Available online at www.sciencedirect.com
ScienceDirect
SAS Journal 4 (2010) 68–74
SAS Journal
www.sasjournal.com

Lumbar arthroplasty

Results from a randomized controlled study between total disc replacement and fusion compared with results from a spine register

Svante Berg, MD, PhD [a,*], Hans Tropp, MD, PhD [b]

[a] *Stockholm Spine Center, Löwenströmska Hospital, Stockholm, Sweden*
[b] *University Hospital, Linköping, Sweden*

Abstract

Background: Difficulties in performing randomized controlled trials (RCTs) to evaluate new treatment options are increasing. Higher costs and patient unwillingness are the main obstacles. A spinal surgery register has been in use in Sweden for 11 years. Our aim was to determine whether this register can provide the same information as an RCT and whether register data compare favorably with RCT data, making RCTs unnecessary. If not the case, was patient selection or follow-up frequency the cause of any differences?

Materials and methods: We compared baseline data and outcome, retrieved from our register, between 2 surgical groups, total disc replacement (TDR) and fusion at 1 or 2 levels, performed for degenerative disc disease. One hundred fifty-two patients were part of an RCT, whereas four hundred fifty-five patients had been treated according to an active decision. These 2 subgroups were the subjects for comparison.

Results: The 2 subgroups were not similar at baseline. Patients who were fused in the non-RCT subgroup were older, had a higher Oswestry Disability Index, and were more frequently smokers than the other patients. The outcome for the non-RCT group showed larger differences in favor of TDR than the RCT did. The nonresponders in the non-RCT group showed worse life quality and disability at baseline, and patients who answered the 1-year follow-up questionnaire but not the 2-year follow-up questionnaire had an inferior clinical result compared with the other patients at 1 year.

Conclusion: Data from our register showed results similar to the RCT, but a register cannot fully replace an RCT study when evaluating a new treatment option if the RCT has narrower selection than just the diagnosis. In this RCT comparing TDR with posterior fusion, the normal exclusion criteria for TDR were used. These were not registered, so the register could not prevent a possible selection bias, which might also be caused by the nonresponders.

Keywords: RCT; Non-RCT; Swespine; Selection biaz; TDR; Fusion

Significant resources are required to perform a randomized controlled trial (RCT). To achieve approval from an ethics committee, there must be a true lack of knowledge in the field that is to be studied, as is the case in chronic low-back pain (CLBP). The diagnosis of degenerative disc disease (DDD) and its surgical treatment are still often questioned, mainly because there are contradicting results from different studies, some of which report a positive result after surgery compared with nonspecific conservative treatment[1] and some of which show no difference between surgery and intensive specific rehabilitation including cognitive behavioral therapy.[2]

To reach the highest degree of evidence, corresponding results from repeated RCTs are mandatory.[3–5] It is important not only to identify a treatment that gives a beneficial result compared with other treatments but also to recognize what an important significant difference is.[6]

Knowledge of treatment options among the public seems to be expanding. It appears to be increasingly difficult to enroll patients to evaluate a new treatment option because many patients have a negative opinion regarding the old and possibly inferior treatment. Some studies concerning spinal surgery have been hampered by crossover.[7] There is often a lack of knowledge about the natural history because it is difficult not to treat the problem.

There are 2 studies published on a specific subgroup of these patients, treated with either total disc replacement (TDR) or fusion.[8,9] These studies report similar results between the 2 treatment options.

* Corresponding author: Svante Berg, MD, PhD, Stockholm Spine Center, Löwenströmska Hospital, 19489 Upplands Väsby, SE-195 89 Stockholm, Sweden.
E-mail address: svante.berg@spinecenter.se

1935-9810
doi:10.1016/j.esas.2010.01.005

The Swedish Spine Register (SweSpine) was started in 1994.[10,11] It originally covered lumbar spine procedures but now includes all types of spinal procedures. Most clinics in Sweden that perform spine surgery report to the register. In all, about 75% of procedures are reported.[12] The register included a total of 41,000 spinal procedures in the spring of 2009 (http://www.4s.nu/4s_eng/index.htm), and several publications are based on SweSpine.[13,14] A similar register, Spine Tango, was started in 2002 in Switzerland[15,16] and has developed into a register for several countries in Europe.

