Impact of quality items on study outcome. Treatment in acute lateral ankle sprains.

International Journal of Technology Assessment in Health Care, **16:4** (2000), 1136–1146.

IMPACT OF QUALITY ITEMS ON STUDY OUTCOME

Treatments in Acute Lateral Ankle Sprains

Arianne P. Verhagen
Maastricht University and Erasmus University Rotterdam

Robert A. de Bie
Maastricht University

Anton F. Lenssen
Maastricht University Hospital

Henrica C. W. de Vet
Maastricht University and Vrije Universiteit

Alphons G. H. Kessels
Academic Medical Hospital

Maarten Boers
Vrije Universiteit and University Hospital

Piet A. van den Brandt
Maastricht University

Abstract

Objective: This study investigates the influence of different aspects of methodologic quality on the conclusions of a systematic review concerning treatments of acute lateral ankle sprain.
Method: A data set of a systematic review of 44 trials was used, of which 22 trials could be included in this study. Quality assessment of the individual studies was performed using the Delphi list. We calculated effect sizes of the main outcome measure in each study in order to evaluate the relationship between overall quality scores and outcome. Next, we investigated the impact of design attributes on pooled effect sizes by subgroup analysis.
Results: The quality of most studies (82%) was low; only 4 of 22 trials were of high quality. Studies with proper randomization and blinding procedure produce a slightly higher (not statistically significant) effect estimate compared to the other studies.
Conclusion: Previous research has suggested that methodologically poorly designed studies tend to over-estimate the effect estimate. Our study does not confirm these conclusions.

Keywords: Review, Methodology, Randomization, Blinding, Effect sizes

Are the conclusions of a systematic review influenced by the methodologic quality of the included studies? Some researchers have not found any difference in results between studies of good and poor quality (1;8;9). Others found that the methodologically sound studies showed less positive treatment effects compared to studies with poor quality (4;10;11;14), or vice versa (2).

Some research on the relationship between design attributes, such as randomization, and outcome has already been performed. Kunz et al. (13) systematically summarized empirical studies comparing randomized versus nonrandomized clinical trials and trials with adequate concealed allocation versus inadequate concealed allocation procedures. Of the latter comparison, most studies indicate that inadequate concealment results in an overestimation of effects (5;16;17). Kunz et al. (13) concluded in their methodologic review that "failure to use adequate concealed random allocation can distort the apparent effects of care in either direction."

A number of studies have examined the influence of blinding on outcome. Shapiro & Shapiro (19) found that when the observer was blinded, the effect sizes were smaller compared to studies with an unblinded observer. Colditz et al. (6) and Miller et al. (15) found conflicting results in their studies. Colditz et al. (6) found lower effect sizes in double-blinded studies compared with no blinding, while Miller et al. (15) found that "double blind comparisons produced the largest average gains (effect sizes), significantly larger than the average for comparisons that involved no blinding." Schulz et al. (18) concluded that lack of double blinding resulted in an overestimation of effects. Also, Linde et al. (14) concluded that, of the design characteristics tested using meta-regression techniques, double blinding had the strongest influence on outcome; namely, double-blinded trials produced less positive results.

Most studies on design characteristics are performed outside the context of a specific therapeutic research question (16;17;18). In this study we chose to place our research within a specific research question, i.e., the efficacy of conservative treatments in acute lateral ankle sprains, to create a more homogeneous data set concerning patients, disease, interventions, and outcome measures. The purpose of this study is to evaluate whether overall trial quality and trial design attributes, such as the randomization procedure and blinding, have an impact on the conclusion of a systematic review.

METHODS

Studies

We used a data set from a systematic review (de Bie RA, Verhagen AP, et al. Efficacy of conservative interventions in the treatment of acute lateral ankle sprains: A systematic review. Unpublished.) of 44 randomized clinical trials on the efficacy of conservative interventions in the treatment of acute lateral ankle sprains. All studies were randomized and compared conservative treatment with either no treatment, a placebo, or nonsurgical treatment. Trials were excluded from this study when they presented a withdrawal rate greater than 50% or when no effect sizes could be calculated.

