

# Methodological quality of a systematic review on physical therapy for temporomandibular disorders: influence of hand search and quality scales

Bart Craane · Pieter Ubele Dijkstra · Karel Stappaerts · Antoon De Laat

Received: 22 January 2010 / Accepted: 16 November 2010 / Published online: 3 December 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** The validity of a systematic review depends on completeness of identifying randomised clinical trials (RCTs) and the quality of the included RCTs. The aim of this study was to analyse the effects of hand search on the number of identified RCTs and of four quality lists on the outcome of quality assessment of RCTs evaluating the effect of physical therapy on temporomandibular disorders. In addition, we investigated the association between publication year and the methodological quality of these RCTs. Cochrane, Medline and Embase databases were searched electronically. The references of the included studies were checked for additional trials. Studies not electronically identified were labelled as “obtained by means of hand search”. The included RCTs (69) concerning

physical therapy for temporomandibular disorders were assessed using four different quality lists: the Delphi list, the Jadad list, the Megens & Harris list and the Risk of Bias list. The association between the quality scores and the year of publication were calculated. After electronic database search, hand search resulted in an additional 17 RCTs (25%). The mean quality score of the RCTs, expressed as a percentage of the maximum score, was low to moderate and varied from 35.1% for the Delphi list to 54.3% for the Risk of Bias list. The agreement among the four quality assessment lists, calculated by the Interclass Correlation Coefficient, was 0.603 (95% CI, 0.389; 0.749). The Delphi list scored significantly lower than the other lists. The Risk of Bias list scored significantly higher than the Jadad list. A moderate association was found between year of publication and scores on the Delphi list ( $r=0.50$ ), the Jadad list ( $r=0.33$ ) and the Megens & Harris list ( $r=0.43$ ).

**Electronic supplementary material** The online version of this article (doi:10.1007/s00784-010-0490-y) contains supplementary material, which is available to authorized users.

B. Craane (✉) · K. Stappaerts  
Faculty of Kinesiology and Rehabilitation Sciences, Department of Rehabilitation Sciences, Catholic University of Leuven, Leuven, Belgium  
e-mail: bart.craane1@telenet.be

P. U. Dijkstra  
Department of Rehabilitation, School for Health Research, University Medical Center, Groningen, The Netherlands

P. U. Dijkstra  
Department of Oral and Maxillofacial Surgery, School for Health Research, University Medical Center, Groningen, The Netherlands

A. De Laat  
Department of Oral and Maxillofacial Surgery, School of Dentistry, Oral Pathology and Maxillofacial Surgery, Catholic University of Leuven, Leuven, Belgium

**Keywords** Physical therapy · Temporomandibular disorders · Randomised clinical trials · Methodological quality · Hand search

## Introduction

Temporomandibular disorders (TMD) is a collective term embracing a number of clinical problems that involve the masticatory musculature, the temporomandibular joint and associated structures, or both [1].

Physical therapy (PT) is defined as “treatment modalities (including exercise, heat and cold application, electrotherapy, massage, stretching, mobilisation, instructions) in order to prevent, correct and alleviate movement dysfunction and pain of anatomic or physiologic origin” and is frequently used as part of the conservative and non-invasive management of

TMD. Although papers on physical treatment for TMD have been published since 1952 [2], the first evidence for its effectiveness based on randomised clinical trials (RCTs) was described in the studies of Kopp and Stenn et al. [3, 4]. In a recent systematic review, 69 RCTs regarding PT for TMD were identified up to February 2010.

Retrieving evidence from large electronic databases such as Medline, Embase and the Cochrane Central Register of Controlled Trials is challenging. The use of adequate search strategies can increase the number of relevant studies while minimising the number of non-relevant studies. In addition to the electronic search strategies, hand searching of all the references of the electronically identified RCTs found, as well as the references of the references of the newly discovered RCTs (manual cross-reference search), may again increase the number of relevant RCTs. The first aim of the present study was to assess the influence of hand searching on the number of RCTs found in a systematic review.

Quality assessment of the identified RCTs is important. Various methods, such as quality scales, criteria lists and checklists can be used [5]. Quality of RCTs defined as ‘the likelihood of the trial design to generate unbiased results’ covers only the dimension of internal validity [6]. Most quality lists however, measure at least three dimensions: internal validity, external validity and statistical validity [7, 8]. Even an ethical component in the concept of quality can be distinguished. The ethical principles of beneficence (doing the best for one’s patients and clients), non-maleficence (doing no harm), patients’ autonomy, justice and equity are positively associated with the quality of a trial [9]. Up to now, it is not clear what the effect is of the different quality lists on the outcome of quality assessment of a particular study. The second aim of the present study therefore was to analyse the effect of four quality lists (Delphi, Jadad, Megens & Harris and Risk of Bias) on the quality assessment of RCTs. The four different lists were applied on the set of 69 RCTs regarding PT for TMD.

PT is a relatively young profession evolving over time. The last decades, the number of published RCTs regarding the effect of the PT interventions on musculoskeletal problems in general and on TMDs in particular, has increased. Assessing the methodological quality of the RCTs in our recent systematic review prompted the question: ‘Has the methodological quality of RCTs increased over time?’, and consequently, the third aim of this study was to analyse the association between publication year and methodological quality as assessed by the different criteria lists.

In summary, based on a recently completed systematic review on the effectiveness of PT on TMD, the aims of the present study were: (1) to analyse the importance of hand search in identifying relevant studies; (2) to analyse the influence of different quality lists on the results of the quality assessment of RCTs; (3) to analyse the association

between publication year and the quality of the RCTs (assessed by four different criteria lists).