There are still no reports comparing RCT with register data in this patient group. We therefore decided to compare the results of an RCT that was undertaken at Stockholm Spine Center, Stockholm, Sweden,[17] comparing TDR and posterior instrumented lumbar fusion, with data using the register on all other patients operated on with 1 of these 2 treatment options for CLBP, diagnosed as DDD at 1 or 2 levels, at the same clinic. The main question was whether comparison between 2 treatment methods can be made by means of evaluation of register data or whether such a method has such serious drawbacks that an RCT is required to provide basic scientific data. We defined 3 aims of the study:

1. To evaluate our selection of patients. Was the RCT performed on a fair sample of patients from this group receiving treatment at the clinic? If not, how was the selection of patients affected by the inclusion and exclusion criteria that were defined in the RCT?
2. To compare results between RCT and retrospective registered data. Were the results after 1 and 2 years for the 2 surgical methods similar in the RCT and non-RCT groups? If not, what could be the reason for the difference? Does a 2-year follow-up contribute to a 1-year evaluation?
3. To analyze the nonresponder problem in the register. Which patients did not respond to the questionnaires? To analyze missing data, we asked how patients who did not answer the 1- and 2-year questionnaires were distributed. Did this jeopardize the results drawn from this study?

The answers to these questions might tell us whether a register study can replace RCTs in the future and, if so, what basic requirements must be reached by the register.

Materials and methods

SweSpine includes data from the attending surgeon on diagnosis, type of surgery, levels treated, bone grafting, implants, complications, antibiotics, and length of stay at the hospital. Patients fill out baseline data on questionnaires immediately before the operation and outcome questionnaires at 1 and 2 years postoperatively.

The patients' baseline data consist of age at treatment, gender, smoking status, previous spine surgery and frequency, work status, earlier back problems, medication, other illness, diagnosis, back and leg pain on a visual analog scale (VAS),[18] EuroQol instruments (EQ-5D and EQ VAS),[19] Short Form 36 (SF-36),[20] and disease-specific disability with the Oswestry Disability Index (ODI).[21]

The outcome variables are complications, reoperation, work status, medication, global assessment of back and leg pain, VAS for back and leg pain, EQ-5D, EQ VAS, SF-36, ODI, and patient's satisfaction with and opinion on the result of the treatment. Global assessment of back pain constituted the primary outcome variable in the RCT at 2 years: total relief, much better, better, unchanged, or worse.[22]

The study group constituted all patients at Stockholm Spine Center with a diagnosis of segmental pain (Fig. 1)

