Conflicting conclusions from two systematic reviews of epidural steroid injections for sciatica: which evidence should general practitioners heed?

KEVORK HOPAYIAN

MIRANDA MUGFORD

SUMMARY

Systematic reviews and meta-analyses are becoming increasingly important in informing clinical practice and commissioning. Two systematic reviews of a treatment for low back pain and sciatica using epidural steroid injections, published in the same year, arrived at conflicting conclusions. Only one was reported in a digest for evidence-based medicine. This paper aims to find the reasons for the discordance between the reviews, and draw conclusions for users of reviews. Using comparative analysis of two published systematic reviews and their source material, it was found that the two reviews had the same overall aims and met the criteria for review methods. They differed in their choice of methods, including the judgement of quality of studies for inclusion and for summing-up evidence. Estimation of summary odds ratios in one review led to stronger conclusions about effectiveness. In conclusion, the choice of methods for systematic review may alter views about the current state of evidence. Users should be aware that systematic reviews include an element of judgement, whatever method is used.

Keywords: sytematic reviews; evidence-based medicine; epidural steroid injections; sciatica.

Introduction

THE time required to search for, read, and synthesize evidence means that, increasingly, clinicians look to systematic reviews to bring order to the welter of original papers which, weed-like, grow ever more numerous and tangled. General practitioners (GPs) in commissioning groups may well look to them for evidence on which to base commissioning decisions. Standards for assessing the quality of reviews have been developed,[1,2] but if two reviews of reasonable quality reach different conclusions about the same topic, what is the busy clinician or GP commissioner to do? The role of systematic review and meta-analysis has been questioned from a range of viewpoints.[3-5] In this paper, we raise questions about the role of such reviews in moulding clinical and commissioning opinion, in the light of recent experience in evaluating treatments for sciatica, on behalf of a GP commissioning group.

K Hopayian, FRCGP, general practitioner, The Surgery, Main Street, Leiston, Suffolk, and GP commissioner, Ipswich and District GP Commissioning Group, Ipswich, Suffolk. M Mugford, DPhil, senior lecturer in health services research, School of Health, University of East Anglia, Norwich.

Submitted: 19 March 1998; final acceptance: 2 July 1998.

Evidence of effectiveness of epidural corticosteroid injections for sciatica

Bandolier is a periodical produced by the NHS Executive Anglia and Oxford which provides a digest of articles relevant to evidence-based medicine, giving brief but straightforward information. In 1996, it reported[6] a meta-analysis of trials of epidural steroid injections in the treatment of sciatica by Watts and Silagy, published in the previous year.[7] The authors of this article had already concluded in their summary:

' ...we present quantitative evidence ... that epidural administration of corticosteroids is effective in management of lumbosacral radicular pain.'

The *Bandolier* editors went on to quote, from a letter written to the same journal in response to the review,[8] a 'number needed to treat' (NNT) which estimated that six patients would need to be treated to see one patient achieve short-term relief. The general reader seeing the review and the *Bandolier* report might have been left with the impression that there is now no doubt that this form of therapy is effective.

A second systematic review of controlled trials of epidural steroid injections in the treatment of low back pain appeared in another journal in the same year as Watts and Silagy's paper.[9] Its conclusion differed:

' ...The best studies showed inconsistent results of epidural steroid injections. The efficacy of epidural steroid injections has not yet been established ... Further research efforts are warranted, but more attention should be paid to the methods of the trials.

In this paper, we investigate the reasons for this discordance and discuss the implications for decision-makers. For brevity, our commentary is restricted to short-term outcomes. For simplicity, we refer to the reviews as 'Watts' and 'Koes'.

Methods

We analysed the trials cited in both reviews to investigate the following possible reasons for the differences between the two reviews:

- aims,
- search criteria,
- criteria for judging quality of trials,
- methods for summarizing evidence, and
- how conclusions are derived.

