Clinical outcomes after treatment with disc prostheses in three lumbar segments compared to one- or two segments

Svante Berg and Nina Gillberg-Aronsson

*Int J Spine Surg* 2015, 9 ()
doi: [https://doi.org/10.14444/2049](https://doi.org/10.14444/2049)
[http://ijssurgery.com/content/9/49](http://ijssurgery.com/content/9/49)

This information is current as of October 4, 2020.

---

**Email Alerts**  
Receive free email-alerts when new articles cite this article. Sign up at:  
[http://ijssurgery.com/alerts](http://ijssurgery.com/alerts)
Clinical outcomes after treatment with disc prostheses in three lumbar segments compared to one- or two segments.

Svante Berg, MD, PhD,1 Nina Gillberg-Aronsson, Med stud.2
1Stockholm Spine Center, Löwenströmska Hospital, Stockholm, Sweden, 2Medical School, Karolinska Institutet, Stockholm, Sweden

Abstract

Background
Fusion surgery in the rare patients suffering from symptomatic degenerative disc disease (DDD) at three segments has been reported to produce poor results and a high frequency of complications, why patients suffering from DDD at three segments are seldom offered surgical treatment.

Purpose
To compare clinical outcome after one- and two years, between patients that have undergone disc replacement surgery (TDR) at three segments and patients that have been treated at less segments.

Methods
The present study is based on data recorded in the Swedish Spine Registry (SweSpine). The study group consisted of 30 patients who underwent three-segment TDR, the comparative group of 700 patients treated in one or two segments. Analyses included comparisons of preoperative data, postoperative results and improvement from baseline.

Results
Our results showed no differences in outcome between groups at one- and two years postoperatively. Improvements achieved after surgery in both groups well exceeded established values for minimally clinically important difference (MCID).

Conclusions
The results of this study show that patients with a diagnosis of therapy-resistant chronic low back pain (CLBP) due to DDD in one, two or even three lumbar segments achieve similar and good results after TDR surgery.

Clinical relevance
The rare patients with severe and convincing DDD from three segments might in carefully selected cases be offered surgery, with a reasonable chance of a beneficial outcome.

Introduction
For patients diagnosed with severe chronic low back pain (CLBP) and segment-specific pain due to degenerative disc disease (DDD), surgery may be an option when initial and prolonged conservative treatment has failed to relieve pain and/or restore quality of life and physical function. However, surgery should only be considered for patients who experience no substantial relief from multimodal conservative treatment and are largely disabled.1

Randomized controlled trials (RCTs) comparing clinical outcomes after non-surgical treatment to clinical outcomes after different surgical interventions in selected patient groups have reported significant positive differences with respect to surgery.2,3 Surgical alternatives mainly consist of different spinal fusion techniques and techniques that aim to preserve motion. These latter techniques often require the implantation of disc prostheses in the affected spinal segment(s), a treatment called total disc replacement (TDR).4-6 Hellum et al. reported no major differences in complication rates between fusion
surgery and TDR surgery. In a recent RCT comparing clinical outcomes for different fusion techniques and TDR for patients suffering from CLBP due to DDD, the majority of patients, irrespective of surgical intervention, exhibited sustained significantly improved clinical outcomes. Compared to the fusion group, the TDR group, however, reported better outcome parameters after both two and five years. A study using human cadavers reported significant differences in favour of TDR when comparing the applied effects of two-segment TDR and two-segment interbody fusion on spinal ROM and adjacent intradiscal pressure properties. Maintained and/or restored mobility after spinal surgery could decrease the frequency of adjacent segment disease postoperatively. In patients with advanced stages of DDD, often accompanied with facet-joint arthritis, fusion still seems to present the only surgical alternative.

TDR surgery is usually performed on one or two spinal segments. A small subgroup of patients suffer from segment-specific pain emanating from three spinal segments; these patients might be treated with TDR on all three segments, if necessary. As a result of earlier treatment options (fusion surgery) and probably due to a more widespread and possibly systemic degenerative disease, these rare patients have been considered much worse from a outcome-prognostic point of view. However, it has not been possible to find differences in results between patients who have undergone TDR in one spinal segment and patients operated on in two spinal segments. Moreover, postoperative clinical outcomes in the small subgroup treated with three-segment TDR have not previously been investigated, although it is of great importance to analyse the clinical results for this patient group as this information can help doctors understand whether it is worthwhile to treat this patient group with TDR surgery. To this end, this study analyses the differences observed in patient-reported clinical outcomes after one and two years between patients who underwent three-segment TDR (group A) and patients who underwent one- or two-segment TDR (group B). The authors are fully aware group A and group B might represent very different patient cohorts, but represents probably the closest comparison groups available.

