.

4.2. Methodological quality

The methodological quality according to our criteria list appeared to be rather low. It should be noticed that the RCTs we included and assessed were not necessarily of poor quality for the purpose they were originally designed for, namely assessing the efficacy of therapeutic interventions. One of the potential explanations is that RCTs are not designed to identify prognostic factors as these studies mainly focus on differences between groups in order to assess the efficacy of one or more treatments.

The most prevalent methodological shortcomings regarding the criteria list appeared to be selection of the study population and analysis techniques. Both items were scored positive by only one study. It was surprising to see that just one study (Takala et al., 1992) used a cohort of patients who were identified at a unique point in the course of the disease

at baseline, which is an important aspect for the assessment of prognosis.

4.3. Limitations of this review

We had to deal with some problems for which no standard solution was available. A potential limitation might be the literature search. Although we used a sensitive set of keywords for RCTs and observational studies, we might have missed some relevant publications. Secondly, we only searched for studies which were published in indexed journals, so unpublished studies and non-indexed journals would have been missed. Thirdly, we only used studies which were published in English, Dutch or German and consequently studies published in other languages have been missed.

The use of RCTs was introduced by a lack of observational studies. We think the use of RCTs was permitted, because these can be viewed as cohort studies in which therapy is one of the potential prognostic factors. Furthermore, we attempted to construct a median and range for the overall improvement (in percentages). Although most studies used a VAS, some used a 10-point scale making it more difficult to compare improvement scores. The same problems occurred when comparing general improvement scores, while improvement could be described by different definitions e.g. 'slight improvement' or 'no symptoms'.

In case a retrospective study design was used, selection bias could have been introduced by including only those cases available for follow-up. Therefore, retrospective studies could not score positive on the criterion referring to drops-outs and loss to follow-up, unless detailed information was provided that there were no drop-outs/loss to follow-up. Consequently, retrospective studies always missed 30% of the maximum attainable score.

4.4. Clinical course

There is not much information on the course of acute neck pain. More information is available for the group of patients with complaints for more than 6 months, who are treated in a secondary care or occupational setting. Most of these studies report on a follow-up period of 6 months or less. In this group approximately 50% of the patients had less pain and a general improvement of 50% with a mean reduction of pain and use of analgesics of about 30%.

4.5. Prognostic factors

There is very limited evidence regarding prognostic factors related to the course of non-specific neck pain. For the few studies reporting on prognostic factors, the main shortcomings are the sample size and the lack of appropriate analyses techniques. Bearing these limitations in mind there are some indications (based on the most frequently reported statistically significant associations) that there is no association between localization (radiation to the arms/neurologic signs) and a worse outcome. Furthermore, there are some indications that there is no association between radiologic findings (degenerative changes in the discs and joints) and a worse prognosis. The severity of pain and a history of previous attacks, however, seems to be associated with a worse prognosis.

4.6. Recommendations

As a result of the most serious gaps we found in the literature on understanding the clinical course and prognostic factors of non-specific neck pain, we suggest that further research is needed on acute neck pain in primary care, using an inception cohort (first period of complaints). To identify patients at risk for developing chronic neck pain or to study the course of chronic non-specific neck pain, a follow-up period of at least 1 year should be used. For the identification of prognostic factors, studies should be designed to generate valid prognostic factors by using a sufficient sample size and adequate analysis techniques.

Acknowledgements

The authors wish to thank Mr. W. Devillé for statistical counselling and Ms. D.L.M. Veldman for her assistance in the preparation of the manuscript. This study was supported by a grant from the Dutch Scientific Research Foundation (NWO), The Hague, The Netherlands.

References

Abbot, P., Rounsefell, B., Fraser, R. and Goss, A., Intractable neck pain, Clin. J. Pain, 6 (1990) 26–31.

Abenhaim, L., Suissa, S. and Rossignol, M., Risk of recurrence of occupational back pain over three years follow-up, Occup. Environ. Med., 45 (1988) 829–833.

Andersson, G.B.J., The epidemiology of spinal disorders. In: J.W. Frymoyer (Ed.), The Adult Spine: Principles and Practice, Lippincott-Raven, Philadelphia, PA, 1997, pp. 130–141.

Anonymous, Pain in the neck and arm: a multi centre trial of the effects of physiotherapy, Br. Med. J., 1 (1966) 253–258.

Berg, M., Sandén, Å., Torell, G. and Järvholm, B., Persistence of musculoskeletal symptoms: a longitudinal study, Ergonomics, 31 (1988) 1281–1285.

Bogduk, N., Neck pain, Aust. Fam. Physician, 13 (1984) 26–30.

Bogduk, N., Neck pain: an update, Aust. Fam. Physician, 17 (1988) 75–80.

Bouter, L.M. and van Dongen, M.C.J.M., Epidemiologisch onderzoek. Opzet en interpretatie, Bohn Stafleu Van Loghum, Houten/Antwerpen, 1991, pp. 103–109.

Coan, R.M., Wong, G. and Coan, P.L., The acupunture treatment of neck pain: a randomized controlled study, Am. J. Chin. Med., 4 (1982) 326–332.

Ceccherelli, F., Altafini, L., Lo Castro, G., Avila, A., Ambrosio, F. and Giron, G.P., Diode laser in cervical myofascial pain: a double-blind study versus placebo, Clin. J. Pain, 5 (1989) 301–304.

