ELSEVIER

CrossMark

The Spine Journal 16 (2016) 579–587

THE SPINE JOURNAL

Clinical Study

The long-term outcome of lumbar fusion in the Swedish lumbar spine study

Rune Hedlund, MD, PhD[a,*], Christer Johansson, MSc[a], Olle Hägg, MD, PhD[b], Peter Fritzell, MD, PhD[c], Tycho Tullberg, MD, PhD[d], Swedish Lumbar Spine Study Group

[a]*Department of Orthopaedics, Salhgrenska University Hospital, Bruna stråket 11, Gothenburg, SE 413 45, Sweden*
[b]*Göteborg Spine Center, Gruvgatan 8, Västra Frölunda, SE 421 30, Sweden*
[c]*Department of Orthopedics, Länssjukhuset, Ryhov, SE 551 85 Ryhov, Sweden*
[d]*Stockholm Spine Center AB, Löwenströmska Sjukhuset, Upplands Väsby, SE 194 89, Sweden*

Received 19 February 2015; revised 30 July 2015; accepted 27 August 2015

Abstract

BACKGROUND CONTEXT: Current literature suggests that in the long-term, fusion of the lumbar spine in chronic low back pain (CLBP) does not result in an outcome clearly better than structured conservative treatment modes.

PURPOSE: This study aimed to assess the long-term outcome of lumbar fusion in CLBP, and also to assess methodological problems in long-term randomized controlled trials (RCTs).

STUDY DESIGN: A prospective randomized study was carried out.

PATIENT SAMPLE: A total of 294 patients (144 women and 150 men) with CLBP of at least 2 years' duration were randomized to lumbar fusion or non-specific physiotherapy. The mean follow-up time was 12.8 years (range 9–22). The follow up rate was 85%; exclusion of deceased patients resulted in a follow-up rate of 92%.

OUTCOME MEASURES: Global Assessment (GA) of back pain, Oswestry Disability Index (ODI), visual analogue scale (VAS) for back and leg pain, Zung depression scale were determined. Work status, pain medication, and pain frequency were also documented.

METHODS: Standardized outcome questionnaires were obtained before treatment and at long-term follow-up. To optimize control for group changers, four models of data analysis were used according to (1) intention to treat (ITT), (2) "as treated" (AT), (3) per protocol (PP), and (4) if the conservative group automatically classify group changers as unchanged or worse in GA (GCAC). The initial study was sponsored by Acromed (US$50,000–US$100,000).

RESULTS: Except for the ITT model, the GA, the primary outcome measure, was significantly better for fusion. The proportion of patients much better or better in the fusion group was 66%, 65%, and 65% in the AT, PP, and GCAC models, respectively. In the conservative group, the same proportions were 31%, 37%, and 22%, respectively. However, the ODI, VAS back pain, work status, pain medication, and pain frequency were similar between the two groups.

CONCLUSIONS: One can conclude that from the patient's perspective, reflected by the GA, lumbar fusion surgery is a valid treatment option in CLBP. On the other hand, secondary outcome measures such as ODI and work status, best analyzed by the PP model, indicated that substantial disability remained at long-term after fusion as well as after conservative treatment. The lack of objective outcome measures in CLBP and the cross-over problem transforms an RCT to an observational study,

FDA device/drug status: Not applicable.

Author disclosures: ***RH:*** Grants: Acromed Corporation (D), Zimmer (B), outside the submitted work; Consulting: K2M (A), Globus Medical (B), Medtronic (B), Zimmer (B), outside the submitted work. ***CJ:*** Nothing to disclose. ***OH:*** Other: Spine Center Göteborg (A), outside the submitted work. ***PF:*** Nothing to disclose. ***TT:*** Dr Tullberg reports being CEO of Stockholm Spine Center (SSC), a private spine clinic, and Stockholder in SSC and in Global Health Partner, which is the main owner of SSC.

* Corresponding author. Department of Orthopaedics, Salhgrenska University Hospital, Bruna stråket 11, Gothenburg, SE 413 45, Sweden. Tel.: +46313434060.

E-mail address: rune.hedlund@vgregion.se (R. Hedlund)

http://dx.doi.org/10.1016/j.spinee.2015.08.065

that is, Level 2 evidence. The discrepancy between the primary and secondary outcome measures prevents a strong conclusion on whether to recommend fusion in non-specific low back pain.