## Material and methods

### Importance of hand search

Three databases, Cochrane, Medline and Embase, were searched electronically via OVID (last search date: February 2010) for relevant RCTs concerning the effects of PT on TMD. The search strategies are based on the search strategy developed for Medline but revised appropriately for each database to take in to account differences in controlled vocabulary (MeSH) and syntax rules (Appendix). All identified studies were screened for their relevance. A study was included in the review process if the title, abstract or full text indicated a RCT regarding PT and TMD. In addition to these databases, the Web of Science was also searched. All studies identified in the database search, published in 2000 and later, were imported in the Web of Science to search for publications citing the studies identified in the searches (Cited Reference Search). The publications found in Web of Science were then again screened for relevance on their title, abstract or full text. In a next step, the references of all the included RCTs were checked manually for relevant RCTs (reference check) and finally the references of (systematic) reviews concerning PT and TMD that were identified through the electronic search were checked manually for relevant RCTs. All RCTs not identified by means of electronic databases were labelled as “obtained by means of hand search”.

### Influence of criteria list used

All included RCTs ( $n=69$ ) were assessed on their methodological quality by one observer (BC) using four different quality lists. The Delphi list was developed by consensus among experts. It consists of ten items (scoring range, 0 to 10). The Delphi list assesses three dimensions of quality: internal and external validity and statistical considerations [10]. The Risk of Bias list was developed by a workgroup of methodologists, editors and review authors and is recommended by The Cochrane Collaboration [11]. It consists of six items (scoring range, 0 to 6). The Megens & Harris list [12] was developed by the McMaster Occupational Therapy Evidence-Based Practice Research Group [13, 14]. It consists of ten items (scoring range, 0 to 11). The Jadad list [6] is a criteria list initially compiled by a multidisciplinary panel of six “judges” and narrowed down by means of the Nominal Group Consensus Technique [7]. It consists of three items which assess internal validity (scoring range, 0 to 5). An overview of the lists has been summarised in Table 1.

**Table 1** Overview of four quality lists: Delphi, Risk of Bias (RoB), Megens and Harris (M&H) and Jadad

Delphi	RoB	M&H	Jadad
<i>Randomization</i>			
Was the method of randomisation performed	Was the allocation sequence adequately generated?	Was the study randomised (this includes the use of words such as randomly, random and randomisation)? +1 Described	Was the study described as randomised (this includes the use of words such as randomly, random and randomisation)? +1 Described and appropriate -1 Described and inappropriate
If subjects were randomly allocated to treatment groups, was the method of random allocation concealed?	Was the allocation adequately concealed?		
<i>Similarity of groups</i>			
Were the groups similar at baseline regarding the most important prognostic characteristics?		Were the groups similar at baseline?	
<i>Inclusion/exclusion criteria</i>			
Were both inclusion and exclusion criteria specified?		Were the inclusion and exclusion criteria listed for the subjects?	
<i>Blinding</i>			
Was the outcome assessor blinded?	Was knowledge of the allocated interventions adequately prevented during the study?	Was the patient, the treatment provider and the assessor blinded?	Was the study described as double blind? (blinding of patients and evaluators, not necessarily therapist) +1 Method of blinding described and appropriate -1 Method of blinding not appropriate
Was the patient blinded?			
Was the care provider blinded?			
<i>Statistics</i>			
Were point estimates and measures of variability presented for primary outcome measure(s)?			
Did the analysis include an 'intention-to-treat' analysis?			
<i>Drop-outs and completeness data</i>			
Were the drop-outs described and acceptable?	Were incomplete outcome data adequately addressed?	Were the drop outs reported?	Was there a description of withdrawals and drop-outs? (explicit statement that all included patients were analysed or if the number and reasons for dropouts in all groups are given separately)
<i>Description of other criteria for trial quality</i>			
	Are reports of the study free of suggestion of selective outcome reporting?	Was the treatment protocol sufficiently described to be replicable?	
	Was the study apparently free of other problems that could put it at a risk of bias?	Was the validity of data obtained with the outcome measures addressed?	
		Was the reliability of data obtained with the outcome measures investigated?	
		Was the follow-up minimum 6 months?	
		Was a home program adherence investigated? If included!	

A score of 1 was given for each item fulfilled by the RCT. A score of 0 was given if the item was not fulfilled or when it was unclearly reported. The scores were summed and for comparison between lists, the percentage of the total possible score was calculated (= quality score (QS)). This percentage was used for the statistical analysis. The agreement among the four quality lists for the complete set of 69 RCTs was calculated by the interclass correlation coefficient (ICC) as described by Portney and Watkins [15]. Since the four scales can be regarded as a random sample of all possible quality lists, the ICC expresses inter-scale agreement in a single rating. Differences between the different quality lists were analysed with repeated measures ANOVA and a post hoc analysis (Bonferroni corrected).

### Quality of RCTs related to the year of publication

The quality of the RCTs, assessed as the percentage number of positive items scored on the different quality lists, was correlated (Pearson's  $r$ ) with the year of publication (from 1978 to 2009). For all statistic calculations, we used SPSS® Software Version 16.

