VU Research Portal

Untitled - Response

van Tulder, M.W.; Bouter, L.M.; Koes, B.W.

published in

Spine

1998

DOI (link to publisher)

[10.1097/00007632-199806010-00025](https://doi.org/10.1097/00007632-199806010-00025)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van Tulder, M. W., Bouter, L. M., & Koes, B. W. (1998). Untitled - Response. *Spine*, 23(11), 1290-1291.
<https://doi.org/10.1097/00007632-199806010-00025>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Letters

To the Editor:

The recent work of van Tulder et al¹³ represents an impressive volume of literature summarized in one meta-analysis. For students of conservative spine care, the evolution and steadfast focus of the group from the Netherlands is obvious. Although this reader is biased, generally in favor of the conclusions they have reached, there is reason to be concerned about the care, consistency, and accuracy with which the group has applied their own criteria for methodologic assessment. Obviously, in a meta-analysis, the veracity of the treatment effects being considered do not change. What changes from one meta-analysis to another is the perspective from which the original research studies are evaluated. The science of meta-analysis alternatively may be considered hypothesis testing for the differences in criteria used to include and rank studies. Similarly apparent is that the processes of rating studies are only as good as the consistency by which they are applied.

The recent work headed by van Tulder is one that appears to have grown to its current mass through a sequence of efforts and earlier publications. The authors, however, indicate in Table 1 that, "The operationalization of the criteria has been published in our previous systematic review [references 42-45 cited]." Those references, however, are the works of Evans et al,⁵ Faas Chavannes et al,⁶ Faas Van Eijk et al,⁷ and Farrel and Twomey.⁸

Over the past decade it has been interesting to observe the development of this group's struggle with the literature on back pain. Unfortunately, earlier publications¹² and the mistakes made in applying the methodologic criteria have been carried forward again into this publication. We take specific issue with the incorrect assignment of methodologic criteria scores for the work published by Triano et al.¹¹ Of the 17 criteria reported by van Tulder et al,¹³ we dispute the consistency of application in six. Table 1 lists the disputed criteria and the ranking that was given.

Table 2 quotes or, for longer sections, paraphrases the relevant sections from Triano et al,¹¹ which, we submit, support our contention that the methodologic criteria were applied incorrectly.

In the case of criterion D, there seems to be little explanation for a rating of 0 out of 3. As an author, one is

Table 1. Criteria and Scores

Criteria List	Scores
Study population	
D. Drop-outs described for each study group separately	0 of 3
Interventions	
G. Interventions standardized and described	5 of 10
I. Co-interventions avoided	0 of 5
Effect	
K. Patients blinded	2 of 5
L. Outcome measures relevant	4 of 10
M. Blinded outcome assessment	0 of 10

always concerned with how to be clear in communicating methods and results. It is difficult to conceive of a better way to account for "Drop-outs described for each study group separately" (Criterion G) than to list them explicitly by group. No other report on manipulation published before 1995 has been as detailed in describing the interventions than our report. The descriptions required nearly one full page, and were referenced to standard teaching texts.¹ Moreover, the manual methods (high velocity low force and high velocity low amplitude) were quantified through biomechanical and biologic quality control measures, randomly, throughout the course of the study. Results confirming the stability of these procedures were reported in the text. Co-interventions (Criterion I) and any unblinding (Criterion K) were identified in an exit interview, in which subjects who believed they had not received legitimate treatment were isolated and removed from the data analysis. The outcome measures focused on the standard measures of pain, function, and psychological status, excluding only patient satisfaction (Criteria L). These measures were and are the categories that are considered relevant to back pain treatment studies.^{2-4,10} In fact, they are considered important categories of outcome by the Agency for Health Care and Policy Research,² which was the foundation for the "levels of evidence criteria" used by van Tulder¹² in his original work. Because our population was ambulatory and still working, more aggressive functional outcome such as return-to-work was irrelevant. Finally, one might argue that the inclusion of the methodologist on teams 2 and 5 would create an unblinding of analysis. However, it is clear that such involvement would not allow access to the treatment and