Assessment of Methodologic Quality

For the assessment of the methodologic quality of individual studies, we used the Delphi list presented in Table 1 (21). The quality score consists of the number of items satisfied and ranges from 0–9. The assessment of the studies was performed independently by two of the authors (APV, AFL) followed by a consensus meeting. Both reviewers have performed quality assessment in other reviews and were therefore regarded as relatively experienced. For the component analysis, studies were divided into several categories according to items concerning the randomization procedure, blinding, and the analysis used. By appropriate

Table 1. The Delphi List Used to Assess the Methodologic Quality of the Investigated Studies

1. Treatment allocation	
a) Was a method of randomization performed?	Yes/No/Don't know
b) Was the treatment allocation concealed?	Yes/No/Don't know
2. Were the groups similar at baseline regarding the most important prognostic indicators?	Yes/No/Don't know
3. Were eligibility criteria specified?	Yes/No/Don't know
4. Was the outcome assessor blinded?	Yes/No/Don't know
5. Was the care provider blinded?	Yes/No/Don't know
6. Was the patient blinded?	Yes/No/Don't know
7. Were point estimates and measures of variability presented for the primary outcome measures?	Yes/No/Don't know
8. Did the analysis include an intention-to-treat analysis?	Yes/No/Don't know

randomization we mean that information about a proper randomization procedure is presented in the paper, instead of just using the word *random*. By concealed randomization we mean that a random (unpredictable) allocation sequence is generated by an independent person not responsible for determining eligibility of the patients, and this sequence is concealed until allocation occurs (17). For blinding we divided the studies into two main categories: blinding reported or not reported. When blinding is reported, we note whether the observer was blinded or the term *double blind* was used. For the statistical analysis items, we divided the studies into two categories: performance of an intention-to-treat (ITT) analysis or not.

Statistical Methods

For the primary outcome measures we calculated the effect sizes and their 95% confidence intervals (CI) according to the methods described in Cooper and Hedges (7). These effect sizes transform the results of continuous data from any parallel group comparison into a standardized metric. For pooling we used one effect size out of each study. When one study allowed for calculating two or more effect sizes, we preferably used pain as primary outcome measure, and swelling if pain was not reported in the original study. Pooled effect sizes were calculated according to a random effects model (7). In a funnel plot we evaluated the possibility of publication bias in this review. If there is publication bias in a meta-analysis, the funnel plot will often be skewed and asymmetrical (8).

Next we calculated overall quality scores (QS) for the individual studies by summing up the "yes" scores. In advance, a cut-off point between "high" and "low" quality studies is set at 50% of the maximum achievable score of 9 points, meaning high-quality studies scored 5 points or more and low-quality studies 4 points or less. For the analysis of major components of quality, i.e., randomization, blinding, and an ITT analysis, we performed component analysis. For the pooling we used the primary outcome of each study.

RESULTS

Studies

Only 23 of the 44 studies allowed for calculation of effect sizes for one or more outcome measures and were included. One study of the 23 (44), is excluded from the analysis because of a high withdrawal rate: over 60% loss to follow-up after 3 months and over 80% after 1 year. The sample of excluded studies was comparable with the included ones concerning patient characteristics, randomization schedule, blinding, interventions, and outcome measures.

In total, five studies compared an intervention such as "short wave" or "laser therapy" with a placebo and six studies compared "brace," "tape," or "bandage" with "cast" or