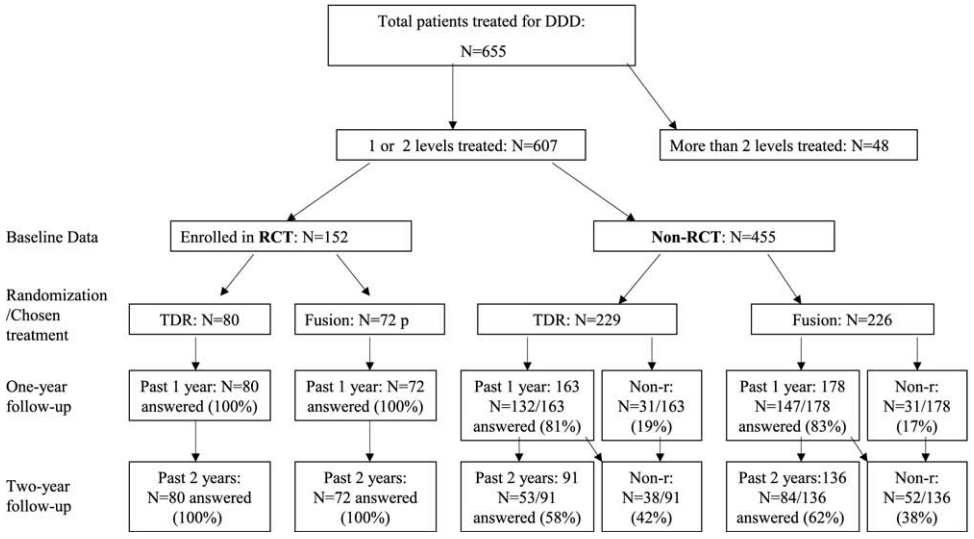

Fig. 1. Total number of patients included in spine register at clinic for DDD surgery (fusion or TDR) from September 2003 until the end of 2008. The figure illustrates how the entire material was divided into gradually smaller subgroups. Non-r, nonresponders.

operated on from September 2003 until the end of 2008, either with TDR or posterior instrumented lumbar fusion, at 1 or 2 levels from L3 and below, aged between 20 and 55 years (criteria for diagnosis are given later). These patients were divided into 2 subgroups: those included in the RCT and those not included in that study (non-RCT group).

The RCT comparing TDR and instrumented posterior lumbar fusion was approved by the Ethics Committee of the Karolinska Institute, Stockholm, in 2003 (03-268).[17] The patients had have symptomatic DDD in 1 or 2 motion segments, with low-back pain as the predominant symptom, although leg pain was not a contraindication. For inclusion in the study, the following conditions were required: back pain diagnosed as mechanical and discogenic in origin with interspinous tenderness on examination, disc narrowing on radiographs, and signs of degeneration on magnetic resonance imaging. Low-grade facet joint arthritis at the index level, as well as low-grade degeneration at other levels, was accepted. Patients who fulfilled the inclusion criteria at the primary consultation but scored lower on the ODI and VAS at the time of surgery were not excluded. The preoperative values served as baseline. After inclusion, patients were randomized between fusion and TDR by use of a closed-envelope technique. The surgeons were not informed of the result of randomization until the patient arrived at the hospital for surgery; the patient was also informed at this time. The inclusion and exclusion criteria for the RCT are summarized in Table 1. Patients in the non-RCT group might not have met these criteria.

Patients with a strong belief that one treatment option was superior to the other were not included in the RCT but could be included in the non-RCT group.

The non-RCT group consisted of the remaining patients in the register after the RCT patients were excluded. We did not exclude patients with low disability or low pain level, because we did not exclude patients who had improved between planning consultation and admittance for surgery in the RCT group. Inclusion was based on medical history and examination at earlier consultations, and the baseline level was based on the preoperative questionnaires. Back pain was mandatory, but it did not have to be predominant over leg pain. According to clinic routine, disc degeneration was mandatory for surgical fusion or TDR. Severe facet joint arthritis was not a strict exclusion criterion for any of these surgical methods but could have affected the type of treatment that was chosen. Locally provoked pain such as interspinous tenderness was not mandatory but was noted in most patients. The treatment each patient received in the non-RCT group was the result of a decision made between the patient and the surgeon at a consultation before admittance to the hospital. The patients were also informed about the register at this time.

Demographics

The total group that was treated at 1 or 2 levels consisted of 607 patients, 54% women, with a mean age of 41 years (range, 21 to 55 years). Of these patients, 309 were treated with TDR and 298 with fusion.

In the RCT subgroup surgical technique was randomly selected, but in the non-RCT subgroup the treatment given was a result of a decision made between the patient and the surgeon.

Demographics in RCT group

Patients referred to the clinic for surgical treatment of DDD who fulfilled the inclusion criteria and had no exclusion criteria were consecutively enrolled for the RCT. In total, 152 patients were included in the RCT: 90 women and 62 men. Of these patients, 80 were treated with TDR and 72

