

Conservative vs. Surgical Care of Lumbar Spinal Stenosis by Dr. Arthur Croft

Abstract

Study Design: Randomized trial and concurrent observational cohort study.

Objective: To compare four-year outcomes of surgery to nonoperative care for spinal stenosis.

Summary of Background Data: Surgery for [spinal stenosis](#) has been shown to be more effective compared to nonoperative treatment over two years, but longer-term data have not been analyzed.

Methods: Surgical candidates from 13 centers in 11 U.S. states with at least 12 weeks of symptoms and confirmatory imaging were enrolled in a randomized cohort (RC) or observational cohort (OC). Treatment was standard decompressive laminectomy or standard nonoperative care. Primary outcomes were SF-36 bodily pain (BP) and physical function scales and the modified Oswestry Disability index assessed at six weeks, three months, six months, and annually up to four years.

Results: A total of 289 patients enrolled in the RC and 365 patients enrolled in the OC. An as-treated analysis combining the RC and OC and adjusting for potential confounders found that the clinically significant advantages for surgery previously reported were maintained through four years, with treatment effects (defined as mean change in surgery group minus mean change in nonoperative group) for bodily pain 12.6 (95% confidence interval [CI], 8.5-16.7); physical function 8.6 (95% CI, 4.6-12.6); and Oswestry Disability index -9.4 (95% CI, -12.6 to -6.2). Early advantages for surgical treatment for secondary measures such as bothersomeness, satisfaction with symptoms, and self-rated progress were also maintained.



Conclusion: Patients with symptomatic spinal stenosis treated surgically compared to those treated nonoperatively maintain substantially greater improvement in pain and function through four years.

Critique

The study reports the results of the Spine Patient Outcomes Research Trial (SPORT), which included a randomized cohort (RC) and a concurrent observational cohort (OC) of patients who declined randomization. The study covered 11 states and 13 medical centers. The lead author of the above study,¹ James Weinstein, is also the editor of *Spine*. (When we were struggling with getting our paper – a critique of the Quebec Task Force on Whiplash-Associated Disorders – published in the late 1990s, by coincidence Jim and I were sitting next to each other at an NIH Grant Proposal committee meeting and I bent his ear – apparently successfully. *Spine* ultimately published our paper.)²

Selection and enrollment was not described in any detail in the Weinstein, et al., paper. Important questions would be how the subjects were recruited, specifically whether they would have been randomly selected from a group of eligible individuals versus being selectively recruited by participating (i.e., cherry-picking) surgeons on a case-by-case basis. In any event, all had either radiculopathy and/or neurogenic claudication, and were deemed surgical candidates. Persons with degenerative [spondylolisthesis](#) and instability were excluded.

Note that in a study by Deyo, et al.,³ 18.4 percent of the lumbar surgical cases had this condition, so it is certainly not a contraindication for surgery. Its exclusion here might be viewed as a bit of cherry-picking. The proportions of prior management were 68 percent PT, 56 percent epidural injections, 28 percent chiropractic and 27 percent opioid analgesics. One might be tempted to kibitz that the lower numbers in the latter two groups might indicate that they are relatively more successful than the others, although there is no indication that is the case with opioids.

The numbers broke down as follows: randomized = 289 (138 assigned to surgery; 151 assigned to nonsurgical status); observational = 365 (219 chose surgery; 146 chose nonsurgical means). In the end, 413 had surgery and 221 did not have surgery. The surgical procedure was a posterior laminectomy.

Nonsurgical protocol included PT, education/counseling, home exercise, and NSAIDs, but not chiropractic treatment. End points were the SF-36 Bodily Pain (BP) and Physical Function (PF) scales, and the AAOS/Modems version of the Oswestry Disability Index (ODI). Of patients in these groups, about a third were lost to follow-up.