Cole, D.C. and Hudak, P.L., Prognosis of nonspecific work-related mus-

culoskeletal disorders of the neck and upper extremity, Am. J. Ind. Med., 29 (1996) 657–668.
Dickersin, K., Scherer, R. and Lefebvre, C., Identifying relevant studies for systematic reviews, Br. Med. J., 309 (1994) 1286–1291.
Foley-Nolan, D., Barry, C., Coughlan, R.J., O'Connor, P. and Roden, D., Pulsed high frequency (27 MHz) electromagnetic therapy for persistent neck pain. A double blind, placebo-controlled study of 20 patients, Orthopedics, 13 (1990) 445–451.
Goldie, I. and Landquist, A., Evaluation of the effects of different forms of physiotherapy in cervical pain, Scand. J. Rehabil. Med., 2–3 (1970) 117–121.
Gore, D.R., Sepic, S.B., Gardner, G.M. and Murray, M.P., Neck pain: a long-term follow-up of 205 patients, Spine, 12 (1987) 1–5.
Gross, A.R., Aker, P.D., Goldsmith, C.H. and Peloso, P., Conservative management of mechanical neck disorders, Online J. Curr. Clin. Trials (serial online) July 30 (1996) 5 (doc No 200).
Horvath, J. and Fellmann, N., Behandlungsergebnisse mit nifluril im doppelten blindversuch bei degenerativen zervikalen veränderungen mit oder ohne zervikalsyndrom, Praxis, 42 (1969) 1342–1345.
Howe, D.H., Newcombe, R.G. and Wade, M.T., Manipulation of the cervical spine. A pilot study, J. Roy. Coll. Gen. Pract., 33 (1983) 574–579.
Jensen, I., Nygren, Å., Gamberale, F., Goldie, I., Westerhom, P. and Jonsson, E., The role of the psychologist in multidisciplinary treatments for chronic neck and shoulder pain: a controlled cost-effectiveness study, Scand. J. Rehabil. Med., 27 (1995) 19–26.
Laupacis, A., Wells, G., Richardson, W.S. and Tugwell, P., Users' guides to the medical literature. V. How to use an article about prognosis, J. Am. Med. Assoc., 272 (1994) 234–237.
Levoska, S. and Keinänen-Kiukaanniemi, S., Active or passive physiotherapy for occupational cervicobrachial disorders? A comparison of two treatment methods with 1-year follow-up, Arch. Phys. Med. Rehabil., 74 (1993) 425–430.
Loy, T.T., Treatment of cervical spondylosis. Electro acupuncture versus physiotherapy, Med. J. Austr., 2 (1983) 32–34.
Nordemar, R. and Thörner, C., Treatment of acute cervical pain. A comparative group study, Pain, 10 (1981) 93–101.
Petrie, J.P. and Hazleman, B.L., A controlled study of acupuncture in neck pain, Br. J. Rheumatol., 25 (1986) 271–275.
Revel, M., Minguet, M., Gergoy, P., Vaillant, P. and Manuel, J.I., Changes in cervicocephalic kinesthesia after a proprioceptive rehabilitation program in patients with neck pain: a randomized controlled study, Arch. Phys. Med. Rehabil., 75 (1994) 895–899.
Rossignol, M., Suissa, S. and Abeheim, L., Working disability due to occupational back pain: three-year follow-up of 2,300 compensated workers in Quebec, J. Occup. Environ. Med., 30 (1988) 502–505.
Sackett, D.L., Haynes, R.B. and Tugwell, P., Clinical Epidemiology, Little, Brown and company, Boston, MA, 1991.
Sloop, P.R., Smith, D.S., Goldenberg, E. and Doré, C., Manipulation for chronic neck pain. A double-blind controlled study, Spine, 7 (1982) 532–535.
Takala, E.-P., Viikari-Juntura, E., Moneta, G.B., Saarenmaa, K. and Kaivanto, K., Seasonal variation in neck and shoulder symptoms, Scand. J. Work. Environ. Health., 18 (1992) 257–261.
Takala, E.-P., Viikari-Juntura, E. and Tynkkynen, E.-M., Does group gymnastics at the workplace help in neck pain? Scand. J. Rehabil. Med., 26 (1994) 17–20.
Tellnes, G., Duration of episodes of sickness certification, Scand. J. Prim. Health Care, 7 (1989) 237–244.
Thorsen, H., Gam, A.M., Svensson, B.H., Jess, S., Jensen, M.K., Piculell, I., Schack, L.K. and Skjøtt, K., Low level laser therapy for myofacial pain in the neck and shoulder girdle. A double-blind, cross-over study, Scand. J. Rheumatol., 21 (1992) 139–142.
Valkenburg, H.A., Laar, A. van Hofman, A. and Haanen, H.C.M., Nek-en lage rugklachten. In: Instituut epidemiologie. Jaarverslag instituut epidemiologie en zesde voortgangsverslag van het epidemiologisch preventief onderzoek Zoetermeer. April 1980, pp. 99-109.
Vasseljen, O., Johansen, B.M. and 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. Rehabil. Med., 27 (1995) 243–252.
Von Korff, M., Studying the natural history of back pain, Spine, 19 (18S) (1994) 2041S–2046S.