Table 1
Inclusion and exclusion criteria for RCT

Inclusion criteria	Exclusion criteria
• Low-back pain with or without leg pain for >1 year; if leg pain occurred, then low-back pain should dominate • Failure of conservative treatment scheduled for >3 months • Radiographic confirmation of disc degeneration on magnetic resonance imaging • Age 20–55 years • ODI ≥30 or back pain (VAS) >50/100 the week before inclusion • Signed informed consent • Patient having an open mind regarding the 2 treatment options	• Spinal stenosis requiring decompression • Moderate or worse facet joint arthritis • Three or more painful levels at clinical examination • No obvious painful level(s) at diagnostic injection evaluation (if done) • Isthmic spondylolysis/spondylolisthesis • Degenerative spondylolisthesis >3 mm • Major deformity • Manifest osteoporosis—if osteoporosis was suspected because of gender and age (women aged >50 years), illness, or medication, osteoporosis should be evaluated and excluded before inclusion • Previous lumbar fusion or decompression with potential instability (eg, facet joint damage or wide laminectomy) • Compromised vertebral body • Previous spinal infection or tumor • Inability to understand information because of abuse or for psychological or medical reasons • Language difficulties with inability to understand follow-up instruments • Pregnancy or other medical condition that would be a contraindication to surgery

Table 2
Key variables

	TDR/RCT	TDR/non-RCT	FUSION/RCT	FUSION/non RCT
Baseline variables in the RCT and the non-RCT group (± SD)				
Demographic				
N	80	163	72	178
Age	40	40	39	43
Gender	61% fem	49% fem	58% fem	54% fem
Smokers	10%	11%	11%	19%, $P < 0.05$
Surgical				
Prior surgery	11%	26%, $P = 0.007$	12%	31%, $P < 0.002$
Levels	1 = 55%	1 = 60%	1 = 45%	1 = 60%
	2 = 45%	2 = 40%	2 = 55%	2 = 40%
Outcome variables at baseline				
Back Pain VAS (0–100)	62 (± 21)	60 (± 20)	58 (± 22)	63 (± 20)
ODI (0–100)	42 (± 12)	41 (± 12)	41 (± 14)	45 (± 14)
EQ5D	0.41 (± 0.31)	0.39 (± 0.31)	0.38 (± 0.33)	0.34 (± 0.32)
EQVAS	46 (± 22)	51 (± 20)	47 (± 23)	47 (± 21)
	TDR/RCT	TDR/non-RCT	FUSION/RCT	FUSION/no
One year outcomes of key variables				
Back pain (VAS)	25 (± 27)	27 (± 26)	34 (± 27)	35 (± 30)
Back Pain Improvement(VAS)	37 (± 30)	33 (± 30)	24 (± 34)	30 (± 30)
ODI (0–100)	19 (± 19)	20 (± 17)	25 (± 16)	30 (± 23)
ODI improvement	23 (± 18)	19 (± 17)	16 (± 18)	15 (± 19)
EDQ5	0.71 (± 0.28)	0.69 (± 0.29)	0.63 (± 0.27)	0.55 (± 0.36)
EDQ5 improvement	0.29 (± 0.35)	0.28 (± 0.37)	0.25 (± 0.37)	0.20 (± 0.37)
Two year outcomes of key variables				
Back pain (VAS)	25 (± 30)	22 (± 25)	30 (± 24)	38 (± 32)
Back Pain Improvement (VAS)	37 (± 31)	34 (± 28)	28 (± 32)	24 (± 31)
ODI (0–100)	20 (± 20)	18 (± 16)	23 (± 17)	30 (± 21)
ODI improvement	22 (± 19)	20 (± 18)	17 (± 19)	14 (± 19)
EQ5D	0.67 (± 0.33)	0.70 (± 0.29)	0.69 (± 0.25)	0.58 (± 0.36)
EQ5D improvement	0.26 (± 0.37)	0.31 (± 0.38)	0.31 (± 0.37)	0.23 (± 0.33)

Abbreviations: ODI, Oswestry disability index (Original version); EQ5D, EuroQol; TDR, total disc replacement; RCT, randomized controlled trial.

with instrumented fusion. Forty-four patients had posterolateral fusions, and twenty-eight had posterior lumbar interbody fusions. There were no differences between the treatment groups in age, gender, smoking status, baseline ODI, SF-36, EQ-5D, surgical levels, prior surgical treatment, or back pain (Table 2).