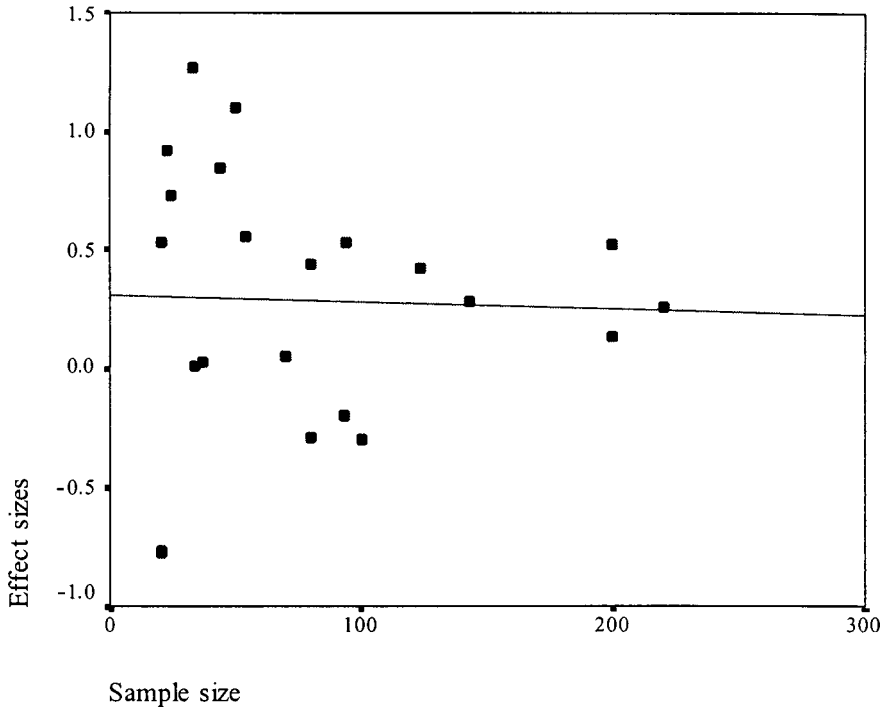

Figure 1. Funnel plot of sample size against effect sizes of the individual studies.

"plaster." In nine studies pain was reported as the main outcome measure, and in nine studies swelling was reported as an outcome measure. All studies included patients with acute ankle sprains (<48 hours).

Effect Sizes

In the 22 included studies we were able to calculate 27 effect sizes. Of these 27 effect sizes, eight (29.6%) were negative, suggesting an effect in favor of the control group. Only one outcome measure of each study, preferably pain, is used in the analysis. In Table 2 we present the characteristics of all 22 studies and their main outcome measures, and effect sizes. In order to assess potential publication bias, Figure 1 presents the funnel plot of the effect sizes, as presented in Table 2, against the sample size. The sample size is presented horizontally and the effect sizes vertically. The funnel plot shows no asymmetry; therefore, we assume that our meta-analysis is probably not biased. The pooled estimate or the overall average of effects is 0.28 (95% CI: 0.08–0.49).

Quality Scores

Table 2 presents the characteristics of the studies included. The Delphi quality scores range from 1 to 6 points. The mean QS is 3.6, which is low compared to the maximum achievable score of 9 points. Only 23% of the studies reported information about the method of randomization and another 23% of the reports presented any information about blinding procedures. Two trials reported information concerning both the randomization and blinding procedures. All high-quality trials (n = 4) are placebo controlled trials. In the subgroup of non–placebo-controlled trials, the mean QS is 3.1 (median = 3).

The two reviewers, who assessed the articles independently, had an initial agreement on the Delphi criteria list of approximately 95%. The 5% disagreement occurred mostly

Table 2. Characteristics of the Studies, Ranked According to the Delphi Quality Score

Study	Sample size	Randomization	Intervention	Blinding	Outcome measures	Intention-to-treat	Effect size (95% CI)	Delphi QS
McGill (34)	n = 37	Concealed	Pulsed short wave vs placebo	Care provider	Pain	No	0.02 (−0.62, 0.66)	6
de Bie (28)	n = 38	Unknown	Laser vs placebo (vs tape)	Care provider and patient	Pain	Yes	0.92 (0.05, 1.78)	6
Pennington (38)	n = 50	Unknown	Diapulse vs placebo	Observer and patient	Swelling	Yes	1.2 (0.6, 1.8)	6
Bradnock (24)	n = 47	Unknown	Low freq ultrasound vs placebo (vs long wave)	Patient	Gait	Yes	0.01 (−0.64, 0.68)	5
Hedges (29)	n = 121	Unknown	Bandage vs plaster	Unknown	Pain [Swelling]	No	−0.19 (−0.59, 0.21) [−0.27 (−0.67, 0.13)]	4
Coté (27)	n = 30	Unknown	Cold vs heat (vs bath)	Unknown	Swelling	Yes	−0.78 (−1.68, 0.12)	4
Sloan (41)	n = 143	Adequate	Cold vs stimulated ther.	Observer	Swelling	No	0.28 (−0.03, 0.59)	4
Rucinski (39)	n = 30	Unknown	Wrap vs elevation (vs intermitt. compression)	Unknown	Swelling	Yes	−0.79 (−1.69, 0.12)	4
O'Hara (37)	n = 220	Adequate	Malleotrain vs tubigrip	Unknown	Pain	Yes	0.26 (0.01, 0.51)	4
Axelsen (23)	n = 48	Adequate	Laser vs placebo	Unknown	Pain [Swelling]	No	1.26 (−0.36, 2.15) [−0.04 (−0.87, 0.79)]	4
Sommer (42)	n = 120	Unknown	Aircast vs plaster (vs cast)	Unknown	Loss of working hours	No	−0.29 (−0.72, 0.14)	4