## Results

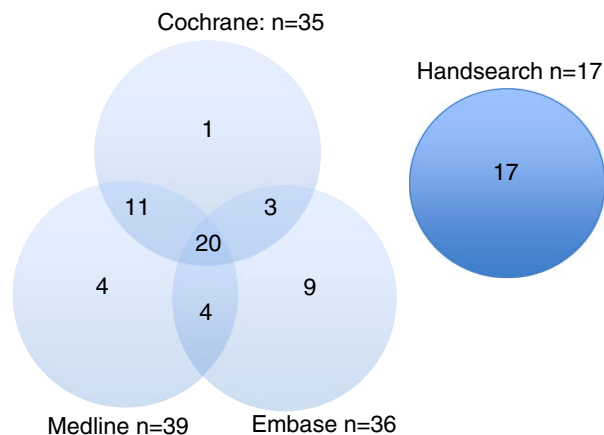
### Importance of hand search

After removing duplicate studies (281), the electronic and hand search of the literature resulted in 407 articles. After applying the inclusion and exclusion criteria, 69 RCTs concerning PT and TMD remained for systematic review. Reasons for exclusion were: no data on treatment effect (251), reviews (29), no randomised controlled trials (37), data of a subsequently published trial (7), physical therapy after neoplastic conditions or systemic diseases (2), no TMD pathology (4), no PT as previously defined (5), irrelevant outcome variables (2), and therapy on painless TMD symptoms (1). The source of identification of the included studies is presented in Fig. 1. The electronic search identified 52 (75%) studies included in the review. Hand search resulted in an additional 17 (25%) RCTs. The Cochrane Central Register of Controlled Trials provided 35 (51%), the Embase database 36 (52%) and the Medline database 39 (57%) of the included studies. Twenty (29%) studies were identified in all three databases.

### Influence of criteria lists

Scrutinising the criteria composing the different quality lists resulted in the following observations: all criteria list includes items to identify randomisation or the procedure of randomisation. The requirement to score positively on

### Source of identification of RCTs



**Fig. 1** Number of RCTs according to the source of identification (Cochrane = the Cochrane Central Register of Controlled Trials)

this item is different for the different lists. All four lists include items about 'randomisation', 'blinding' and 'dropouts'. The Delphi list differentiates between the 'levels of blinding' (patient, therapist or observer) whereas the Jadad list includes 'a description of the blinding method'. The Delphi list and the Risk of Bias list, assess 'treatment allocation' and 'statistical analysis'. 'The presentation of the data' is assessed only in the Delphi list. The Megens & Harris list is the only one that scores, 'the length of follow-up', 'home programme', 'reliability' and 'validity of the outcome measurement' and 'description of treatment protocol'. Only the Delphi and the Megens & Harris lists assess 'the similarity of the groups at baseline'. The Risk of Bias list contains 'selective outcome reporting' and 'other potential threats to validity'.

In Table 2, the included studies are presented with their quality scores according to the different quality assessment methods. The Delphi scores varied between 0 and 8 points out of 10. The Risk of Bias scores varied between 0 and 6 out of 6. The Megens & Harris scores varied between 2 and 9 out of 10 and between 2 and 11 out of 11 (if 'home programme adherence' was investigated). The Jadad scores varied between 0 and 4 out of 5. Two studies scored maximum scores for the Risk of Bias list and one study scored maximum in the Megens & Harris list. None of the studies were assigned maximum scores on any other criteria lists. The mean (SE) quality score of the 69 RCTs, expressed as a percentage of the maximum possible score, varied from 35.1 (2.2) for the Delphi list, 48.7 (2.4) for the Jadad list, 49.5 (2.2) for the Megens & Harris list to 54.3 (2.4) for the Risk of Bias list. The agreement between the four quality assessment lists (ICC) was 0.603 (95% CI, 0.389; 0.749). In repeated measures ANOVA, a significant difference was found between the scores of the different scales. ( $F_{3,204}=44.2819$  ( $p<0.001$ )). Post hoc analysis (Bonferroni corrected) made it clear that the Delphi list scored significantly lower than the

**Table 2** Results of the quality score for the different criteria lists expressed as a percentage of the maximum possible positive items scored

Author	Delphi (%)	RoB (%)	M&H (%)	Jadad (%)
Al-Badawi 2004	40	67	40	40
Alvaraz 2002	10	33	20	20
Bakke 2008	40	33	60	40
Bender 1991	20	33	30	20
Bertolucci 1995	20	33	30	20
Brooke 1983	10	50	30	40
Burgess 1988	30	50	40	40
Carlson 2001	60	67	64	80
Carmeli 2001	30	50	50	60
Conti 1997	40	50	50	60
Crockett 1986	20	33	36	20
Dahlstrom 1984	0	50	27	40
Dalen 1986	10	0	27	20
De Abreu 2005	40	67	50	60
DeLaat 2003	40	67	60	60
Dogu 2009	20	50	40	20
Dohrman 1978	30	50	36	60
Dworkin 1994	60	83	73	80
Dworkin 2002a	40	50	64	40
Dworkin 2002b	40	50	36	40
Erlandson 1989	10	50	30	40
Funch 1984	30	33	73	60
Gardea 2001	60	67	70	60
Gavish 2006	40	67	60	60
Glaros 2007	50	67	64	60
Glas 2000	30	67	46	60
Gray 1994	30	33	40	80
Ismaïl 2007	50	50	30	40
Kavuncu 1999	20	50	30	40
Klobas 2006	40	100	73	80
Komiyama 1999	30	50	60	40
Kopp 1979	20	67	30	60
Kruger 1998	20	33	30	20
Kulekcioglu 2003	40	50	50	40
Linde 1985	30	50	36	40
Magnussen 1999	20	0	46	40
Maloney 2002	20	33	40	20
Mazzetto 2007	40	50	50	80
Michelotti 2004	50	50	64	40
Minakuchi 2004	70	83	70	80
Monteiro 1988	0	33	18	20
Moystad 1990	40	50	50	40
Mulet 2007	60	83	82	80
Nunez 2006	30	50	40	40
Okeson 1983	20	50	36	40
Olson 1987	20	67	40	60
Peroz 2004	80	100	70	80
Reid 1994	50	50	60	60

