

Entertaining Patients and Dazzling Reviewers



— *The function of a doctor is to entertain while you improve* [trad.]

Archie Cochrane (1909-1988) was best known for the quality of his fieldwork, building on strong and personal empathy with the people in the Welsh valleys. The studies conducted by his team, and with colleagues such as Kellgren and Lawrence, were observational studies. It seems that he never published a randomized clinical trial (RCT).

But in 1979, he published an essay in which he stated firmly, “It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomised controlled trials”¹. Following a long gestation, this notion led to international and interdisciplinary “Cochrane Collaborations” in 1993, to the registration of a Musculoskeletal Group, and to a first meeting of the Cochrane Collaboration Steering Group, at McMaster University, Hamilton, Canada, in 1994.

These developments were quickly followed by a *British Medical Journal* publication in 1996, by Aker, Gross, Goldsmith, and Peloso, titled “Conservative management of mechanical neck pain: systematic overview and meta-analysis”². To quote from the report:

In a survey about attitudes on treatment of musculoskeletal disease, active exercise, traction, TENS [transcutaneous electrical nerve stimulation], and ultrasound were perceived to be the best methods for the treatment of neck pain. The results of this overview clearly do not support these approaches nor others commonly used in practice today. What becomes most clear from this overview is the lack of evidence for many of the standard approaches to neck pain used in health care today. Even for the treatments found in this overview to have some early evidence of support, such as manual treatments in combination with other treatments, conclusions must be

made cautiously because of the small number of trials on which they are based. In general, no treatments have been studied in enough detail to adequately assess either efficacy or effectiveness. When we consider the varied treatment approaches to neck pain, with their potential risks and costs, much further work is needed to determine optimal treatment approaches.

These authors and others formed a Cervical Overview Group, which has diligently re-reviewed this field, “adapted periodically” as urged by Cochrane. In 1994, Gross, *et al* published “A Cochrane Review of Manipulation and Mobilization for Mechanical Neck Disorders” in *Spine*³.

The current Cochrane Collaboration review⁴ updates and summarizes evidence about the many treatment strategies covered in more narrowly focused reviews over the past decade. From 88 selected RCT, 52 of “acceptable quality,” Gross and colleagues conclude, “Exercise combined with mobilization/manipulation, exercise alone, and lidocaine (IM) for chronic MND [mechanical neck disorders]; glucocorticoid (IV) for acute WAD [whiplash associated disorders]; have demonstrated either intermediate or longterm benefits.” But in a very similar Cochrane Collaboration update, Verhagen and colleagues concluded⁵: “We can draw no conclusion about the most effective therapy for patients with chronic WAD... All but one study mentioned positive results, but the actual data of the high quality studies were conflicting.” Both groups targeted conservative treatments for whiplash and mechanical neck disorders. We look back with respect, but before looking forward, we must review problems that have become apparent.

Selection instruments such as the Jadad, the van Tulder, and the more recent CONSORT criteria concentrate on editorial content and structure; they neglect quality issues identified as crucial as early as 1994, but which are increasingly apparent. Some of these include: the need for factorial design, related to which is statistical fitting for multiple

See Conservative management of mechanical neck disorders, page 1083

responses; adequate sample size and power; and adequate modeling for time. Most clinics studying neck (and back) problems use multiple treatment strategies. Most of their patients improve, and patients, therapists, and researchers have been content to attribute improvement to the therapy given. The struggle has been to decide which therapy is best. Perhaps none of the above?

The need for factorial design was recognized in 1994 (a need largely ignored in the studies reviewed), and the need to model the effects of time became more obvious in 1995, with the publications of the Quebec Task Force On Whiplash-Associated Disorders⁶ and the report by Radanov, *et al*⁷ on the rate of recovery after acute neck injury. In both studies, there was very rapid recovery in the first weeks: over half were asymptomatic by 3 months. But of those still affected at 6 months, well over half were still symptomatic at 1 year. In the study by Helewa, *et al*⁸, the variance explained by time was more than double the largest intergroup treatment effect.

When the effects of time are not considered during the design phase, sample sizes will be too small. If they are not considered during the analysis, before/after differences will wrongly be attributed to the treatments given, and results wrongly considered statistically significant and clinically meaningful. In the report by Gross, *et al* in this issue of *The Journal* these concerns apply to most studies listed as providing “strong evidence” of benefit (as in their Table 4)⁴.

The control group is a source of concern in many of the studies. At baseline, the patients assigned to the control group should be identical in every way to those in the treatment groups, except for random variation. During the “treatment” period, they should be eager participants, believing they are valued and are receiving the best possible care. They need to be “entertained” or they will seek treatment elsewhere. If they become nonrandom dropouts, there is no satisfactory statistical remedy. Design and recruitment for such studies can present difficult problems, so that comparison groups created may be fundamentally different from the treatment groups. After the fact, it may be difficult for a reviewer to identify that there is a problem. Of course, anomalous control groups can arise purely by chance. From 88 studies, differences at the 0.05 level could be expected to occur by chance in about 4 or 5 studies: we do not have to infer evil intent.