Brakenbury (25)	n = 400	Unknown	Bandage (+ chymoral or placebo) vs cast (+ chymoral or placebo)	Unknown	Range of motion	No	0.52 (0.25, 0.79)	3
Muwanga (36)	n = 144	Unknown	Nottingham ankle support vs tubigrip (vs strapping)	Unknown	Range of motion	No	0.53 (0.12, 0.94)	3
Michlovitz (35)	n = 30	Unknown	Ice vs ice + (placebo or real) pulsed stimulation	Unknown	Pain [Swelling]	Yes	0.58 (0.31, 1.47) [−0.13 (−1.0, 0.74)]	3
Airaksinen (22)	n = 44	Unknown	Bandage + compression vs bandage alone	Unknown	Pain [Swelling]	No	0.85 (0.23, 1.46) [0.85 (0.23, 1.46)]	3
Klein (32)	n = 60	Concealed	Bandage vs cast	Unknown	Health index	No	0.56 (0.02, 1.1)	3
Hølmer (30)	n = 200	Unknown	Wrap vs bandage	Unknown	Pain [Swelling]	No	0.13 (−0.14, 0.40) [0.06 (−0.21, 0.33)]	3
Scotece (40)	n = 184	Unknown	Gel cast vs daily strapping (vs tape)	Unknown	Return to duty	Yes	0.42 (−0.07, 0.77)	3
Jongen (31)	n = 100	Unknown	Tape vs malleotrain	Unknown	Health index	No	−0.3 (−0.69, 0.09)	3
Konradsen (33)	n = 80	Unknown	Brace vs aircast	Unknown	Pain	No	0.44 (0, 0.88)	2
Wilkerson (43)	n = 34	Unknown	Tape vs compression (with or without ice)	Unknown	Function	No	0.73 (0.09, 1.55)	2
Caro (26)	n = 132	Unknown	Tape vs cast (vs hydrocortisone injection)	Unknown	Time to cure	No	0.08 (−0.38, 0.55)	1

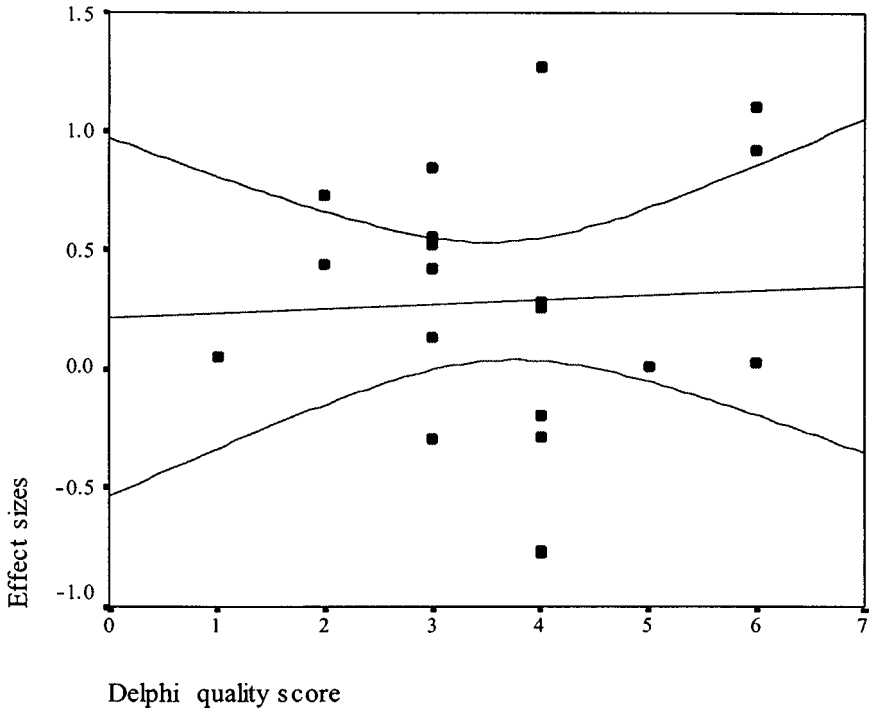

Figure 2. Plot of the Delphi QS of the individual studies against effect size.

because one reviewer had missed some information (4%), but rarely because of a difference in interpretation of the information (1%).