Where relevant, we re-estimated pooled odds ratios (ORs) using meta-analytic software from Update Software.[10]

Results

Definition of hypothesis and search criteria

Table 1 summarizes features of the two reviews. At first glance, the discordance could be explained by differences in the subject: the Watts study was concerned with sciatica, while the Koes study was concerned with low back pain, with or without sciatica. Furthermore, the Watts study was not restricted to English language papers: two trials were in Polish.[11,12] Both reviews restricted their search to MEDLINE, with additional references from the papers retrieved in these searches. The Watts study also sought unpublished papers, but found none. Both reviews covered broadly the same time period.

Despite differences in the search and selection of trials, it is still fair to compare the two reviews for two reasons. First, the two sets of trials differed very little: nine papers out of 10 were common to the final analysis of short-term outcomes in both reviews. Only one trial was exclusive to the Watts study.[11] Three papers excluded by Watts were included in the Koes study.[13-15] Secondly, although ostensibly restricting their study to sciatica, Watts included onetrial that included patients with spinal stenosis,[16] further diminishing the difference between the two reviews.

Differing methods for selecting and summarizing data in reviews

The approach adopted by the reviewers differed in two important ways: the measurement of methodological quality, and the method for summarizing and/or pooling results. Once both sets of reviewers had decided on these issues, a further decision had to be made about whether and how to derive overall results on short-term pain from the range of methods used in different trials.

The two reviews used different methods to measure methodological quality. Watts used a system which looked only at sources of bias in the estimation of difference in outcome between the treatment and control groups. Three parameters were assessed: the quality of random allocation, the extent to which primary analysis included every patient randomized to each group, and the extent to which assessors were blinded. Each parameter was rated on a scale of one to three. The Koes review measured several additional parameters which indicate other dimensions of the quality of a trial, including sample size, relevance of outcome measures, and the avoidance of co-interventions. These parameters were weighted according to their perceived importance to give a maximum possible score of 100. If the two systems of measuring methodological quality were equivalent, the ranking of their common papers ought to be the same. However, the ranking turns out to be different. Although three papers score highest in both,[17-19] there are noteworthy differences. Three trials[20-22] rank equal first in Watts but all rank fifth or sixth in Koes. One trial[16] ranked fourth by Koes is ranked eighth by Watts: although it was judged as relatively well-designed by Koes, it was judged by Watts to be a more potentially biased study than other trials. The quality score used by Watts did not take account of potential bias from unexplained high drop-out rates and so one trial was ranked equal first,[21] while the Koes review ranked it equal sixth.

The Watts method for summarizing and/or pooling results was a pooled meta-analysis, while Koes constructed a league table arranged by methodological quality. When pooling results, Watts included only those trials achieving a satisfactory score for methodological quality based on minimization of bias, excluding the two papers that had low scores.[12,15] In contrast, Koes ranked each trial by score for methodological quality and divided trials into 'positive' trials, which showed a statistically significant ($P \leq 0.05$) improvement in outcome with epidural steroid, or 'negative' trials: meaning any trial not classed as positive. All trials meeting their initial inclusion criteria were kept in the final analysis.

Scrutiny of the original trials reveals that even those meeting inclusion criteria for the Watts review were of inadequate design. For example, in one trial,[16] short-term outcome was assessed at 24 hours, which is considered by many clinicians to be too short a period to detect the effects of steroids in musculoskeletal conditions. Another trial,[19] ranked first in quality by both reviews, used unvalidated and unrecognized measures of pain relief not capable of easy interpretation, despite the existence at that time of valid pain assessment tools.[23,24] Even excluding these studies, the typical odds ratio for short-term pain relief was 2.4 (95% confidence interval (CI) = 1.61–3.57).

Differing conclusions from the two reviews

Watts' analysis of pain relief found a pooled odds ratio of 2.61 favouring epidural steroids (95% CI = 1.90–3.77). They conclud-

Table 1. Criteria used in two reviews of epidural steroids for sciatica.