Materials and Methods

Based on baseline data and clinical outcomes reported to the Swedish Spine Registry (SweSpine), this retrospective register study includes patients from one surgical clinic who received TDR as treatment for CLBP due to DDD. All patients that were treated had previously undergone a prolonged period of physiotherapy and training, most frequently with a multimodal approach at the same clinics rehabilitation department. The patients had also been screened for psychosocial disorders as well as medication dependence.

To observe true differences between the two groups, clinical outcome parameters widely used and validated in the field of low back pain were examined. Using preoperative and postoperative questionnaires, the patients provided preoperative data as well as data on outcome after one and two years. These data were reported to and registered in SweSpine. All postoperative questionnaires were sent by post to patients and completed in private, strategies intended to limit the influence of caregivers. The completed questionnaires were sent to and registered at one location, and administrative personnel at the treating clinic acquired the results from this centre. All included questionnaire parameters were registered preoperatively at admission to the hospital and at one and two years postoperatively except the Global Assessment of back pain that was included only postoperatively.

Before surgery, all patients in the study were diagnosed with CLBP due to DDD by clinical examination, medical history and radiographic investigation. In addition, all patients in group A were examined using discography as a further source of information. Discography was only performed in discs that had signs of degeneration on MRI. These degenerative changes were dehydration and at least some loss of disc-height compared to non-dehydrated lumbar discs in the same patient. Modic changes were frequently but not always present. The method for discography was with pressure provocation followed by injection of local anaesthetics for each examined disc. For surgery to be warranted, discography had to be positive for all three degenerated discs. Group B consisted of patients where one or two disc
was dehydrated and had a loss of disc-height, frequently accompanied by Modic changes. 16% of the patients in group B were investigated with discography, due to uncertainty whether to treat one- or two segments. During the time of this study, three different brands of disc-prostheses were used at the study cite, but each patient was treated with only one of these brands respectively. This study was approved by The Ethics Committee of the Karolinska Institute (2013/2199-31).

Inclusion and exclusion criteria

All patients treated with TDR from September 2003 until September 2012 were included, except hybrid cases and cases where TDR was used to treat adjacent segment disease post fusion surgery. Group A consisted of 30 patients treated from L3 to S1. Group B consisted of 700 patients treated with TDR in one or two segments during the same period as group A, this to minimise any possible influence of technical advances. Patients operated on after September 2012 were excluded due to the requirement of at least one-year follow-up. Figure 1 and Figure 2 show flowcharts of patients included in group A and group B. All patients in group A participated in at least one follow-up. No patient in group A had a reoperation at the index segment. In group B, a total sum of 81 patients did not participate in any of the follow-ups (i.e. 1- and/or 2-year follow-up) to the index operation. Two of these patients (0.3%) had a reoperation at the index segment before one-year follow-up and another three of these patients (0.4%) had a reoperation at the index segment before two-year follow-up. The baseline data and demographics of these 81 patients did not differ from the patients that answered follow-up questionnaires.

Outcome measures

Over the last several years, the use of patient-reported clinical outcomes has increased. Hägg et al., for example, concluded that global assessment of back pain (GA) adequately summarizes other outcome variables. GA is a five step postoperative scale: 0= no LBP before surgery, 1= LBP is completely gone, 2= LBP is largely reduced, 3= somewhat reduced LBP, 4= unchanged LBP, 5= worsened LBP. In this study, clinical outcomes were assessed based on the following previously validated patient-reported outcome measurement instruments: GA, VAS back and leg pain, European Quality of Life 5D (EQ-5D), 36-Short Form health survey (SF-36), and Oswestry Disability Index (ODI), as advised by Chapman et al., for evaluating the importance of findings and for power calculations, we considered the Minimally

---

Fig. 1. Flowchart of included patients in groups A and B (n = number of patients, follow-up percentage) at 1-year follow-up. Numbers deviating in separate text boxes represent patients lost to follow-up. 1 = both patients lost to follow-up at one year answered the two-year follow-up questionnaire in group A. 2 = out of the patients lost to follow-up at one year in group B, 81 patients did not answer the two-year follow-up questionnaire.