**Table 2** (continued)

Author	Delphi (%)	RoB (%)	M&H (%)	Jadad (%)
Schiffman 1996	60	67	60	60
Schiffman 2007	70	100	100	80
Shin 1997	30	50	40	40
Stam 1984	30	67	60	40
Stegenga 1993	20	50	50	40
Stenn 1979	30	33	30	40
Talaat 1986	0	33	20	20
Taube 1988	30	67	40	60
Taylor 1987	50	67	40	60
Taylor 1994	40	50	40	40
Townsend 2001	20	50	70	20
Treacy 1999	30	67	50	40
Truelove 2006	70	83	82	80
Tullberg 2003	70	83	60	80
Turk 1993	20	33	27	40
Turner 2008	70	83	90	60
Wahlund 2003	30	67	64	60
Wright 1995	50	50	80	60
Wright 2000	50	67	73	60
Yuasa 2001	10	17	30	0
Yoshida 2005	40	67	60	60

other three lists and that the Risk of Bias list scored significantly higher than the Jadad list (Table 3).

Quality of RCTs related to year of publication

The correlation between trial quality and the year of publication was 0.497 (95% CI, 0.295; 0.656) for the Delphi list, 0.329 (95% CI, 0.101; 0.525) for the Risk of Bias list, 0.481 (95% CI, 0.276; 0.644) for the Megens & Harris list, and 0.219 (95% CI, -0.018; 0.433) for the Jadad list.

**Discussion**

Hand search identified 17 RCTs (25%) that were not found in the electronic databases. In a recent study, Egger and

**Table 3** The mean quality scores (+standard error) expressed as a percentage of the maximum possible score

Scale	Mean score	Std. error	95% Confidence interval
Delphi	35.1	2.2	30.6; 39.5
Risk of bias	54.3	2.4	49.5; 59.2
MH	49.5	2.2	45.1; 53.9
Jadad	48.7	2.4	43.9; 53.5

Smith concluded that the Cochrane Central Register of Controlled Trials is still likely to be the best source of information and should be the first one to be examined by those carrying out systematic reviews [16]. In the present study, 51% of the studies were found in the Cochrane Central Register of Controlled Trials, 52% in Embase and 57% in Medline. This illustrates that consulting also other databases is important to reduce the selection bias in identifying studies to be included. In addition, since Cochrane, Medline and Embase searches together resulted in only 75% of the included reports, our present study indicates that hand search plays a valuable role in identifying randomised controlled trials. Similar results were found in a previous report in which 82% of the studies were identified by means of complex electronic searches [17]. The present results, therefore, concur with Richards [18] who commented that although complex electronic searches using a range of databases may identify the majority of trials, hand searching is still valuable in identifying randomised trials. Also Crumley et al. highlighted the importance of searching multiple sources for conducting a systematic review [19]. For example, only 23 of 33 (67%) studies were found while searching Embase in a study of Al-Hajeri et al. [20]. Possible reasons why electronic searches fail are multiple: lack of relevant indexing terms, inconsistency by indexers, reports published as abstracts and/or included in supplements that are not routinely indexed by electronic databases [21, 22]. The Cochrane Collaboration has recognised the importance of searching journals page-by-page and reference-by-reference to trace as many relevant articles as possible and has set up a worldwide journals hand searching programme to identify RCTs [23].

The use of a criteria list allows estimating the methodological quality of the design and conduct of the trial. The items of the different criteria lists focus on different methodological aspects of RCTs and enable assessment of methodological quality by a summation of criteria scores. Calculating summary scores inevitably involves assigning a particular ‘weight’ to different items in the scale, and it is difficult to justify the weights assigned. Therefore, the summation scores must be simply interpreted as a ‘number of items scored positively’ on the list. The summation of these quality scores results in a hierarchical list in which more positive items indicate a better methodological quality [24]. However, different sets of criteria applied to the same set of trials do not always provide similar results [25]. The present study compared the overall QS resulting from different quality lists and showed significant differences in mean scores expressed as a percentage. These observed differences probably result in part from the variation of items included in the different lists. Only 3 out of 15 different items used in the four quality

scales are represented in all four of them: ‘randomisation’, ‘blinding’ and ‘drop-outs’. Additionally, the ‘wording’ of similar items is different in the different lists. In the Delphi and Risk of Bias lists, assessment of randomisation requires more specific information, while in the Megens & Harris and the Jadad list, the simple use of words such as randomly, random and randomisation is sufficient to score positive for this item. ‘Blinding’ is represented in all four lists, but the Delphi list discriminates between outcome assessor, therapist and patient and consequently ‘blinding’ scores 3 items out of 10. By contrast, in the Risk of Bias method, blinding is represented as only 1 item out of 6, and in the Megens & Harris list as 1 item out of 10 or 11. In the Jadad list, an extra point can be earned if the method of randomisation is explicitly described and therefore ‘blinding’ accounts for 2 items out of 5. In most of the PT interventions, blinding of the therapist and patient is impossible. Consequently the ‘weight’ of blinding as 3 out of 10 items for the Delphi list and 1 out of 6 for the Risk of Bias list could cause lower quality scores for PT studies using the Delphi list. A typical example in the present review was the study of Carmeli et al. [26] that scored 3 on the Risk of Bias list and also 3 on the Delphi list. Whereas ‘blinding’ represents 1 item out of 6 for the Risk of Bias list (=17%), it counts for 3 items out of 10 for the Delphi list (=33%).