Demographics in non-RCT group

When patients included in the RCT were excluded from the total group, 455 patients remained in the non-RCT group (52% women). There were 163 treated with TDR (11% smokers), and 178 were fused (19% smokers). All fusions were instrumented; 82 were posterolateral fusions, and 96 were posterior lumbar interbody fusions (Table 2).

Methods

Baseline data were obtained from the preoperative questionnaire and surgeon's register chart. Baseline data from patients in the RCT group were compared with the non-RCT group to evaluate patient selection. Patients in the non-RCT group treated with TDR were compared with those treated with fusion to determine whether any differences in baseline data influenced the decision on what treatment option each respective patient was offered.

To compare RCT results with non-RCT data, outcome data from the TDR group and the fusion group in the RCT subgroup were compared with the corresponding patients in the non-RCT subgroup. ODI success was described as greater than 25% improvement compared with greater than 15% in Food and Drug Administration studies on TDR.[8,9]

Questionnaires were administered and clinical follow-up visits performed after 1 and 2 years. In the non-RCT group not all patients answered the questionnaire at 1 year and even fewer did so at 2 years. To evaluate to what extent the lower reply frequency in this group affected the results, an analysis was made to determine which patients had not answered the questionnaires, for example, whether it was patients with the worst or best prognosis or whether the nonresponders were evenly distributed in the material and between the 2 treatment options. This analysis was performed by comparing baseline with 1-year questionnaire answers and comparing 1-year with 2-year questionnaire answers in the non-RCT group to determine whether the nonresponders differed at baseline or 1-year follow-up from those patients who answered.

Statistical methods

The RCT study was dimensioned to compare TDR and fusion with global assessment of back pain at 2 years as the primary outcome variable. "Total relief" was considered the optimum result and the primary endpoint, and "much better" was interpreted as essential improvement in contrast to "better," "unchanged," and "worse." The Lehr formula was used to provide crude estimates of sample size.[23] With 80% power at 5% significance level, the size of each group was estimated to be 64 patients, which was increased to 72 to allow for potential dropout. The results in this study were from our comparison of the clinical results of the RCT group with those of the non-RCT group. For this comparative study, no extra power calculation was made because the material was larger.

Statistical analysis was performed by use of Statistica, version 7 (StatSoft, Tulsa, Oklahoma). Results are given as mean ± standard deviation, and confidence intervals (CIs) for differences between groups were calculated. For comparison between the treatment groups, as well as for some subgroup analyses, 2-tailed Mann-Whitney *U* and Wilcoxon rank sum tests were used. For ordinal data, the Student *t* test was used, and for categorical data (eg, global assessment), Spearman *r*, Fisher exact, and χ^2 tests were used. Statistical significance was defined as $P < .05$.

Results

The total group of registered patients who were treated more than 1 year before this study was 493 (54% women and 46% men). The RCT group comprised 152 patients, and the non-RCT, 341 patients.

Examination of baseline data

In the RCT group the mean age was 39.5 years (range, 21 to 55 years), and in the non-RCT group the mean age was 41.3 years (range, 22 to 55 years) ($P = .014$; CI, −3.19 to −0.35). Fusion patients had a mean age of 38.7 years in the RCT group and 42.7 years in the non-RCT group ($P < .00009$; CI, −5.91 to −2.00). There was no difference in mean age between TDR patients in the RCT group (40.3 years) and those in the non-RCT group (39.8 years). There was a difference in age in the non-RCT group between patients treated with TDR (39.8 years) and those treated with fusion (42.7 years) ($P < .0002$; CI, −4.36 to −1.37), and there were fewer smokers in the TDR group ($P = .041$). In the non-RCT group the fusion patients had a higher ODI than the TDR patients (45 vs 41) ($P < .005$; CI, 1.30 to 7.16).

The patients selected for TDR in the non-RCT group resembled the total RCT group at baseline. There were differences in patients that received a fusion between the RCT and the non-RCT group, where in the later group patients were older, more frequently smokers and had a higher ODI preoperatively.