Table 2. Evidence From Triano et al (1995)

Criterion	Evidence of Incorrect Application
D	Page 951: "From all three treatment groups, only 18.7% (7 HVLA; 14 HVLF; 18 BEP) elected to drop out of the study before completion." The n size for each separate group and variable after removing confounding factors from the exit interview were given explicitly in the tables reported by the study.
G	Page 949-50: The HVLA, HVLF mimic, and BEP treatments were described in detail and referenced, where appropriate, to standard teaching texts for the procedures (HVLA). Standardization of treatment was similarly made with respect to frequency of provider contact with the patient and time spent with the patient across all three treatment groups. Page 950: "Quality control measures were put in place to independently verify the performance of the HVLA and HVLF on a random basis." Quality control measures were biomechanical and biologic, with descriptions given and references from detail of previous work provided. "This quality control showed the modifications of the HVLF to be highly successful and further biologic testing was discontinued after 8 months." Figure 2 (page 951) further quantified the differences between the HVLA and HVLF experimental treatments that were administered.
I, K	Page 951: "During the debriefing interview on completion of the study protocol for each participant, 25 patients were identified who had serious confounding factors. Types of confounding factors included unblinding of the participant, new trauma occurring during the intervening time between accepting the patient into the study and protocol completion, and cointerventions capable of altering results on study outcome measures. Data from these subjects were eliminated from analysis"
L	Page 950: "Three primary outcome measures designed to evaluate perceived pain, functional activity, and limited emotional status were selected on the basis of a feasibility study 38, which confirmed the findings of Deyo 15 regarding their reliability and validity within our patient population."
M	Page 949: "Every effort was made to ensure that investigators were masked, by establishing separate evaluation and treatment teams that were rigorously segregated in terms of duties and familiarity with patient information. Patient screening for candidacy was carried out by team 1. Recruitment, primary outcome measures, and administration of the randomized treatment assignment tables were the responsibility of team 2. Secondary outcomes measures were obtained by team 3. Treatment was administered. By team 4 according to random assignment. Finally, data analysis was carried out at the end of the study by team 5, which consisted of the principal investigator and project methodologist. No information crossed team boundaries with the exception of the methodologist who participated on teams 2 and 5."

outcome measures that were obtained by other teams. Moreover, a presumption of bias further presumes that an individual investigator is capable of retaining the details of 209 separate patients interviewed over a period of several months or that such an investigator would decode data intentionally in an effort to bias it.

The problems and cautions related to meta-analysis have been described previously by other authors.⁹ There is no good way to determine what can constitute the correct criteria; therefore multiple reports, even those

from the same authors, can be published, just by changing criteria applications. Does this provide any greater understanding of the issues related to the treatment effects or does it tend to foster greater confusion? This is a greater quandary when there is inconsistency in the application of the selected criteria.

Granted, all writing is not as successful in communication as a writer might prefer, but the issues that are raised in the criteria of van Tulder et al¹³ are addressed prominently in the manuscript by Triano et al.¹¹ Even given different viewpoints on relative merit, it is difficult to understand how criteria D, G, and M were given no value by the reviewers. One must ask how the integrity of this meta-analysis, as a whole, can be presumed if nearly 30% of the criteria judgments made on just one referenced study can be challenged so obviously by a simple review.