Well-conducted RCTs provide the best evidence on the efficacy of a particular treatment. Since the publication of a study undertaken for Britain’s Medical Research Council by Hill in 1948, that may have been the first to have all the methodological elements of a modern RCT [27], the number of RCTs published each year increases immensely: according to Pubmed, over 9,000 new RCTs were published in 2008. For the practising clinician, it becomes impossible to keep up with the recent evidence. To appraise and synthesise this information, systematic reviews can be of great help. Of course, the validity of the conclusions of a systematic review depends on the quality of the included studies, and one could wonder whether the methodological quality of RCTs improved over the years. The present study analysed the correlation of the different quality scores with the year of publication and showed improvement of the methodological quality of RCTs as assessed by the Delphi list, the Megens & Harris list, the Jadad list and the Risk of Bias list. The correlation between year of publication and the results obtained with the Jadad list was not significant. A possible reason for this finding is the low number of items included (3 items versus 10 or 11 for Delphi and Megens & Harris lists). Similar to our findings, Falagas et al. [28] observed a temporal evolution of methodological quality of RCTs in various research fields (including PT), but he concluded that only certain aspects of the methodological quality improved significantly over time. In our study, we did not

analyse the temporal trend for the different items separately. The results of the study of Falagas et al. may explain the different correlations for the different lists since the contents of the assessment differ per list. However, it must be noted that the 95% confidence intervals around the correlations found in the present study overlap for all lists. Our findings are in contrasts with those of Koes et al. [29] who did not find an association between the year of publication and the methodological quality of physiotherapeutic interventions studies. Although the highest methodological scores were attained during the last decade, Fernández-de-las-Peñas compared the methodological quality of RCTs evaluating PT in tension-type headache, migraine and cervicogenic headache, published before and after 2000 and found no significant differences [30].

## Conclusion

- Hand searching contributes considerably to the search results for RCTs.
- Different quality lists lead to significantly different scores. Therefore, a specific criteria list must be carefully chosen when quality scores are taken into account in drawing conclusions on evidence.
- The quality of RCTs regarding PT for TMD does improve over time if assessed by the Delphi list, the Megens & Harris list and the Risk of Bias list.

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Okeson JP (1996) Orofacial pain: guidelines for assessment, diagnosis and management. Quintessence, Chicago, pp 113–184
2. Panagopoulos P (1952) Physical treatment of trismus and partial ankylosis of the temporomandibular joint. *J Oral Surg* 10:36–39
3. Kopp S (1979) Short term evaluation of counselling and occlusal adjustment in patients with mandibular dysfunction involving the temporomandibular joint. *J Oral Rehabil* 6:101–109
4. Stenn PG, Mothersill KJ, Brooke RI (1979) Biofeedback and a cognitive behavioral approach to treatment of myofascial pain dysfunction syndrome. *Behav Ther* 10:29–36
5. Moher D, Jadad AR, Tugwell P (1996) Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care* 12:195–208
6. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 17:1–12
7. Chalmers TC, Smith H Jr, Blackburn B, Silverman B, Schroeder B, Reitman D, Ambroz A (1981) A method for assessing the quality of a randomized control trial. *Control Clin Trials* 2:31–49
8. Colditz GA, Miller JN, Mosteller F (1989) How study design affects outcomes in comparisons of therapy. I: Medical Stat Med 8:441–454
9. Lumley J, Bastian H (1996) Competing or complementary? Ethical considerations and the quality of randomized trials. *Int J Technol Assess Health Care* 12:247–263
10. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, Knipschild PG (1998) The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol* 51:1235–1241
11. The Cochrane Collaboration (2009) Cochrane handbook for systematic reviews of interventions. Available at <http://www.cochrane-handbook.org/>
12. Megens A, Harris SR (1998) Physical therapist management of lymphedema following treatment for breast cancer: a critical review of its effectiveness. *Phys Ther* 78:1302–1311
13. Harris SR (1996) How should treatments be critiqued for scientific merit? *Phys Ther* 76:175–181
14. Law M, Stewart D, Letts L, Pollock N, Bosch J, Westmorland M (1998) Guidelines for critical review form-quantitative studies. Available at <http://fhs.mcmaster.ca/rehab/ebp/pdf/qualguidelines.pdf>
15. Portney LG, Watkins MP (2000) Foundations of clinical research, application of practice, 2nd edn. Prentice Hall, Englewood Cliffs, pp 557–586
16. Egger M, Smith GD (1998) Bias in location and selection of studies. *BMJ* 316:61–66
17. Hopewell S, Clarke MJ, Lefebvre C, Scherer RW (2008) Hand searching versus electronic searching to identify reports of randomized trials. *Cochrane Database of Syst Rev* (4):MR000001
18. Richards D (2008) Hand searching still a valuable element of the systematic review. Does hand searching identify more randomized controlled trials than electronic searching? *Evid Based Dent* 9:85
19. Crumley ET, Wiebe N, Cramer K, Klassen TP, Hartling L (2005) Which resources should be used to identify RCT/CCTs for systematic reviews: a systematic review. *BMC Med Res Methodol* 5:24
20. Al-Hajeri AA, Fedorowicz Z, Amin FA, Eisinga A (2006) The hand searching of 2 medical journals of Bahrain for reports of randomized controlled trials. *Saudi Med J* 27:526–530
21. McDonald S, Lefebvre C, Antes G, Galandi D, Gøtzsche P, Hammarquist C, Haugh M, Jensen KL, Kleijnen J, Loep M, Pistotti V, Rütthter A (2002) The contribution of hand searching European general health care journals to the Cochrane Controlled Trials Register. *Eval Health Prof* 25:65–75
22. Blümle A, Antes G (2008) Handsuche nach randomisierten kontrollierten Studien in deutschen medizinischen Zeitschriften. *Dtsch Med Wochenschr* 133:230–234
23. Bickley SR, Harrison JE (2003) How to ... find the evidence. *J Orthod* 30:72–78
24. de Bie RA (1996) Methodology of systematic reviews: an introduction. *Phys Ther Rev* 1:47–51
25. Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S (1995) Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials* 16:62–73
26. Carmeli E, Sheklow SL, Bloomenfeld I (2001) Comparative study of repositioning splint therapy and passive manual range of motion techniques for anterior displaced temporomandibular discs with unstable excursive reduction. *Physiotherapy* 87:26–37
27. Randal J (1998) How randomized clinical trials came into their own. *J Natl Cancer Inst* 90:1257–1258