Fig. 2. Flowchart of included patients in groups A and B (n = number of patients, follow-up percentage) at 2-year follow-up. Numbers deviating in separate text boxes represent patients lost to follow-up. One patient had not reached 2 years. Discrepancies in total numbers between initial patient selection and total numbers at two-year follow-up are due to patients who had not yet reached the two-year follow-up.

Downloaded from http://ijssurgery.com/ by guest on October 4, 2020
Clinically Important Difference (MCID).\textsuperscript{18-21}

**Statistical Methods**
Statistical analyses were performed with Statistica version 7 (StatSoft Inc. Tulsa, OK, USA). A demographic analysis was performed. Because of the large difference in total numbers between the groups, statistically significant differences concerning demographic parameters were tested with the two-tailed Mann-Whitney U test. In both groups, the results from outcome measurements were considered to be normally distributed. Results from outcome measurements (except the GA score) and including the ODI- and VAS-improvement were calculated as means and standard deviations with descriptive statistics. To analyse whether differences in reported results from outcome measurements (except the GA score) between the groups were statistically significant, we used the two-tailed Mann-Whitney U test. The proportion of patients in each group reporting each GA category was calculated as percentages. Fisher’s exact test was used for statistical significance analysis regarding differences in GA results between the groups. Power calculations, performed using two means t-test for individual samples, were designed to detect a difference of at least ten points between ODI-improvement means in both groups. At 5% significance level, the actual group sizes were calculated to achieve 84.26% power. Statistical significance was defined as $p < 0.05$.

**Results**
Demographics of the groups were similar (Table 1). Pre- and postoperative data, given as means and standard deviations of all outcome parameters and their improvement, are presented in Table 2, Table 3, Table 4, and Table 5. All stated p-values are calculated for the differences between group A and B.

**Primary outcome, GA**
No statistically significant differences were found between the groups when evaluating the mean in GA scoring at one- and two-year follow-up. In addition, there were no statistically significant differences when comparisons were made between percentages of patients in each group who reported at the different GA categories. GA results, presented as percent-ages of the included patients reporting to each GA category (GA1-5), are displayed in Figure 3. In group B, two (0.3%) out of all of included patients at both

<table>
<thead>
<tr>
<th>Table 1. Demographics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number n = 700 n = 30 p</td>
</tr>
<tr>
<td>Mean age at operation date 41 ± 9 (range 18-64) 40 ± 9 (range 21-54) 0.81</td>
</tr>
<tr>
<td>Women n = 354 n = 12 0.25</td>
</tr>
<tr>
<td>Men n = 346 n = 18 0.25</td>
</tr>
<tr>
<td>Preoperative smoking (percentage of group) n = 67 (9.5%) n = 4 (13%) 0.47</td>
</tr>
</tbody>
</table>

Demographics of the included patients. P-values describe intergroup differences. Abbreviations: n = numerical value; NA = not applicable.

<table>
<thead>
<tr>
<th>Table 2.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome measurement Preoperative Group B Group A p</td>
</tr>
<tr>
<td>VAS back pain 60 ± 20 61 ± 21 0.80</td>
</tr>
<tr>
<td>VAS leg pain 34 ± 28 32 ± 26 0.72</td>
</tr>
<tr>
<td>EQ-5D 0.4 ± 0.3 0.3 ± 0.3 0.17</td>
</tr>
<tr>
<td>SF-36 PCS 40.0 ± 13.5 38.9 ± 11.5 0.65</td>
</tr>
<tr>
<td>SF-36 MCS 33.2 ± 9.1 33.0 ± 7.6 0.94</td>
</tr>
<tr>
<td>ODI 40 ± 13 43 ± 12 0.23</td>
</tr>
</tbody>
</table>

Outcome measurement results preoperatively for group A and group B. Abbreviations: VAS = visual analogue scale; EQ-5D = European Quality of Life 5D; SF-36 PCS = 36-short form health survey, physical component summary; SF-36 MCS = 36-short form health survey, mental component summary; ODI = Oswestry disability index.

<table>
<thead>
<tr>
<th>Table 3.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome measurement At 1-year follow-up Group B Group A p</td>
</tr>
<tr>
<td>VAS back pain 23 ± 25 24 ± 28 0.95</td>
</tr>
<tr>
<td>VAS leg pain 16 ± 24 16 ± 24 0.85</td>
</tr>
<tr>
<td>EQ-5D 0.7 ± 0.3 0.7 ± 0.3 0.72</td>
</tr>
<tr>
<td>SF-36 PCS 47.5 ± 12.4 44.2 ± 12.7 0.13</td>
</tr>
<tr>
<td>SF-36 MCS 45.1 ± 11.3 44.1 ± 12.0 0.79</td>
</tr>
<tr>
<td>ODI 18 ± 17 20 ± 21 0.89</td>
</tr>
</tbody>
</table>

Outcome measurement results at 1-year follow-up for study group and control group. Abbreviations: VAS = visual analogue scale; EQ-5D = European Quality of Life 5D; SF-36 PCS = 36-short form health survey, physical component summary; SF-36 MCS = 36-short form health survey, mental component summary; ODI = Oswestry disability index.
one-year and two-year follow-up reported a GA score of zero. Corresponding numbers in the group A were zero at both occasions.