Outcome variables at 1 year

In the RCT group results in the TDR group were better than those in the fusion group regarding the number of totally pain-free patients ($P = .003$), global assessment of back pain ($P = .013$), reduction in back pain VAS ($P = .017$; CI, −22.8 to −2.3), leg pain VAS ($P = .007$), reduction in ODI ($P = .019$; CI, −12.5 to −0.9), and 3 of 10 SF-36 domains (Table 3).

In the non-RCT group results in the TDR group were better both numerically and regarding the number of variables that improved. Patients treated with TDR had better results than patients treated with fusion at 1 year in global assessment of back pain ($P = .024$), return to work ($P < .001$; CI, −0.89 to −0.38), back pain VAS ($P = .031$; CI, 0.41 to 13.65), EQ VAS ($P = .013$; CI, −13.78 to −2.58), EQ-5D ($P < .003$; CI, −2.21 to 0.06), ODI ($P < .001$; CI, 4.76 to 14.34), and all 10 SF-36 domains.

Outcome variables at 2 years

Results in the RCT TDR group were better than in the fusion group regarding totally pain-free patients ($P = .031$), back pain VAS ($P = .048$), and leg pain VAS ($P = .037$). The other differences seen at 1 year were no longer apparent (Table 4).

In the non-RCT group the TDR group still had better results than the fusion group for the majority of outcome measures, with small changes from 1-year data.

Table 3
Clinical outcome at one year in RCT group and non-RCT group (± SD). Percentage indicates the responding patients that have passed one year

	TDR/RCT n = 80 100% answer	TDR/non-RCT n = 163 80% answer	FUSION/RCT n = 72 100% answer	FUSION/non-RCT, n = 178 83% answer
Global assessment of back pain	Totally pain-free 29% Unchanged o worse 13%	Totally pain-free 20% Unchanged o worse 11%	Totally pain-free 10% Unchanged o worse 17%	Totally pain-free 11% Unchanged o worse 21%
Totally pain-free or better	88%	89%	83%	79%
On full sick-leave	32%	24%	40%	37%
Satisfied with result	76%	75%	65%	65%
Reop	1%	8%	7%	10%
EQVAS	69 (25)	71 (22)	67 (22)	62 (26)
EQVAS improvement	24 (26)	20 (23)	20 (26)	16 (23)

Table 4
Clinical outcome at two years in RCT group and non-RCT group (± SD). Percentage indicates the responding patients that have passed two years

	TDR/RCT n = 80 100% answered	TDR/non-RCT n = 91 64% answered	FUSION/RCT n = 72 100% answered	FUSION/non-RCT n = 136 62% answered
Global assessment of back pain	Totally pain-free 30% Unchanged o worse 13%	Totally pain-free 32% Unchanged o worse 15%	Totally pain-free 15% Unchanged o worse 14%	Totally pain-free 14% Unchanged o worse 21%
Totally pain-free or better	88%	85%	86%	76%
Satisfied with result	71%	79%	67%	58%
Reop	5%	0%	12%	10%
EQVAS	67 (26)	74 (21)	69 (22)	62 (27)
EQVAS improvement	22 (27)	20 (24)	22 (28)	14 (29)

There were clear differences between the RCT group and the non-RCT group in outcome. At 1 year, both studies showed a better result for TDR than fusion, but the advantages in favor of TDR were greater in the non-RCT group. At 2 years, several differences in results between TDR and fusion were no longer seen in the RCT group but remained in the non-RCT group.

Effect of nonresponders on results

Follow-up in the RCT group was 100% at both 1 and 2 years, whereas in the non-RCT group, follow-up was 80% in TDR patients and 83% in fusion patients at 1 year and 64% and 62%, respectively, at 2 years. Patients who did not answer the 1-year follow-up questionnaire differed in EQ-5D ($P < .04$), ODI ($P = .05$), and gender proportion in terms of baseline data from the rest of the patients.

Patients who answered the 1-year but not the 2-year follow-up questionnaire had a somewhat worse outcome at 1 year compared with the rest of the group regardless of type of treatment.