28. Falagas ME, Grigori T, Ioannidou E (2009) A systematic review of trends in the methodological quality of randomized controlled trials in various research fields. *J Clin Epidemiol* 62:227–231
29. Koes BW, Bouter LM, van der Heijden GJ (1995) Methodological quality of randomized clinical trials on treatment efficacy in low back pain. *Spine* 20:228–235
30. Fernández-de-las-Peñas C, Alonso-Blanco C, San-Roman J, Miangolarra-Page JC (2006) Methodological quality of randomized controlled trials of spinal manipulation and mobilization in tension-type headache, migraine, and cervicogenic headache. *J Orthop Sports Phys Ther* 36:160–169

### References of the studies included in the systematic review on PT for TMD

- Al Badawi EA, Mehta N, Forgbione AG, Lobo SL, Zawawi KH (2004) Efficacy of pulsed radio frequency energy therapy in temporomandibular joint pain and dysfunction. *Cranio* 22:10–20
- Alvarez-Arenal A, Junquera LM, Fernandez JP, Gonzales I, Olay S (2002) Effect of occlusal splint and transcutaneous electric nerve stimulation on the signs and symptoms of temporomandibular disorders in patients with bruxism. *J Oral Rehab* 29:858–863
- Bakke M, Eriksson L, Thorsen NM, Sewerin I, Petersson A, Wagner A (2008) Modified condylotomy versus conventional conservative treatment in painful reciprocal clicking—a preliminary prospective study in eight patients. *Clin Oral Invest* 12:353–359
- Bender T, Gidofalvi E (1991) The effect of ultrasonic therapy in rheumatoid arthritis of the temporomandibular joint [Az ultrahang terapias hatasanak vizsgalata temporomandibularis izületre rheumatoid arthritisen]. *Fogorv Sz* 84:229–232
- Bertolucci LE, Grey T (1995) Clinical comparative study of microcurrent electrical stimulation to mid-laser and placebo treatment in degenerative joint disease of the temporomandibular joint. *Cranio* 13:116–120
- Brooke RI, Stenn PG (1983) Myofascial pain dysfunction syndrome. How effective is biofeedback-assisted relaxation training? In: Bonica JJ et al, editor(s). *Adv Pain Res Ther* Vol. 5. Raven Press, New York, pp 809–812
- Burgess JA, Sommers EE, Truelove EL, Dworkin SF (1988) Short-term effect of two therapeutic methods on myofascial pain and dysfunction of the masticatory system. *J Prosthet Dent* 60:606–610
- Carlson CR, Bertrand PM, Ehrlich AD, Maxwell AW, Burton RG (2001) Physical self-regulation training for the management of temporomandibular disorders. *J Orofac Pain* 15:47–55
- Carmeli E, Sheklow SL, Bloomenfeld I (2001) Comparative study of repositioning splint therapy and passive manual range of motion techniques for anterior displaced temporomandibular discs with unstable excursive reduction. *Physiotherapy* 87:26–37
- Conti PC (1997) Low level laser therapy in the treatment of temporomandibular disorders (TMD): a double-blind pilot study. *Cranio* 15:144–149
- Crockett DJ, Foreman ME, Alden L, Blasberg B (1986) A comparison of treatment modes in the management of myofascial pain dysfunction syndrome. *Biofeedback Self Regul* 11:279–292
- Dahlström L, Carlsson SG (1984) Treatment of mandibular dysfunction: the clinical usefulness of biofeedback in relation to splint therapy. *J Oral Rehabil* 11:277–284
- Dalen K, Ellertsen B, Espelid I, Gronningsaeter AG (1986) EMG feedback in the treatment of myofascial pain dysfunction syndrome. *Acta Odontol Scand* 44:279–284
- De Abreu VR, Camparis CM, De Fatima ZR (2005) Low intensity laser therapy in the treatment of temporomandibular disorders: a double-blind study. *J Oral Rehabil* 32:800–807