Table 4.

<table>
<thead>
<tr>
<th>Outcome measurement</th>
<th>At 2-year follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group B</td>
</tr>
<tr>
<td>VAS back pain</td>
<td>22 ± 26</td>
</tr>
<tr>
<td>VAS leg pain</td>
<td>15 ± 23</td>
</tr>
<tr>
<td>EQ-5D</td>
<td>0.7 ± 0.3</td>
</tr>
<tr>
<td>SF-36 PCS</td>
<td>48.2 ± 12.1</td>
</tr>
<tr>
<td>SF-36 MCS</td>
<td>46.0 ± 10.9</td>
</tr>
<tr>
<td>ODI</td>
<td>17 ± 17</td>
</tr>
</tbody>
</table>

Outcome measurement results at 2-year follow-up for study group and control group. Abbreviations: VAS = visual analogue scale; EQ-5D = European Quality of Life 5D; SF-36 PCS = 36-short form health survey, physical component summary; SF-36 MCS = 36-short form health survey, mental component summary; ODI = Oswestry disability index.

Table 5.

<table>
<thead>
<tr>
<th>Outcome measurement</th>
<th>Improvement from baseline to 2-year follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group B</td>
</tr>
<tr>
<td>VAS back pain</td>
<td>37 ± 28</td>
</tr>
<tr>
<td>ODI</td>
<td>22 ± 17</td>
</tr>
</tbody>
</table>

Mean improvement from baseline to 2-year follow-up as reported with the VAS and ODI instruments. Abbreviations: VAS = visual analogue scale; ODI = Oswestry disability index.

Fig. 3. Postoperative results reported with the GA score at one and two years. Outcome given as percentages of patients in each population stating a total relief of LBP (GA 1), much less LBP (GA 2), less LBP (GA 3), unchanged LBP (GA 4), and worsened LBP (GA 5). GA = global assessment of postoperative low back pain (LBP). Patients who reported a GA score of zero are not represented.

**VAS back pain**
No statistically significant differences were found between the groups when evaluating participants’ LBP preoperatively and at the one- and two-year follow-up. Analyses concerning mean VAS-improvement from baseline to two-year follow-up revealed no statistically significant differences in improvement (p = 0.59). Both groups well exceeded the suggested MCID (18-20 mm) using the VAS scoring instrument for LBP.19

**VAS leg pain**
No statistically significant differences were found between the groups when evaluating participants’ leg pain preoperatively and at one- and two-year follow-up.

**EQ-5D**
No statistically significant differences were found between the groups preoperatively or postoperatively; however, before the operation group B reported less deterioration in quality of life due to general health (score difference 0.1, p = 0.17).

**SF-36**
No statistically significant differences were found between the groups preoperatively or postoperatively; however, at the one- and two-year follow-up. Group B initially reported less deterioration in quality of life due to general health (p = 0.65 physical component summary; p = 0.94 mental component summary). A similar result was also found in the EQ-5D, the other HRQoL questionnaire used in this study.

**ODI**
No statistically significant differences were found between the groups when exploring the mean scores measured with the ODI preoperatively and at the one- and two-year follow-up. Group B initially reported less disability due to CLBP than group A (p = 0.23). When using the ODI instrument, both groups reported a mean improvement from baseline to the two-year follow-up exceeding the suggested MCID (reduction of ten points).19 The intergroup improvement difference was not significant (p = 0.66). The original full-length questionnaire was used.
Complications

Two patients in group A were re-operated the day after surgery due to subcutaneous hematoma. None of the patients in the study group have required any other reoperation or developed venous thrombosis, nerve injury or deep infection.