Discussion

Patient selection

Patients who were treated with TDR were similar, whether they were in the RCT group or the non-RCT group. To a large extent, this is because the same inclusion and exclusion criteria were used before a TDR was suggested to any patient. On the other hand, only fusions in the RCT were selected under the same inclusion and exclusion criteria, which is why the register data alone are not suitable for comparison between these 2 surgical methods.

When trying to understand these differences, we examined our inclusion and exclusion criteria for the RCT. The criteria on age and what levels and number of levels to treat were similar between the RCT group and the non-RCT group. It was not possible to track the remaining inclusion and exclusion criteria in the register. Patients with more than slight facet joint arthritis were not included in the RCT, or if they were in the non-RCT group, TDR was not suggested. Furthermore, patients in the non-RCT group with severe spondylotic changes, making the possibility to restore mobility questionable, were not offered TDR treatment. Patients in the non-RCT group chosen for fusion were on average 3 years older than patients chosen for TDR despite the fact that these patients were from the same age group (range, 20 to 55 years), indicating an age difference within the group. The fusion patients also had a higher ODI than patients chosen for TDR. The patients treated with fusion in the non-RCT group may have had more advanced degenerative changes than patients treated with TDR. It is possible that TDR, when used outside an RCT, is offered to a subgroup of patients with a better prognosis. The RCT group was not a fair sample of patients treated at our clinic for this diagnosis, and the results of the RCT cannot be applied to all patients with DDD and a definite subgroup could not be defined. More preoperative baseline data in the register, especially X-ray and magnetic resonance imaging findings, could probably help in this respect. An increase in the number of variables to be registered preoperatively might, on the other hand, lead to a lower registration frequency.

However, in most patients who presented with adjacent-level disease in this age group, who were in the non-RCT group, TDR was suggested. These patients also represent a subgroup with a poorer prognosis.

Outcome

The results of the RCT showed that at 1-year follow-up, the group treated with TDR rated their improvement higher and their pain scores lower for several of the outcome measures. At the 2-year follow-up, these differences were smaller.

Compared to the RCT, the clinical outcome differences seen among the non-RCT patients were larger to the advantage of TDR and remained unchanged between the one- and two-year follow-up. The differences were larger in absolute scores than in changes from baseline, indicating that the TDR patients and the fusion patients in the non-RCT group differed from the start.

The results in the non-RCT group support the results in the RCT group that, for the subgroup of patients with mechanical low back pain, diagnosed as DDD, and fitting in with the inclusion and exclusion criteria of the RCT, a

somewhat better result could be expected with TDR than with fusion. When one just looks at the register, the results appear somewhat exaggerated because the fusion group in the register seems to have been worse from the start.

In the RCT the 1- and 2-year follow-up was 100%, as compared with 88% and 90%, respectively, in the Food and Drug Administration studies.[8,9] There was a lower response to the questionnaire in the non-RCT group. In SweSpine the 1- and 2-year follow-up figures are 70% to 75% and 50% to 60%, respectively (P. Fritzell, oral personal communication P Fritzell, August 2009), so the our study's follow-up frequency is better than average. Patients who did not answer at 1 year differed slightly at baseline from the responding patients, and patients who did not respond at 2 years had a somewhat worse result at 1-year follow-up. We therefore conclude that nonresponders slightly jeopardized the 2-year results in a positive direction for both TDR and fusion.

All things considered, the difference in results between the RCT and non-RCT groups seems mainly to be caused by the fact that the fusion subgroup in the non-RCT group differed from the rest of the patients. Nonresponders may also have contributed to the difference in results.

Conclusion

We find Sweden's spine register (SweSpine) to be very useful in many aspects, but the fact that the results from the RCT are different from the non-RCT results at both 1 and 2 years indicates that the RCT study was performed in a subgroup of patients with CLBP and that the worst cases might not have been included. This difference appears mainly to be the result of a deliberate clinical decision on what type of surgical procedure to offer the patient. In all register studies on the general population, there will be differences between outcomes for different treatment options because of selection differences for each method. The results of this comparative study show the absolute necessity for RCT studies when comparing treatment options.