John J. Triano, DC
Texas Back Institute
Plano, Texas

References

1. Bergmann TF, Peterson DH, Lawrence DJ. Chiropractic Technique. 1st Edition. New York, Churchill Livingstone, 1993:450-7.
2. Bigos S, Bowyer O, Braen G, et al. Acute Low Back Problems in Adults: Clinical Practice Guideline No. 14. AHCPR Publication no. 95-0642. Rockville, Md: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services. December, 1994.
3. Cherkin DC, Deyo RA, Street JH, Barlow W. Predicting poor outcomes for back pain seen in primary care using patients' own criteria. *Spine* 1996;21:2900-7.
4. Deyo RA. Drug therapy for back pain: Which drugs help which patients? *Spine* 1996;21:2840-8.
5. Evans DP, Burke MS, Newcombe RG. Medicines of choice in low back pain. *Curr Med Res Opin* 1980;6:540-7.
6. Faas A, Chavannes AW, Van Eijk JThM, Gubbels JW. A randomized placebo-controlled trial of exercise therapy in patients with acute low back pain. *Spine* 1993;18:1388-95.
7. Faas A, Van Eijk JThM, Chavannes AW, Gubbels JW. A randomized trial of exercise therapy in patients with acute low back pain. *Spine* 1995;20:941-7.
8. Farrell JP, Twomey LT. Acute low-back pain: Comparison of two conservative treatment approaches. *Med J Aust* 1982;1:160-4.
9. Hasselkorn JK, Turner JA, Diehl PK, Ciol MA, Deyo RA. Meta-analysis: A useful tool for the spine researcher. *Spine* 1994;19(Suppl):S2076-82.
10. Roland M, Morris R. A study of the natural history of low-back pain: Part II. Development of guidelines for trials of treatment in primary care. *Spine* 1983;8:145-50.
11. Triano J, McGregor M, Hondras M, Brennan P. Manipulative therapy versus education programs in chronic low back pain. *Spine* 1995;20:948-55.
12. Van Tulder MW. Diagnostics and treatment of chronic low back pain in primary care. Amsterdam: Thesis Publishers, 1996.
13. Van Tulder MW, Koes BW, Bouter LM. Conservative

treatment of acute and chronic nonspecific low back pain. *Spine* 1997;22:2128-56.

In Response:

We would like to respond to the letter written by Dr. John Triano regarding our article⁷ published in *Spine*.

Dr. Triano's first comment concerns the legend of Table 1, in which we refer to our previous reviews using the wrong references. We are thankful to Dr. Triano for noticing this, and we admit that this must have escaped our attention when checking the proofs of the article. The correct references, however, are given at several places in the text, for example just below Table 1 on page 2129 in the last sentence of the left column and the first sentence of the right column, "the same criteria list was used in our previously published systematic reviews [references 78, 79, 84, and 85 cited]."

The most important issue in Dr. Triano's letter is his concern about the consistency and accuracy of our methodologic quality assessment. Dr. Triano disputes our application of six of the methodologic quality criteria to his randomized clinical trial. However, we believe that the main problem he has with our application does not relate to the consistency of the application, but solely concerns our operationalization of the criteria. We would like to illustrate this with two examples:

1) Regarding item D, the drop-out rate, Dr. Triano states that from all three treatment groups only 18.7% dropped out, and that the corresponding numbers were listed explicitly by group. Our operationalization of this item was "drop-outs described for each study group separately and with reason for withdrawal [references 78, 79, 84, and 85 cited]." In Dr. Triano's study, only 117 of the 209 randomly selected subjects are included in the analysis of disability, and only 129 are included in the analysis of pain. Reasons for noncompliance with treatment or for missing follow-up measurements can be correlated with treatment results. Therefore, the conclusions of this study should be interpreted with caution. In our opinion, Dr. Triano et al have described only some characteristics of a part of the drop-out group; they certainly did not describe all drop-outs for each intervention group with reason for withdrawal. Dr. Triano's study has been included in several systematic reviews,^{1,2,4,5,7} and has been assessed by various reviewers independently from each other. The application of this item was very consistent, not only for reviewers within the same research group or country, but also for reviewers from other groups and countries. All reviewers of Triano's randomized clinical trial scored 0 points for this item!

2) With regard to the relevance of outcome measures, item L, Dr. Triano states in his letter that our score of

4 out of 10 possible points is inconsistent. We have operationalized this item as "the measurement and report of pain, global measure of improvement, functional status (activities of daily living), spinal mobility, and medical consumption (2 points each) [references 78, 79, 84, and 85 cited]." The three outcome measures in Dr. Triano's study were pain (visual analog score), functional status (Oswestry), and psychological depression (modified Zung). Consistent application of our operationalization clearly leads to 4 points: 2 for pain and 2 for functional status.