Relationship Between Overall Quality and Outcome

We made a scatter plot between the overall quality scores and effect sizes (Figure 2). The scatter plot shows no relation between the QS and the effect sizes (intercept = 0.217; slope = 0.045). The pooled effect size of high-quality studies (n = 4) is 0.53 (95% CI: −0.21–1.27) and of low-quality studies (n = 18) is 0.19 (95% CI: 0.009–0.38). This difference is not statistically significant. When we divide the subgroup of non–placebo-controlled trials in high (n = 6) and low (n = 11) quality, using a cut-off score of the mean QS (3.1; median = 3), the pooled effect sizes are −0.14 (95% CI: −0.51–0.24) and 0.37 (0.14–0.61), respectively.

Component Analysis

The effect sizes, pooled for subgroups according to the various design attributes, are presented in Figure 3.

Randomization. Of all 22 studies, two reported a concealed randomization procedure (32;34), and three an adequate method of randomization (23;37;41). Because of the small numbers of studies in both categories, we combined them in the component analysis (n = 5). Only one trial with a proper randomization procedure is regarded of high quality. When the randomization method is unknown, the pooled effect size is lower (0.21; 95% CI: 0.00–0.44) than when the method is appropriate or concealed (0.34; 95% CI: 0.02–0.68).

Blinding. When double blinding is mentioned (28;34;38), all studies described at least one level of blinding, and all described the method of blinding. Two studies (38;41) described blinding of the outcome measurement (observer), and two studies (28;38) described

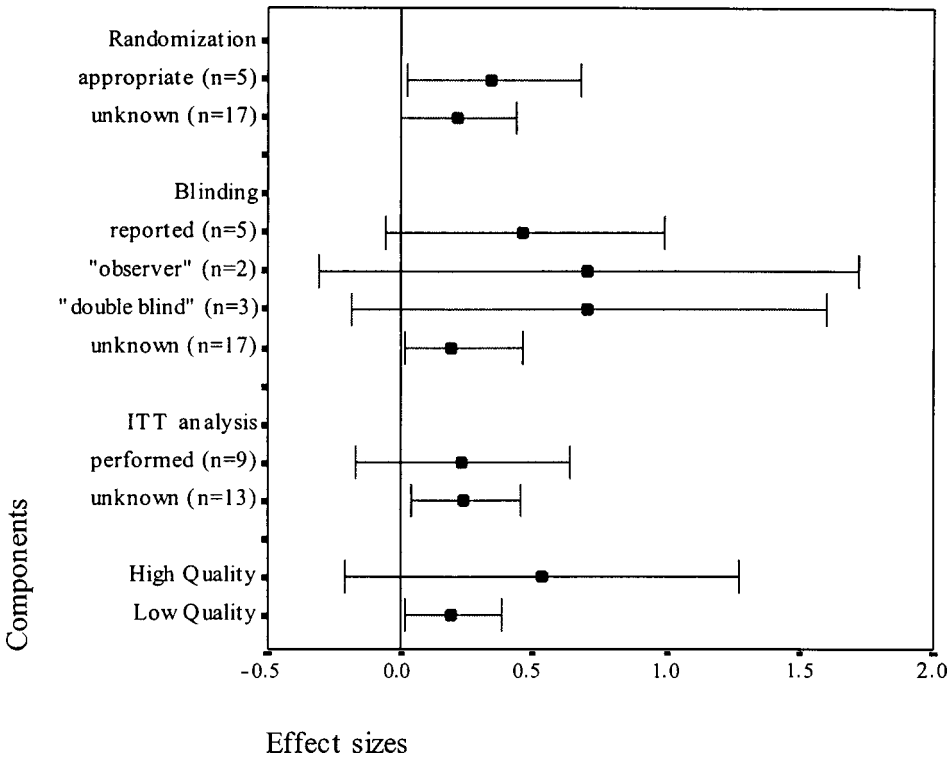

Figure 3. Pooled effect size of all studies according to various design attributes.