Review	Koes et al[9]	Watts and Silagy[7]
Year published	1995	1995
Type of study	RCTs	RCTs
Patients' complaints	Back pain or sciatica	Sciatica
Treatments compared	One arm had one or more epidural steroid injections	One arm had epidural steroid injections
Language of publication	English	Any language
Databases searched	MEDLINE	MEDLINE
Search terms used	MEDLINE keywords (MeSH): 'backache', 'low back pain' (including all minor subheadings)	Medline keywords: 'epidural', 'caudal corticosteroid or methylprednisolone', 'treatment of sciatica (lumbosacral radiculopathy)', in combination with 'randomized', 'double-blind controlled prospective trial'
Unpublished studies	Not sought	Search for unpublished papers by contacting manufacturers of a depot methylprednisolone and 'published authors in the field'
Time period covered by search	1966–1993	'From the time of inception (of epidural steroids) to the present'

Table 2a. Review methods and assessment of their quality.

	Koes et al[7]	Watts and Silagy[9]
Method of review	To compare outcome of results with methodological quality	To pool results and calculate odds ratios of numbers achieving 75% pain relief
Assessment of quality of method	Sixteen parameters differently weighted giving a percentage score (maximum 100) similar to, but not exactly the same as, criteria used in other studies of musculoskeletal and pain therapies	Scale of 0–3 for each of three parameters (maximum of 9) similar to criteria used in most reviews in Cochrane Library[10]

Table 2b. Reasons for inclusion or exclusion of papers.

	Koes et al[9]		Watts and Silagy[7]	
	Study	Reason	Study	Reason
Papers exclusive to one study	Rocco[13] Serrao[14,14]	Broader definition of back pain and control treatment	Popiolek[11] Czarski[12]	Non-English language papers included
Papers excluded from final analysis	None		Czarski[12] Yates[15]	Follow-up period not stated Cross-over design, data could not be extracted

Table 2c. Ranking of common papers according to methodological quality.

Study	Score	
	Koes et al[9] (maximum score = 100)	Watts and Silagy[7] (maximum score = 9)
Snoek[19]	72	9
Mathews[19]	67	9
Breivik[17]	63	9
Cuckler[16]	62	9
Bush[20]	59	9
Klenerman[22]	50	9
Dilke[21]	50	8
Ridley[28]	47	7
Beliveau[29]	45	3
Yates[15]	17	0

ed that epidural steroids are more effective in the short term. Koes found that, out of 12 papers, six showed a statistically significant 'positive' effect for epidural steroids and six did not (referred to as 'negative'). They found no correlation between the methodological quality of the trial and its outcome. They concluded that the 'efficacy of epidural steroid injections has not been established'.

We considered whether the papers exclusive to one review determined the conclusion. Only one paper was exclusive to the Watts review.[11] The recalculated OR after removing this study was 2.41. Thus, the presence or absence of the paper does not significantly alter the conclusions that flow from the method chosen by Watts. The inclusion by Koes of two trials[13,14] is surprising. Rocco and his colleagues[13] studied patients with post-laminectomy pain; a group that is unrepresentative of the population of back pain sufferers. Serrao and colleagues[14] did not study steroid versus a placebo but steroid versus another active, though unproven, treatment (intrathecal midazolam). It is unclear if these two papers ought to be included in the systematic review at all, but it is clear that their presence or absence makes little difference to the conclusion. Removing these two papers leaves 10 papers, four reporting a 'negative' outcome and six reporting a 'positive' one. Of the five papers with quality scores above 50. three had 'positive' outcomes and two 'negative'. Once again, there was no association between methodological score and outcome. Following the logic of the method of this review, the conclusion would be unchanged.

Conclusions

Our starting point for this paper was that the conclusions of the two reviews differ in a way that might affect decisions about care. If the authors of two well-conducted reviews disagree about their findings, what should decision-makers do? We have shown that readers can check for themselves why this has occurred. In this case, it does not appear that slight differences in selection of trials caused the discordance. We also found that the judgement of which trials to include in the reviews differed, but would not have affected their overall conclusions. We conclude that the reason for the discordance was the difference in method for pooling the results of trials.