Let us return to the 4 studies listed in Gross’s Table 4⁴ dealing with therapies involving stretching and strengthening exercises, mobilization (low velocity), and manipulation (high velocity thrusts), and considered to yield evidence judged as “strong.” The lead authors were Allison⁹, Brodin¹⁰, Karlberg¹¹, and Jull¹².

The Allison study involved only 30 subjects in 3 groups, and was considered by the authors as “a pilot study.” At baseline and 8 weeks there were 2 response measures, the Northwick Park Neck Pain Questionnaire (NPQ) and a pain

visual analog scale (VAS). NPQ differences were not significant, but VAS response in the control group was “significantly” less than in the other 2. This was due to differences at baseline, when the VAS was only 3.3 for the control group, and 4.8 and 5.1 for the other 2. All statistical calculations were within subgroup endpoint minus baseline, ignoring the effect of time.

The Brodin study was larger, involving 63 participants in 3 groups, observed over 1 week. There were 2 control groups. One received only oral analgesics, the other “mock therapy” in a multimodal treatment center, a combination “package” of superficial massage, manual traction, heat, and analgesics. The test group received the same package, plus cervical mobilization. Outcome was pain on a 9-point scale. The 2 groups treated within the center did not differ significantly, but did better than the excluded group, but statistical power was calculated at 18%. In the 1996 review, this result was listed with “none significant.”

In the Karlberg study, the focus was on postural performance in patients with dizziness of suspected cervical origin. Patients were randomized either to receive immediate physiotherapy (n = 9), or to wait 2 months (n = 8). Need I go on?

Jull and colleagues were interested in the effectiveness of physiotherapy in reducing the frequency of headaches of cervical origin. Two hundred participants who met diagnostic criteria for cervicogenic headache were randomized into 4 factorial groups: manipulative therapy, exercise therapy, combined therapy, and a control group. The treatment period was 6 weeks with followup assessment after treatment, then at 3, 6, and 12 months. In addition to headache frequency, outcome measures included the Northwick Park Neck Pain Index, medication intake, and patient satisfaction. At the 12-month followup assessment, all center treated groups improved equally, when calculated as endpoint minus baseline, ignoring the effect of time.

There is one study of interest that was not included in their list of studies of acceptable quality, despite an ambitious, balanced 2*2*2 factorial design, with 336 patients randomly allocated to one of the following groups: manipulation with and without heat, manipulation with and without electrical muscle stimulation (EMS), mobilization with and without heat, or mobilization with and without EMS. In essence, there were no differences among the treatment groups¹³, and all differences could be attributable to time. There was no control group. The study was of interest, because at least 1 adverse event (pain and headaches) was described by 30% of respondents, more commonly after manipulation than low velocity mobilization. Subjects reporting adverse reactions were less satisfied with care and less likely to have clinically meaningful improvements in pain and disability¹⁴.

A single study of the use of high-dose intravenous methylprednisolone for whiplash patients was included in the review by Gross and colleagues. There were 20 patients

and 20 placebo-treated controls. Symptom benefit after 6 months for the treated group was claimed¹⁵. However, a high incidence of adverse reactions was reported in a study of similar size and design, not referenced in the current review. There were 8 instances of pulmonary complication with steroid (34.8%) and one instance (4.34%) with placebo ($p = 0.009$). There were 4 instances of gastrointestinal complication (17.4%) with steroid and none with placebo ($p = 0.036$)¹⁶.

Now let us return to Quebec, and to a new study recently published in *The Journal*¹⁷. Recall the rapid return to work documented in the 1995 report: “among those whose pain persisted longer than 7 days, about 35% were still receiving compensation after 13 weeks.” Now “data from 2000–2001 show that the proportion of patients still receiving compensation at 13 weeks [was] 60% in the experimental program. This apparently important increase in time on compensation between 1987 and 2001 needs to be investigated.” Multimodal therapy is expensive, in both direct costs and the costs associated with disability.

EVIDENCE FROM OBSERVATIONAL STUDIES — CONTROLLED OR OTHERWISE

Evidence-based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. In the Cochrane Reviews, it is common to consider only evidence derived from RCT.

“There’s a paradox here. Ask most research physicians how a profession can advance, and they will talk about the model of ‘evidence-based medicine’ — the idea that nothing ought to be introduced into practice unless it has been properly tested and proved effective by research centers, preferably through a double-blind RCT. But, in a 1978 ranking of medical specialties according to their use of hard evidence from randomized clinical trials, obstetrics came in last. Obstetricians did few randomized trials, and when they did they ignored the results... Doctors in other fields have always looked down their masked noses on their obstetrical colleagues. Obstetricians used to have trouble attracting the top medical students to their specialty, and there seemed little science or sophistication to what they did. Yet almost nothing else in medicine has saved lives on the scale that obstetrics has.”¹⁸ We would add the definition of the health problems related to smoking, or the lung diseases in Welsh coal miners described by Cochrane, all based on high quality observational studies.