Discussion

Although a few studies have described postoperative evaluation of patients post three-segment TDR, no previous study has separately reported group results for clinical outcome after three-segment TDR. To the best of our knowledge, this is the first study to report results from a significant number of three-segment TDR procedures due to DDD causing severe CLBP. The lack of previous studies is not surprising since the number of patients where three-segment treatment is considered, compared to the more frequent one- and two-segment cases, is very low. Nevertheless, we found it important to investigate these results as the results from previous studies investigating three-segment fusions as treatment for DDD might lead to this patient-group being neglected.

The preoperative baseline data were equal between the two groups, but still it must be considered that they are basically different. The results did not reveal any differences in any of the outcome parameters reported after one- and two-years between the groups. Both groups exhibited equal significant benefit from their respective procedures with improvement ratios well over MCID for both VAS and ODI. These good results, which are in contrast to previous reports after fusion, could be due to TDR being a more favourable treatment than fusion surgery for patients with multilevel symptomatic DDD. In addition, it could be argued that the maintained/restored spinal mobility emanating from TDR enables physiological functioning of the spine, leading to both decreased disability and increased quality of life for multilevel affected patients. On the other hand, results from three-segment fusions might be affected by the possibly more advanced degeneration, including facet-joint arthritis, which was not present in the above groups.

Potential complication risks must always be taken into consideration when viewing spinal surgery as an option. Higher complication rates and higher frequency of adjacent segment disease have been reported for patients after multi-segment fusion surgery. SweSpine includes data on reoperations, thrombosis, infections, and nerve injuries. This study did not examine rates of minor complications after three-segment TDR, such as urinary tract infections or prolonged wound-pain. From the perspective of our results, it would seem that the benefit might exceed the potential risk of negative outcome for the individual patient.

Our results suggest that patients with a confirmed diagnosis of therapy-resistant and severely disabling CLBP due to DDD in one, two, or even three spinal segments might in highly selected cases benefit from TDR-surgery. The large number of patients included in the comparative group and the high follow-up frequencies in the study group strengthen this conclusion. Although the follow-up rate was less in the comparative group, the results from this group were similar to previous findings. Our power calculations support the validity of our findings and our results for clinical outcome also support pre- and postoperative data from previous studies on one- and two-segment TDR with the same outcome measurement set.

This present study does have some potential weaknesses. First, it is not clear that patients suffering from DDD in three segments do not differ in other ways than just the number of segments affected, from the one- and two-segment cases, even if baseline data was similar. It is worth mentioning again that the groups are per definition different. Also, the surgeon’s threshold to propose three-segment surgery probably is higher than for cases with fewer segments involved. The preoperative examinations differed because all the patients in the study group had provocative discography before surgery, but rarely in the control group. The authors, however, believe that every available examination was justified for this special patient-group. Also, it is well described that provocative discography might produce both false positive and false negative results, even if that error is reduced when the provocation is com-

Downloaded from http://ijssurgery.com/ by guest on October 4, 2020
bined with an attempt to relieve pain with LA. Although a well performed RCT is always considered the most reliable study design, it has previously been established that register outcome from SweSpine after TDR, as was used in this study, corresponds well with an RCT.23 The strengths of this study include independent data collection and registration of results, and the fact that all the patients were treated at the same clinic.

For patients treated in one or two spinal segments, studies have demonstrated that TDR may at some indications be a better alternative than fusion surgery.7,13 However, patients treated with fusion surgery also reported results that indicated significant benefit from the treatment. Compared to one- or two-segment fusion, three-segment fusion surgery in this patient group seem to produce worse results.10,12 The findings from these previous studies make the results from this study – i.e., results from clinical outcomes when treating patients with multi-level severely symptomatic DDD with a method other than fusion surgery even more interesting. The beneficial results in both groups justify consideration of surgery for patients with CLBP due to DDD when conservative treatment has failed.

Conclusions
We conclude that patients might in selected and severe cases benefit from TDR surgery if they have an accurate diagnosis of prolonged, therapy-resistant CLBP due to DDD in one, two, or even three lumbar spinal segments. The patients included in this study group were specifically chosen, so our results cannot be extrapolated to the general group of patients with prolonged low back pain, and must not lead to an addition of treated segments at surgery.

References

Disclosures

This study had no financial support or funding. The authors report no relevant financial disclosures.

Corresponding Author

Svante Berg, Stockholm Spine Center, Löwenströmska Hospital, SE-19489 Upplands Väsby, Stockholm, Sweden. svante.berg@spinecenter.se.

Published 30 September 2015.
This manuscript is generously published free of charge by ISASS, the International Society for the Advancement of Spine Surgery. Copyright © 2015 ISASS. To see more or order reprints or permissions, see http://ijssurgery.com.