The nature of the discordance is one of degree: the Watts study found evidence for future practice, while the Koes study does not find sufficient evidence for a recommendation for practice. Watts and Silagy set clear rules to exclude trials with unacceptable bias, and to pool statistical results of the remainder. A statistically significant pooled typical odds ratio, then, based on the least biased available evidence, implies that there is an effect of treatment. The Koes review also used a formal procedure which met criteria for an acceptable systematic review, but the method was both less conducive to reaching a simple decision and more open to illustrate variations in quality other than bias in the trials. One important difference between the two reviews is in the way that studies which show no statistically significant difference in outcomes are treated. In the Koes review, these studies weigh the evidence against accepting the intervention, while in the Watts review, studies which show a non-significant trend nevertheless contribute to the estimate of the overall odds that the treatment is effective. The clear statistical result of the meta-analysis by

Watts and Silagy, on the one hand, makes more use of the information from the trials than does Koes's review, but it also has a seductive 'bottom line' quality for decision-makers, as is clear from the report in *Bandolier*. Watts and Silagy did not themselves make recommendations for practice but reported a significant result. This is in keeping with the objectives of reviewers within the Cochrane Collaboration, for example, who aim to provide information for decision-makers but not decisions.[10]

Watts and Silagy should have been less willing to do pooled meta-analysis on the basis of the studies that are of questionable comparability and quality. But in defence, excluding more studies increases the risk of finding a positive effect that does not exist. Including more (homogeneous) studies in different sorts of patient groups provides more widely generalizable findings.

However, if there are shortcomings to the pooling of data, then there are also problems with assessing net effect on the basis of a 'head count' of 'positive' and 'negative' studies. The very use of the term 'negative' for a study which does not demonstrate a significant result at an arbitrary level, denies the potential value of the evidence gathered in that trial. Even if the result is not significant at the 95% level, there may be a significant trend in one direction rather than another. So-called positive studies can be overemphasized in 'head count' approaches to reviewing. Small studies have a higher risk of giving a statistically significant result by chance. Pooling of evidence irons out both of these effects. Neither review assessed the possibility of publication and other selection bias in identifying studies. A simple graphical test, the 'funnel plot' of effect size and sample size to look for predicted funnel-shaped convergence, would have helped the reader.[25]

Our study shows that reviews should not be seen as a replacement for the judgement of decision makers, be they clinicians or commissioners. Decision-makers should read a review as critically as they would a trial. In particular, they should consider the assumptions and shortcomings of the review as we have demonstrated here. We suggest that they should seek all available reviews of a topic: This is easier now that the Cochrane Library[10] is widely available, including not only reports of systematic reviews of controlled trials (CDSR), but also the Database of Abstracts of Reviews of Effectiveness (DARE) and the Cochrane Controlled Trials Register. Our comparison is clearly not the last word on the effectiveness of epidural corticosteroid injections for sciatica. The case for an updated review is clear in the light of new evidence.[26] A new review protocol[27] included in the Cochrane Library will look at different types of injection therapy for chronic back pain, but it is not clear from the protocol whether or how the reviewers will consider sciatica as an indication for treatment.