Is multimodal care as helpful for neck and back problems as is assumed by the authors and the reviewers in this field? In 1990, the Workers’ Compensation Board of Ontario sponsored an early, active, exercise and education program available in more than 100 clinics. Health and cost outcomes were compared with “usual” care. Duration of benefits showed no significant difference between attenders and

non-attenders, and healthcare costs for clinic attenders were significantly higher. Functional status, health-related quality of life, and pain measures all improved significantly throughout time for both groups, but there were no statistically significant differences in rate of improvement¹⁹.

CONCLUSION

I would return to the conclusions stated in the 1996 review by this group²:

What becomes most clear from this overview is the lack of evidence for many of the standard approaches to neck pain used in health care today. Even for the treatments found in this overview to have some early evidence of support, such as manual treatments in combination with other treatments, conclusions must be made cautiously because of the small number of trials on which they are based... In general, no treatments have been studied in enough detail to assess either efficacy or effectiveness adequately. When we consider the varied treatment approaches to neck pain, with their potential risks and costs, much further work is needed to determine optimal treatment approaches.

HUGH A. SMYTHE, MD, FRCPC,
Associate Editor,
The Journal of Rheumatology

Address reprint requests to Dr. H.A. Smythe, University of Toronto,
2 Heathbridge Park, Toronto, Ontario M4G 2Y6, Canada.
E-mail: hasmythe@rogers.com.

REFERENCES

1. Cochrane AL. 1931-1971: a critical review, with particular reference to the medical profession. In: *Medicines for the year 2000*. London: Office of Health Economics;1979:1-11.
2. Aker PD, Gross AR, Goldsmith CH, Peloso P. Conservative management of mechanical neck pain: systematic overview and meta-analysis. *BMJ* 1996;313:1291-6.
3. Gross AR, Hoving JL, Haines TA, et al. A Cochrane Review of manipulation and mobilization for mechanical neck disorders. *Spine* 2004;29:1541-8.
4. Gross AR, Goldsmith CH, Hoving JL, et al. Conservative management for mechanical neck disorders: A systematic review. *J Rheumatol* 2007;34:1083-102.
5. Verhagen AP, Scholten-Peeters GG, de Bie RA, Bierma-Zeinstra SM. Conservative treatments for whiplash. *Cochrane Database Syst Rev* 2004;1:CD003338.
6. Spitzer WO, Skovron ML, Salmi LR, et al. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: redefining “whiplash” and its management. Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec, Canada. *Spine* 1995;20:1S-73S.
7. Radanov BP, Sturzenegger M, Di Stefano G. Long-term outcome after whiplash injury. A 2-year follow-up considering features of injury mechanism and somatic, radiologic, and psychosocial findings. *Medicine* 1995;74:281-97.
8. Helewa A, Goldsmith CH, Smythe HA, Lee P, Obricht K, Stitt L. Effect of therapeutic exercise and sleeping neck support on patients with chronic neck pain: a randomized clinical trial. *J Rheumatol* 2007;34:151-8. Epub 2006 Nov 1.
9. Allison GT, Nagy BM, Hall T. A randomized clinical trial of

- manual therapy for cervico-brachial pain syndrome — a pilot study. *Man Ther* 2002;7:95-102.
10. Brodin H. Cervical pain and mobilization. *Int J Rehabil Res* 1984;7:190-1.
 11. Karlberg M, Magnusson M, Malmstrom EM, Melander A, Moritz U. Postural and symptomatic improvement after physiotherapy in patients with dizziness of suspected cervical origin. *Arch Phys Med Rehabil* 1996;77:874-82.
 12. Jull G, Trott P, Potter H, et al. A randomized controlled trial of exercise and manipulative therapy for cervicogenic headache. *Spine* 2002;27:1835-43.
 13. Hurwitz EL, Morgenstern H, Harber P, Kominski GF, Yu F, Adams AH. A randomized trial of chiropractic manipulation and mobilization for patients with neck pain: clinical outcomes from the UCLA neck-pain study. *Am J Public Health* 2002;92:1634-41.
 14. Hurwitz EL, Morgenstern H, Vassilaki M, Chiang LM. Adverse reactions to chiropractic treatment and their effects on satisfaction and clinical outcomes among patients enrolled in the UCLA Neck Pain Study. *J Manipulative Physiol Ther* 2004;27:16-25.
 15. Pettersson K, Toolanen G. High-dose methylprednisolone prevents extensive sick leave after whiplash injury. A prospective, randomized, double-blind study. *Spine* 1998;23:984-9.
 16. Matsumoto T, Tamaki T, Kawakami M, Yoshida M, Ando M, Yamada H. Early complications of high-dose methylprednisolone sodium succinate treatment in the follow-up of acute cervical spinal cord injury. *Spine* 2001;26:426-30.
 17. Suissa S, Giroux M, Gervais M, et al. Assessing a whiplash management model: a population-based non-randomized intervention study. *J Rheumatol* 2006;33:581-7.
 18. Gawande A. The score. *The New Yorker*; October 9, 2006.
 19. Sinclair SJ, Hogg-Johnson SA, Mondloch MV, Shields SA. The effectiveness of an early intervention program for workers with soft tissue injuries: the Early Claimant Cohort Study. *Spine* 1997;22:2919-31.