blinding of two different levels. One study (28) evaluated whether the blinding procedure was successful. Four of the five studies with a description of blinding are considered of high quality. The pooled effect size for the category "blinding not reported" is lower (0.19; 95% CI: 0.008–0.40) than when blinding is reported (0.46; 95% CI: −0.06–0.99) and is comparably low with the pooled effect size in the category "randomization procedure unknown" (Figure 3).

Analysis. In eight studies the performance of the analysis was carried out according to the ITT principle. There is no difference in pooled effect size between studies with an ITT and those with no ITT analysis (Figure 3).

DISCUSSION

This study evaluates the impact of design characteristics in systematic reviews. Retrospectively, our data set appeared to be a less than ideal set of data for this purpose. Nevertheless, some results are worth discussing. In general, design factors, such as proper randomization and blinding procedures, do influence the interpretation of the results of individual clinical trials. The use of design factors in the interpretation of aggregated research in systematic reviews or meta-analysis is more difficult. In our study the overall methodologic quality scores varied between poor and reasonably good (range between 11% and 66% of maximum available score), but most studies (18 of 22) scored less than half of the maximum available score. More important, we had to exclude 50% of the identified studies solely because of a poor data presentation. Our findings support the conclusion of other researchers that only a few clinical trials meet the minimum standards of methodologic rigor to be validly interpretable from a scientific point of view (3;12;20).

A leading paradigm in empirical research is that clinical trials that do not meet some design criteria, such as concealed randomization or double blinding, will be biased in favor of the intervention, and will therefore more likely produce positive treatment effects. We cannot confirm this paradigm. We found a trend, although not statistically significant, toward a higher effect estimate in high-quality trials. Concerning design factors, we found a trend toward a higher effect size in trials with an appropriate randomization procedure or where blinding was reported, compared to the ones using an unknown randomization and blinding schedule.

There are several possible explanations why our findings do not confirm this paradigm. The validity of our investigation is limited by the small number of trials ($n = 22$), the small number of patients involved, and the quality of the data presented. Our results could be affected by the fact that we had to exclude almost half of our studies, because data enabling calculation of effect sizes was not presented. However, the included and excluded studies were similar with regard to the most important design characteristics. The difference in effect size between high- and low-quality trials might be due to the difference in control group: all high-quality trials were placebo-controlled trials.

Contrary to other studies addressing design characteristics (16;17), we chose to place our research within a specific research question. This, and the exclusion of half of the identified studies, resulted in a loss of power and may have increased the risk of a type II error. Because of the small number of trials, we were unable to perform meta-regression analysis on the separate design characteristics. According to Kunz et al. (13), evidence about the influence of randomization is less clear in comparisons across interventions compared to empirical studies using studies with more or less the same intervention. Combining trials concerning varying interventions in varying diseases or disorders leads to such a large heterogeneity that an estimation of an overall average of effects cannot be given. In that case the assumption that nonconcealed randomized trials, or not-blinded trials, provide an overestimation of the treatment effect cannot be tested.

The problem remains that we do not know what the "true" treatment effect is, we can only estimate it. There are two possible ways of estimating a true effect size. One is to infer it from the methodologic best studies. Another way is to assume that larger studies present a more precise estimate of the true effect size and that the results of small trials show a random variation around this true effect size. We based our decision about the impact of design characteristics on the latter way of inferring a possible true effect size, or the overall average of effects of 0.28 (95% CI: 0.08–0.49), because there were too few adequate methodologic studies in this field. Stating this, the pooled estimates of the studies, which reported a concealed or appropriate randomization or blinding procedure, provide a slightly higher estimate compared to the overall average of effects. The pooled effect estimate of the high-quality studies (0.53) only provides a much higher estimate of the probable treatment effect, although not statistically significant.