The Spine Tango register has mainly been used for evaluation of the safety of the use of TDR in Switzerland. Neither selection of patients nor characteristics of dropout patients are described, and it therefore cannot serve as a basis for further decisions regarding the method and its usefulness or as a basis for quality control.

References

1. Fritzell P, Hägg O, Wessberg P, et al. Lumbar fusion versus nonsurgical treatment for chronic low back pain. A multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine (Phila Pa 1976)* 2001;26:2521–34.
2. Brox JI, Sørensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine (Phila Pa 1976)* 2003;28:1913–21.
3. Resnick DK. Evidence-based spine surgery. *Spine (Phila Pa 1976)* 2007;32:15–9.
4. Chou R, Loeser JD, Owens DK. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain. An evidence-based clinical practice guideline from the American Pain Society. *Spine (Phila Pa 1976)* 2009;34:1066–77.
5. Chou R, Baisden J, Carragee EJ, et al. Surgery for low back pain. A review of the evidence for an American Pain Society clinical practice guideline. *Spine (Phila Pa 1976)* 2009;34:1094–109.
6. Copay AG, Glassman SD, Subach BR, et al. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and Pain Scales. *Spine J* 2008;8: 968–74.
7. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation. The spine patient outcomes research trial (SPORT): a randomized trial. *JAMA* 2006;296:2441–50.
8. Blumenthal S, McAfee PC, Guyer RD, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the Charité artificial disc versus lumbar fusion. Part I: evaluation of clinical outcomes. *Spine (Phila Pa 1976)* 2005;30:1565–75.
9. Zigler J, Delamarter R, Spivak JM, et al. Results of the prospective, randomized, multicenter Food and Drug Administration Investigational Device Exemption Study of the ProDisc-L total disc replacement *versus* circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine (Phila Pa 1976)* 2007; 32:1155–62.
10. Strömqvist B, Jonsson B, Fritzell P, et al. The Swedish national register for lumbar spine surgery. *Acta Orthop Scand* 2001;72:99–106.
11. Strömqvist B, Fritzell P, Hägg O. The Swedish Spine Register: development, design and utility. *Eur Spine J* 2009;18(Suppl 3):294–304.
12. Strömqvist B, Fritzell P, Hägg O, et al. One-year report from the Swedish National Spine Register. *Acta Orthop Scand* 2005;76:1–24.
13. Fritzell P, Strömqvist B, Hägg O. A practical approach to spine registers in Europe: the Swedish experience. *Eur Spine J* 2006;15:57–63.
14. Strömqvist B. Evidence-based lumbar spine surgery. The role of national registration. *Acta Orthop Scand* 2002;73(Suppl 305):34–9.
15. Melloh M, Staub L, Aghayev E, et al. The international spine registry SPINE TANGO: status quo and first results. *Eur Spine J* 2008;17: 1201–9.
16. Aebi M, Grob D. SSE Spine Tango: a European Spine Registry promoted by the Spine Society of Europe (SSE). *Eur Spine J* 2004; 13:661–2.
17. Berg S, Tullberg T, Branth B, et al. Total disc replacement compared to lumbar fusion: a randomized controlled trial with 2-year follow-up. *Eur Spine J* 2009;8:1512–9.
18. Zanoli G, Strömqvist B, Jönsson B. Visual analog scales for interpretation of back and leg pain intensity in patients operated for degenerative lumbar spine disorders. *Spine (Phila Pa 1976)* 2001;26:2375–80.
19. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37: 53–72.
20. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 1992;30:473–83.
21. Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.
22. Hägg O, Fritzell P, Odén A, et al. Simplifying outcome measurement. Evaluation of instruments for measuring outcome after fusion surgery for chronic low back pain. *Spine (Phila Pa 1976)* 2002;27:1213–22.
23. Lehr R. Sixteen S-squared over D-squared: a relation for crude sample size estimates. *Stat Med* 1992;11:1099–102.