- De Laat A, Stappaerts K, Papy S (2003) Counselling and physical therapy as treatment for myofascial pain of the masticatory system. *J Orofac Pain* 17:42–49
- Doğu B, Yilmaz F, Karan A, Ergöz E, Kuran B (2009) Comparative the effectiveness of occlusal splint and TENS treatments on clinical findings and pain threshold of temporomandibular disorders secondary to bruxism [Bruksizme baliğ temporomandibuler rahatsızlığında oklüzal splint ve TENS tedavilerinin klinik ve ağrı eşiği üzerine olan etkinliklerinin karşılaştırılması]. *Türk Fiz Tip Rehab Derg* 55:1–7
- Dohrmann RJ, Laskin DM (1978) An evaluation of electromyographic biofeedback in the treatment of myofascial pain-dysfunction syndrome. *JADA* 96:656–662
- Dworkin F, Turner JA, Wilson L, Massoth D, Whitney C, Huggins KH et al (1994) Brief group cognitive-behavioral intervention for temporomandibular disorders. *Pain* 59:175–187
- Dworkin SF, Turner JA, Mancl L, Wilson L, Massoth D, Huggins KH et al (2002) A randomized clinical trial of a tailored comprehensive care treatment program for temporomandibular disorders. *J Orofac Pain* 16:259–276
- Dworkin SF, Huggins KH, Wilson L, Mancl L, Turner J, Massoth D et al (2002) A randomized clinical trial using research diagnostic criteria for temporomandibular disorders-Axis II to target clinic cases for a tailored self-care TMD treatment program. *J Orofac Pain* 16:48–64
- Erlanson PM, Poppen R (1989) Electromyographic biofeedback and rest position training of masticatory muscles in myofascial pain-dysfunction patients. *J Prosthet Dent* 62:335–338
- Funch DP, Gale EN (1984) Biofeedback and relaxation therapy for chronic temporomandibular joint pain: predicting successful outcomes. *J Consult Clin Psychol* 52:928–935
- Gardea MA, Gatchel RJ, Mishra KD (2001) Long-term efficacy of biobehavioral treatment of temporomandibular disorders. *J Behav Med* 24:341–359
- Gavish A, Winocur E, Astandzelov-Nachmias T, Gazit E (2006) Effect of controlled masticatory exercise on pain and muscle performance in myofascial pain patients: a pilot study. *Cranio* 24:184–191
- Glaros AG, Weroha NK, Lausten L, Franklin KL (2007) Comparison of habit reversal and a behaviorally-modified dental treatment for temporomandibular disorders: a pilot investigation. *Appl Psychophysiol Biofeedback* 32:149–154
- Gray RJ, Qalby AA, Hall CA, Schofield MA (1994) Physiotherapy in the treatment of temporomandibular joint disorders: a comparative study of four treatment methods. *Brit Dent J* 176:257–261
- Ismail F, Demling A, Heßling K, Fink M, Stiesch-Scholz M (2007) Short-term efficacy of physical therapy compared to splint therapy in treatment of arthrogenous TMD. *J Oral Rehabil* 34:807–813
- Kavuncu V, Danisger S, Kozkciolu M, Omer S, Aksoy C, Yucel K (1994) Comparison of the efficacy of TENS and ultrasound in temporomandibular joint dysfunction syndrome. *Romatoloji ve tibbi rehabilitasyon dergisi* 5:38–42
- Klobas L, Axelsson S, Tegelberg A (2006) Effect of therapeutic jaw exercise on temporomandibular disorders in individuals with chronic whiplash-associated disorders. *Acta Odontol Scand* 64:341–347
- Komiyama O, Kawara M, Arai M, Asano T, Kobayashi K (1999) Posture correction as part of behavioural therapy in treatment of myofascial pain with limited opening. *J Oral Rehabil* 26:428–35
- Kopp S (1979) Short term evaluation of counselling and occlusal adjustment in patients with mandibular dysfunction involving the temporomandibular joint. *J Oral Rehabil* 6:101–109
- Kruger LR, Van der Linden WJ, Cleaton-Jones PE (1998) Transcutaneous electrical nerve stimulation in the treatment of myofascial pain dysfunction. *South African J Surg* 36:35–39