Our research, although hampered by a lack of power, contributes to the (still very small) body of scientific knowledge concerning quality assessment in clinical trials. Our data do not confirm the leading paradigm in this field of research, that lack of proper randomization and blinding lead to overestimation of effects. This means that the impact of design characteristics is not clear and simple. We are convinced that empirical research should be performed within a specific research question, hereby preventing heterogeneity caused by a different study population, different interventions, or different outcome measures. The resulting lack of power can be overcome in meta-analysis of empirical studies such as performed by Kunz and colleagues (13).

In conclusion, we found a consistent trend, although not statistically significant, toward a higher effect estimate in studies with higher quality or when randomization and blinding procedures are properly done or described, respectively. The direction in effects we found

were contrary to the ones suggested by the paradigm. Thus, the direction and magnitude of this effect is unpredictable and may depend on the research question. Quality assessment is seen as an important part of a meta-analysis, but the influence of quality on outcome remains unclear and needs further research.

REFERENCES

1. Aker PD, Gross AR, et al. Conservative management of mechanical neck pain: Systematic overview and meta-analysis. *BMJ.* 1996;313:1291-1296.
2. Assendelft WJ, Koes BW, et al. The efficacy of chiropractic manipulation for back pain: Blinded review of relevant randomized clinical trials. *J Manipulative Physiol Ther.* 1992;15:487-494.
3. Assendelft, WJ, Koes BW, et al. The effectiveness of chiropractic for treatment of low back pain: An update and attempt at statistical pooling. *J Manipulative Physiol Ther.* 1996;19:499-507.
4. Beckerman H, de Bie RA, et al. The efficacy of laser therapy for musculoskeletal and skin disorders: A criteria-based meta-analysis of randomized clinical trials. *Phys Ther.* 1992;72: 483-491.
5. Chalmers TC, Celano P, et al. Bias in treatment assignment in controlled clinical trials. *N Engl J Med.* 1993;309:1358-1361.
6. Colditz GA, Miller JN, et al. How study design affects outcomes in comparisons of therapy, I: Medical. *Stat Med.* 1989;8:441-454.
7. Cooper H, Hedges LV. *The handbook of research synthesis*. New York: Russel Sage Foundation; 1994.
8. Egger M, Davey Smith G, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315:629-634.
9. Emerson JD, Burdick E, et al. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. *Control Clin Trials.* 1990;11:339-352.
10. Kleijnen J, Knipschild P, et al. Clinical trials of homoeopathy [published erratum appears in *BMJ*. 1991;302:818] [see comments]. *BMJ.* 1991;302:316-323.
11. Kleijnen J, ter Riet G, et al. Acupuncture and asthma: A review of controlled trials. *Thorax.* 1991;46:799-802.
12. Koes BW, Assendelft WJ, et al. Spinal manipulation for low back pain. An updated systematic review of randomized clinical trials. *Spine.* 1996;21:2860-2871.
13. Kunz R, Oxman AD. The unpredictability paradox: Review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ.* 1998;317:1185-1190.
14. Linde K, Scholz M, et al. Impact of study quality on outcome in placebo-controlled trials of homeopathy. *J Clin Epidemiol.* 1998;52:631-636.
15. Miller JN, Colditz GA, et al. How study design affects outcomes in comparisons of therapy, II: Surgical. *Stat Med.* 1989;8:455-466.
16. Moher D, Pham B, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet.* 1998;352:609-613.
17. Schulz KF, Chalmers I, et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273:408-412.
18. Schulz KF, Grimes DA, et al. Blinding and exclusions after allocation in randomised controlled trials: Survey of published parallel group trials in obstetrics and gynaecology. *BMJ.* 1996;312:742-744.
19. Shapiro DA, Shapiro D. Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychol Bull.* 1982;92:581-604.
20. Tyson JE, Furzan JA, et al. An evaluation of the quality of therapeutic studies in perinatal medicine. *J Pediatr.* 1983;102:10-13.
21. Verhagen AP, de-Vet HCW, et al. The Delphi List: A criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol.* 1998;51:1235-1241.

References of the Systematic Review

22. Airaksinen O, Kolari PJ, Miettinen H. Elastic bandages and intermittent pneumatic compression for treatment of acute ankle sprains. *Arch Phys Med Rehabil.* 1990;71:380-383.