Keywords: Chronic low back pain; Conservative treatment; Long-term outcome; Lumbar fusion; Physical therapy; Randomized trial

Introduction

Despite an abundance of clinical studies, the outcome of fusion for chronic low back pain (CLBP) remains a highly controversial subject in spine surgery. In the first randomized controlled trial (RCT) comparing lumbar fusion with conservative therapy, The Swedish Lumbar Spine Study Group reported a positive short-term effect of surgery compared with unstructured physiotherapy [1]. In contrast, in British and Norwegian short-term as well as long-term studies, no statistically or clinically relevant difference could be demonstrated comparing fusion with physiotherapy and cognitive therapy [2–5].

In an evidence based medicine perspective, well-designed and executed RCTs are considered Level 1 studies on which recommendations on treatment can be based. However, an important limitation of long-term RCTs that compare surgical with conservative treatment is crossover between treatments, which undermines the randomization process. A further problem in long-term studies is follow-up rate; there is risk of selection bias with suboptimal number of patients available for follow-up. In a recently combined British-Norwegian study on CLBP [5], the follow-up rate was 55%, which is generally considered too low for robust conclusions.

The present study was performed to determine the long-term outcome of fusion for CLBP treated with fusion or an unstructured physiotherapy program. The data presented are the long-term RCT follow-up of the Swedish Lumbar Spine Study [1]. A further objective was to analyze the methodological problem associated with crossover patients in long-term RCTs.

Materials and methods

Consecutively referred patients with CLBP aged 25–65 years [1] were eligible to participate in the study (Table 1). The inclusion criteria were patients aged 25–65 years, male and female, with severe CLBP of at least 2 years' duration, and with no signs of nerve root compression. Further inclusion criteria were sick leave or "equivalent" major disability for at least 1 year and unsuccessful medical interventional treatment efforts. Radiological inclusion criteria were degenerative changes at L4–L5 or L5–S1 ("spondylosis") on plain radiographs or computed tomography, or magnetic resonance imaging.

Exclusion criteria were previous spine surgery except for successful removal of a herniated disc, spondylolysis, spondylolisthesis, new or old fractures, infection, inflammatory process, or neoplasm. No patient was diagnosed with a spondylolysis intraoperatively.

Table 1
Baseline demographic and clinical characteristics

	Surgical group (n=222)	Medical interventional group (n=72)
Age (range)	43 (25–64)	44 (26–63)
Sex (female)	112 (50.5%)	37 (51.4%)
Smoking	40.6%	49.3%
Comorbidity	39.1%	23.5%
Mean pain duration, years (range)	7.8 (2–34)	8.5 (2–40)
Mean time of sick leave, years (range)	3.2 (0.1–18)	2.9 (0.1–8)
Working part or full time	20.9%	23.6%
ODI (0–100)	47.3 (11.4)	48.4 (11.9)
VAS back pain (0–100)	64.2 (14.3)	62.6 (14.3)
VAS leg pain (0–100)	35.3 (25.4)	35.6 (25.2)
Zung depression scale (20–80)	39.1 (13.3)	39.4 (13.9)

ODI, Oswestry Disability Index; VAS, Visual Analogue Scale.
Data are means (standard deviation [SD]) or numbers (%) unless stated otherwise.

There were 294 patients, 144 women and 150 men, with a mean age of 47 years (range, 28–72). A 3:1 randomization between different types of fusion and physiotherapy was performed by a computer-generated random sequence, which resulted in 222 patients in the fusion group and 72 in the conservative group. The different types of fusion were (1) non-instrumented posterolateral fusion, (2) instrumented posterolateral fusion with internal fixation, (3) instrumented circumferential fusion with additional interbody bone graft, either as anterior lumbar interbody fusion or posterior lumbar interbody fusion. Only autografts were used. The different types of fusion resulted in a similar outcome at 2 years [6], as well as at long-term follow-up, as observed in the present study. Therefore, only the combined results of the fused patients are presented. Crossover, that is, change of treatment group post randomization, can be followed by the flowchart shown in Fig. 1.

The randomization resulted in a 40.6% smokers in the surgical group and 49.3% in the conservative group, introducing the risk of bias. The difference was accounted for by multiple linear regression analysis, with adjustment for age, gender, smoking, pretreatment pain duration, previous spine surgery, and Oswestry Disability Index (ODI) at baseline, all well-known risk factors for a worse outcome in spine surgery.

The primary outcome measure was Global Assessment (GA) in which the patient classified the outcome as "much better," "better," "unchanged," or "worse." Secondary outcome