- Kulekcioglu S, Sivrioglu K, Ozcan O, Parlak M (2003) Effectiveness of low-level laser therapy in temporomandibular disorder. *Scand J Rheumatol* 32:114–118
- Linde C, Isacson G, Jonsson BG (1995) Outcome of 6-week treatment with transcutaneous electric nerve stimulation compared with splint on symptomatic temporomandibular joint disk displacement without reduction. *Acta Odontol Scand* 53:92–98
- Magnusson T, Syrén M (1999) Therapeutic jaw exercises and interocclusal appliance therapy. *Swed Dent* 23:27–37
- Maloney GE, Mehta N, Forgiione AG, Zawawi KH, Al Badawi EA, Driscoll SE (2002) Effect of a passive jaw motion device on pain and range of motion in TMD patients not responding to flat plane intraoral appliances. *Cranio* 20:55–66
- Mazzetto MO, Carrasco TG, Bidinelo EF, deAndrade Pizzo RC, Mazzetto RG (2007) Low intensity laser application in temporomandibular disorders: a phase I double-blind study. *Cranio* 25:186–193
- Michelotti A, Steenks MH, Farella M, Parisini F, Cimino R, Martina R (2004) The additional value of a home physical regimen versus patient education only for the treatment of myofascial pain of the jaw muscles: short-term results of a randomized clinical trial. *J Orofac Pain* 18:114–125
- Minakuchi H, Kuboki T, Matsuka Y, Maekawa K, Yatani H, Yamashita A (2001) Randomized controlled evaluation of non-surgical treatments for temporomandibular joint anterior disk displacement without reduction. *J Dent Res* 80:924–928
- Monteiro AA, Clark GT (1988) Mandibular movement feedback vs occlusal appliances in the treatment of masticatory muscle dysfunction. *J Craniomandib Disord* 2:41–48
- Moystad A, Krogstad BS, Larheim TA (1990) Transcutaneous nerve stimulation in a group of patients with rheumatic disease involving the temporomandibular joint. *J Prosthet Dent* 64:596–600
- Mulet M, Decker KL, Look JO, Lenton PA, Schiffman EL (2007) A randomized clinical trial assessing the efficacy of adding 6×6 exercises to self-care for the treatment of masticatory myofascial pain. *J Orofac Pain* 21:318–329
- Nunez SC, Garcez AS, Suzuki SS, Ribeiro MS (2006) Management of mouth opening in patients with temporomandibular disorders through low-level laser therapy and transcutaneous electrical neural stimulation. *Photomed Laser Surg* 24:45–49
- Okeson JP, Moody PM, Kemper JT, Haley JV (1983) Evaluation of occlusal splint therapy and relaxation procedures in patients with temporomandibular disorders. *JADA* 107:420–425
- Olsen RE, Malow RM (1987) Effects of biofeedback and psychotherapy on patients with myofascial pain dysfunction who are nonresponsive to conventional treatments. *Rehabil Psychol* 32:195–205
- Peroz I, Chun YH, Karageorgi G, Schwerin C, Bernhardt O, Roulet JF et al (2004) A multicenter clinical trial on the use of pulsed electromagnetic fields in the treatment of temporomandibular disorders. *J Prosthet Dent* 7:180–188
- Reid KI, Dionne RA, Sicard-Rosenbaum L, Lord D, Dubner RA, Bethesda (1994) Evaluation of iontophoretically applied dexamethasone for painful pathologic temporomandibular joints. *Oral Surg Oral Med Oral Pathol* 77:605–609
- Schiffman EL, Braun BL, Lindgren BR (1996) Temporomandibular joint iontophoresis: a double-blind randomized clinical trial. *J Orofac Pain* 10:157–165
- Shin SM, Choi JK (1997) Effect of indomethacin phonophoresis on the relief of temporomandibular joint pain. *Cranio* 15:345–349
- Stam HJ, McGrath PA, Brooke RI (1984) The effect of a cognitive-behavioral treatment program on temporo-mandibular pain and dysfunction syndrome. *Psychosom Med* 46:534–545
- Stegenga B, de Bont LG, Dijkstra PU, Boering G (1993) Short-term outcome of arthroscopic surgery of temporomandibular joint osteoarthritis and internal derangement: a randomized controlled clinical trial. *Brit J O Maxillofac Surg* 31:3–14
- Stenn PG, Mothersill KJ, Brooke RI (1979) Biofeedback and a cognitive behavioral approach to treatment of myofascial pain dysfunction syndrome. *Behav Ther* 10:29–36
- Talaat AM, el-Dibany MM, el-Garf A (1986) Physical therapy in the management of myofascial pain dysfunction syndrome. *Ann Otol Rhinol Laryngol* 95:225–228
- Taube S, Ylipaavalniemi P, Kononen M, Sunden B (1988) The effect of pulsed ultrasound on myofascial pain. A placebo controlled study. *Proc Finnish Dent Soc* 84:241–246
- Taylor K, Newton RA, Personius WJ, Bush FM (1987) Effects of interferential current stimulation for treatment of subjects with recurrent jaw pain. *Phys Ther* 67:346–350
- Taylor M, Suvinen T, Reade P (1994) The effect of Grade IV distraction mobilisation on patients with temporomandibular pain-dysfunction disorder. *Physiotherapy* 10:129–136
- Townsend D, Nicholson RA, Buenaver L, Bush F, Gramling S (2001) Use of a habit reversal treatment for temporomandibular pain in a minimal therapist contact format. *J Behav Ther Exp Psychiatry* 32:221–239
- Treacy K (1999) Awareness/relaxation training and transcutaneous electrical neural stimulation in the treatment of bruxism. *J Oral Rehabil* 26:280–287
- Truelove E, Hanson-Huggins K, Mancl L, Dworkin SF (2006) The efficacy of traditional, low-cost and nonsplint therapies for temporomandibular disorder: a randomized controlled trial. *JADA* 137:1099–1107
- Tullberg M, Alstergren PJ, Ernberg MM (2003) Effect of low-power laser exposure on masseter muscle pain and microcirculation. *Pain* 105:89–96
- Turk DC, Zaki HS, Rudy TE (1993) Effects of intraoral appliance and biofeedback/stress management alone and in combination in treating pain and depression in patients with temporomandibular disorders. *J Prosthet Dent* 70:158–164
- Turner AJ, Mancl LL, Aaron LA (2006) Short- and long-term efficacy of brief cognitive-behavioral therapy for patients with chronic temporomandibular disorder pain: a randomized, controlled trial. *Pain* 121:181–194
- Wahlund K, List T, Larsson B (2003) Treatment of temporomandibular disorders among adolescents: a comparison between occlusal appliance, relaxation training, and brief information. *Acta Odontol Scand* 61:203–212
- Wright E, Anderson G, Schulte J (1995) A randomized clinical trial of intraoral soft splints and palliative treatment for masticatory muscle pain. *J Orofac Pain* 9:192–199
- Wright E, Domenech MA, Fisher JR (2000) Usefulness of posture training for patients with temporomandibular disorders. *JADA* 131:202–211
- Yoshida H, Fukumura Y, Suzuki S, Fujita S, Kenzo O, Yoshikado R, Nakagawa M, Inoue A, Sako J, Yamada K, Morita S (2005) Simple manipulation therapy for temporomandibular joint internal derangement with closed lock. *Asian Ass Oral Maxillofac Surg* 17:256–260
- Yuasa H, Kurita K (2001) Randomized clinical trial of primary treatment for temporomandibular joint disk displacement without reduction and without osseous changes: a combination of NSAIDs and mouth-opening exercise versus no treatment. *Oral Surg Oral Med Oral Pathol Oral Radiol Endo* 91:671–675