23. Axelsen SM, Bjerno T. Laserbehandling af fodledsdistorsion (in Scandinavian). *Ugesk Laeger.* 1993;155:3908-3911.
24. Bradnock B, Law HT, Roscoe K. A quantitative comparative assessment of the immediate response to high frequency ultrasound and low frequency ultrasound in the treatment of acute ankle sprains. *Physiother.* 1995;81:78-84.
25. Brakenbury PH, Kotowski J. A comparative study of the management of ankle sprains. *Br J Clin Pract.* 1983;37:181-185.
26. Caro D, Craft IL, Howells JB, Shaw PC. Diagnosis and treatment of injury of lateral ligament of the ankle joint. *Lancet.* 1964;1:720-723.
27. Coté DJ, Prentice WE, Hooker DN, Shields EW. Comparison of three treatment procedures for minimizing ankle sprain swelling. *Phys Ther.* 1988;68:1072-1076.
28. de Bie RA, Steenbrugge RA, Bouter LM. Effect van lasertherapie op inversietraumata van de enkel (in Dutch). *Ned Tijdschr v Fysiother.* 1988;98:108-112.
29. Hedges JR. Management of ankle sprains. *Ann Emerg Med.* 1980;9:296-302.
30. Hølmer P, Carstensen NC, Merrild UB. Stottestromper kontra stottebind i behandlingen af akutte ankeldistorsioner. En prospektiv randomiseret undersogelse (in Scandinavian). *Ugesk Laeger.* 1991;153:430-432.
31. Jongen SJM, Pot JH, Dunki Jacobs PB. De behandeling van de verzwikte enkel; Tape versus malleotrain (in Dutch). *Geneesk en Sport.* 1992;25:98-101.
32. Klein J, Rixen D, Albring T, Tiling T. Funktionelle versus Gipsbehandlung bei der frischen Außenbandruptur des oberen Sprunggelenks: Eine randomisierte klinische Studie (in German). *Unfallchirurg.* 1991;94:99-104.
33. Konradsen L, Hølmer P, Søndergaard L. Early mobilizing treatment for grade III ankle ligament injuries. *Foot Ankle.* 1991;12:69-73.
34. McGill SN. The effects of pulsed shortwave therapy on lateral ligament sprain of the ankle. *N Z J Physiother.* 1988:21-24.
35. Michlovitz S, Smith W, Watkins M. Ice and high voltage pulsed stimulation in treatment of acute lateral ankle sprains. *JOSPT.* 1988;9:301-304.
36. Muwanga CL, Quinton DN, Sloan JP, Gillies P, Dove AF. A new treatment of stable lateral ligament injuries of the ankle joint. *Injury.* 1986;17:380-382.
37. O'Hara J, Valle-Jones JC, Walsch H, et al. Controlled trial of an ankle support (Malleotrain) in acute ankle injuries. *Br J Sp Med.* 1992;26:139-142.
38. Pennington GM, Danley DL, Sumko MH, Bucknell A, Nelson JH. Pulsed, non-thermal, high-frequency electromagnetic energy (diapulse) in the treatment of grade I and grade II ankle sprains. *Milit Med.* 1993;158:101-104.
39. Rucinski TJ, Hooker DN, Prentice WE, Shields EW, Coté-Murray D. The effects of intermittent compression on edema in postacute ankle sprains. *JOSPT.* 1991;14:65-69.
40. Scotece GG, Guthrie MR. Comparison of three treatment approaches for grade I and II ankle sprains in active duty soldiers. *JOSPT.* 1992;15:19-23.
41. Sloan JP, Hain R, Pownall R. Clinical benefits of early cold therapy in accident and emergency following ankle sprain. *Arch Emerg Med.* 1989;6:1-6.
42. Sommer HM, Schreiber H. Die fruh funktionelle konservative Therapie der frischen fibularen Kapsel-Band Ruptur aus Sozial konomischer Sicht (in German). *Sportverl Sportschad.* 1993;7:40-46.
43. Wilkerson GB, Horn-Kingerly HM. Treatment of the inversion ankle sprain: Comparison of different modes of compression and cryotherapy. *JOSPT.* 1993;17:240-246.
44. Zeegers AVCM. *Supinatie letsel van de enkel* (in Dutch). Academisch proefschrift, thesis, Utrecht; 1995.