Although the RC was initially analyzed on an intention-to-treat basis, they were subsequently analyzed on the basis of any cross-over that they made. This approach, which is called on-protocol analysis, creates a bias in favor of surgery. A more conservative approach is to conduct the analysis as the no group change had taken place. This method is referred to as intention to treat, and it makes it more difficult to demonstrate a positive treatment effect.

The best approach is to run the analysis using both methods. If the results are very divergent, it implies that the conclusions may be suspect.

In a strict randomized design, the surgical group would all have surgery and the nonsurgical group would not, but ethical prerequisites allow enrollees to cross-over if they wish. After four years, only 68 percent of the RC had undergone surgery, compared to 49 percent of the OG.

Complications were broken down as to the type of complication. Four deaths were reported as being the result of complications from surgery. Intraoperative complications were reported to be 10 percent and postoperative ones 13 percent. The four-year re-operation figure was also 13 percent.

The authors reported that their results were "generally similar to those in previous surgical series." There are other factors to consider, however, and not all studies have shown surgical approaches to be most favorable. Fundamentally, an overriding question that public health officials and health administrators share is the comprehensive cost to society and the individual. How many days of productive work were lost during the four years of the study? What was the cost to the employer and to the individual? What were the costs of medication, orthotics, and other health aids? Were there any untoward side effects from taking the medication, such as GI bleeds, narcotic dependency problems, etc.? What were the costs of these episodes? What proportion was working in their usual job description full time at follow-up, as opposed to part time or in a changed job description; or not working at all? What were the total costs of surgery, including the second, third and in some cases, fourth surgeries? What were the long-term consequences of the complications? How do the total [medical costs](#) compare in both groups?

Further questions that seem appropriate would be, why was a smaller percentage of presurgical care chiropractic? More fundamentally, why did the nonsurgical group not have chiropractic treatment as one of the alternative treatments? After all, if only 28 percent had previously seen a chiropractor, perhaps as much as 72 percent of these subjects could have potentially benefited.

The Takeaway

Upon closer examination, this study is another example of stacking the deck for success, even forgetting the potential of selective recruitment of patients likely to have a more favorable surgical outcome (and I have no reason to suspect that such a thing occurred other than a suspicious mind, more than half a century of personal experience with human nature, many years in the research business, and the fact that the authors did not make that process terribly clear).

A degree of selection bias exists merely by dint of the fact that these people were recruited from a spine specialty centers. Individuals whose spinal stenosis is manageable without surgery are less likely to end up in such a center in the first place. And, the "nonsurgical" groups had all been previously deemed "surgical candidates." So, the study did not honestly compare surgery to no-surgery in a representative cohort of individuals with spinal stenosis. It compared surgery and no-surgery among a surgical subgroup of patients with spinal stenosis.

Moreover, rather than comparing surgery to a viable competing management approach, they essentially compared surgery to "no treatment." Why do I say that? Because the nonsurgical alternatives had already been shown to be ineffective in this group, so the probability of their being effective after randomization or self-allocation to the OG was marginal at best. This is a variant of Gail's fallacy.⁴

Chiropractic management would have been a more SPORTing contest, of course, but that may have been deemed too risky. It might have been seen to be much less expensive, with far fewer and much less serious side effects/complications, and no downtime for recovery.

And that leads us to the final question: why only four years of follow-up? Some of the people who had undergone multiple surgeries were probably still recovering from the previous one. The biggest question is how these people fare 10 years down the road and beyond. I hope the authors attempt to get some follow-up in time.

References

1. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonoperative treatment for lumbar spinal stenosis: four-year results of the Spine Patient Outcomes Research Trial. *Spine* 2010;35(14):1329-38.
2. Freeman MD, Croft AC, Rossignol AM. Whiplash-associated disorders (WAD) - redefining whiplash and its management" by the Quebec Task Force: a critical evaluation. *Spine*, 1998;23(9):1043-1049.
3. Deyo RA, Mirza SK, Martin BI, et al. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. *JAMA*, 2010;303(13):1259-1265.
4. Gail MH. Does cardiac transplantation prolong life? *Ann Intern Med*, 1972;76:815.