We think that this clarification concerning two of six disputed criteria is sufficient to make our point: Dr. Triano did not identify inconsistencies; he simply does not agree with our operationalization of the criteria. We are prepared to react in detail to the other four criteria, but we do not consider this to be very informative to the readers of *Spine*.

We are thankful to Dr. Triano for his letter, because it gives us another opportunity to strongly emphasize that the methodology of systematic reviews is still in an early stage of evolution. Within the framework of the Cochrane Collaboration, we have been working on the improvement of the methodology of systematic reviews, including the assessment of the methodologic quality. Just recently in *Spine*, method guidelines were published that were developed by the editorial board of the Cochrane Back Review Group for Spinal Disorders.⁸ These guidelines aim at improving the quality of systematic reviews, facilitating comparisons across reviews, and enhancing consistency among reviewers.

However, these guidelines are not a "gold standard," and do not provide solutions for all methodologic problems. Some subjective decisions are unavoidable in assessing the quality of a study. We can see three categories of problems: the choosing and operationalizing of criteria, its consistent application, and the lack of information in publications. Regarding the first category, Dr. Triano's letter shows that there is still plenty room for discussion. The second category of problems can be solved in part by using two (or more) reviewers and letting them apply the criteria independently and by using a consensus method to resolve disagreements, such as we did in our systematic review. The third category of problems might be tackled by asking the authors of the original studies to provide more information if their reports are unclear and by asking them if they disagree with any of our application of criteria. If the authors disagree, they will be given the opportunity to provide the relevant data from the analysis or research protocol to convince the reviewers that they should change their scores. A relevant recent development, the so-called CONSORT statement, consists of recommendations for the report of randomized controlled trials, which have been accepted as a gold standard by the leading medical journals.³ In the near future, we hope to see this and similar initiatives

being implemented. Clearly this will make further evolution of systematic reviews possible and will enable the acceptance of the methods and results of systematic reviews even by the authors of the randomized clinical trials being reviewed.

Maurits van Tulder, PhD

Lex Bouter, PhD

Bart Koes, PhD

Institute for Research in Extramural Medicine

Vrije Universiteit Amsterdam

Amsterdam

the Netherlands

References

1. Assendelft WJJ, Koes BW, Van der Heijden GJMG, Bouter LM. The effectiveness of chiropractic for treatment of low back pain: An update and attempt of statistical pooling. *J Manipulative Physiol Ther* 1996;19:499-507.
2. Assendelft WJJ, Shekelle PG, Koes BW. Spinal Manipulation for Low Back Pain: Protocol for the Cochrane Collaboration Back Review Group for Spinal Disorders. Oxford, England: Update Software Ltd, Cochrane Library 1997, issue 4.
3. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *JAMA* 1996;276:637-9.
4. Bronfort G, Hass M, Bouter LM, Assendelft WJJ. Efficacy of spinal manipulation and mobilization for low back and neck pain: A systematic review and best evidence synthesis. In: Bronfort, G, ed. *Efficacy of Manual Therapies of the Spine* (Thesis). Amsterdam, the Netherlands: Vrije Universiteit, 1997.
5. Koes BW, Assendelft WJJ, Van der Heijden GJMG, Bouter LM. Spinal manipulation for low back pain: An updated systematic review of randomized clinical trials. *Spine* 1996;21:2860-73.
6. Van Tulder MW. *Diagnostics and Treatment of Chronic Low Back Pain in Primary Care* (Thesis). Amsterdam, the Netherlands: Vrije Universiteit, 1996.
7. Van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain: A systematic review of randomized controlled trials of the most common interventions. *Spine* 1997;22:2128-56.
8. Van Tulder MW, Assendelft WJJ, Koes BW, Bouter LM and the Editorial Board of the Cochrane Collaboration Back Review Group. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine* 1997;22:2323-30.