References

1. Thompson SG, Pocock SJ. Can meta-analyses be trusted? *Lancet* 1991; **338:** 1127.
2. Oxman A, Cook D, Guyatt G, for the Evidence-Based Medicine Working Group. User's guides to the medical literature: how to use an overview. *JAMA* 1994; **272:** 1367-1371.
3. Petticrew M, Kennedy SC. Detecting the effects of thrombosis: the case of the rogue reviews. *BMJ* 1997; **315:** 665-668.
4. Anon. Meta-analysis under scrutiny. [Editorial.] *Lancet* 1997; **350:** 675.
5. Naylor DC. Meta-analysis and the meta-epidemiology of clinical research. *BMJ* 1997, **315:** 617-619.
6. Anon. Putting it into your back! [Editorial.] *Bandolier* 1996; **3:** 3-4.
7. Watts RW, Silagy CA. A meta-analysis on the efficacy of epidural corticosteroids in the treatment of sciatica. *Anaesth Intensive Care* 1995; **23:** 564-569.
8. McQuay HJ, Moore A. Epidural steroids for sciatica. *Anaesth Intensive Care* 1996; **24:** 284-285.
9. Koes BW, Scholten RJPM, Mens JMA, Bouter LM. Efficacy of epidural steroid injections for low back pain and sciatica: a systematic review of randomised clinical trials. *Pain* 1995; **63:** 279-288.
10. The Cochrane Collaboration. *The Cochrane Library CD ROM 1998*. Issue 2. Oxford: Update Software Ltd, 1998.
11. Popiolek A, Domanik A, Mazurkiewicz G. Leczenie sterydowymi ostrzknieciami nodopow owymi chorychz prewlekla rwa kulsowa w przebiegu dyskopatti. *Neur Neurochir Pol* 1991; **T25 NR5:** 640-646.
12. Czarski Z. Leczenie rwy kulsowej wstrzykin aniem hydrokotyzonu I nowokainy do rozworn krzuzuwego. *Przeglad Lekarski* 1965; **21:** 511-515.
13. Rocco AG, Frank E, Kaul AF, *et al*. Epidural steroids, epidural morphine and epidural steroids combined with morphine in the treatment of post laminectomy syndrome. *Pain* 1989; **36:** 297-303.
14. Serrao JM, Marks RL, Morley SJ, Goodchild CS. Intrathecal midazolam for the treatment of chronic mechanical low back pain: a comparison with epidural steroid in a pilot study. *Pain* 1992; **48: -**5-12.
15. Yates DW. A comparison of the types of epidural injection commonly used in the treatment of low back pain and sciatica. *Rheum Rehab* 1978; **17:** 181-186.
16. Cuckler J, Bernini P, Weisel SW, *et al*. The use of steroids in the treatment of lumbar radicular pain. *J Bone Joint Surg* 1985; **67:** 63-66.
17. Breivik H, Helsa PE, Molnar I, Lind B. Treatment of chronic low back pain and sciatica: comparison of caudal epidural injections of bupivacaine and methylprednisolone with bupivacaine followed by saline. *Adv Pain Res Ther* 1976; **1:** 927-932.
18. Mathews JA, Mills SB, Jenkins VM, *et al*. Back pain sciatica: controlled trials of manipulation, traction, sclerosant and epidural injections. *Br J Rheumatol* 1987; **26:** 416-4**4**23.
19. Snoek W, Weber H, Jorgenson B. Double blind evaluation of extradural methyl prednisolone for herniated lumbar discs. *Acta Orthop Scand* 1977; **48:** 635-641.
20. Bush K, Hillier S. A controlled study of caudal epidural injections of triamcinolone plus procaine for the management of intractable sciatica. *Spine* 1991; **16:** 572-575.
21. Dilke TFD, Burry HC, Grahame R. Extradural corticosteroid injection in management of lumbar nerve root compression. *BMJ* 1973; **2:** 635-637.
22. Klenerman L, Greenwood R, Davenport HT, *et al.* Lumbar epidural injections in the treatment of sciatica. *Br J Rheumatol* 1984; **23:** 35-38.
23. Melzck R. The McGill Pain questionnaire: major properties and scoring methods. *Pain* 1975; **1:** 277-299.
24. Huskisson EC. Measurement of pain. *Lancet* 1974; **ii**: 1127-1131.
25. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634.
26. Carette S, Leclaire R, Marcoux S, *et al*. Epidural corticosteroid injections for sciatica due to herniated nucleus pulposus. *New Engl J Med* 1997; **336:** 1634-1650.
27. de Bie RNP, de Vet HSH. Injection therapy for patients with chronic benign low back pain. [Protocol.] *The Cochrane Library 1998*. Issue 2. Oxford: Update Software Ltd, 1998.
28. Ridley MG, Kingsley GH, Gibson T, Grahame K. Outpatient lumbar epidural corticosteroid injection in the management of sciatica. *Br J Rheumatol* 1988; **27:** 295-299.
29. Beliveau P. A comparison between epidural anaesthesia with and without corticosteroid in the treatment of sciatica. *Rheum Phys Med* 1971; **11:** 40-43.

Acknowledgements

We are grateful to colleagues and anonymous referees for comments on earlier drafts. This research was not funded from external sources, and we acknowledge support from the University of East Anglia.

Address for correspondence

Dr Miranda Mugford, Senior Lecturer in Health Services Research, School of Health, University of East Anglia, Norwich NR4 7TJ.