Patient-Reported Outcome Measures after Multilevel Lumbar Total Disc Arthroplasty for the Treatment of Multilevel Degenerative Disc Disease

Scott-Young, Matthew; McEntee, Laurence; Zotti, Mario; Schram, Ben; Furness, James; Rathbone, Evelyne; Hing, Wayne

Published in:
Spine

DOI:
10.1097/BRS.0000000000003201

SPINE An International Journal for the study of the spine, Publish Ahead of Print

DOI: 10.1097/BRS.0000000000003201

Patient reported outcome measures following multi-level lumbar total disc arthroplasty for the treatment of multi-level degenerative disc disease

Matthew Scott-Young (FRACS) [1,2], Laurence McEntee (FRACS) [1,2], Mario Zotti (FRACS)[1,2], Ben Schram (PhD)[2] , James Furness (PhD)[2], Evelyne Rathbone (MSc) [2], Wayne Hing (PhD)[2]

[1] Gold Coast Spine, Gold Coast, QLD AUSTRALIA 4229

[2] Faculty of Health Science & Medicine, Bond University, Gold Coast, QLD AUSTRALIA 4229

Corresponding Author:

Matthew Scott-Young

27 Garden Street, Southport

Gold Coast, Australia, 4215

fax number: +617 5503 1933; telephone number: +617 5528 6477

email: swalter@goldcoastspine.com.au

The device(s)/drug(s) is/are FDA-approved or approved by corresponding national agency for this indication.

No funds were received in support of this work.

Relevant financial activities outside the submitted work: consultancy, royalties.

ABSTRACT

Study Design: Case series

Objective: The aim of this study was to assess the patient reported outcome measures (PROMs) and patient satisfaction of multi-level lumbar total disc arthroplasty (TDA) for symptomatic multilevel degenerative disc disease (MLDDD).

Summary of Background Data: TDA has been shown to be safe and effective for the treatment of symptomatic single level degenerative disc disease (DDD). There is minimal PROMs data on the mid- to long-term outcomes of multi-level TDA constructs.

Methods: Prospectively collected PROMs were analysed from patients receiving multi-level TDA for symptomatic MLDDD. Data was collected pre-operatively and post-operatively at 3, 6 and 12 months, then yearly. PROMs included patient satisfaction, Visual Analogue Score back and leg, Oswestry Disability Index and Roland Morris Disability Questionnaire.

Results: 122 patients (77 males, 45 females) who had pre-operative and at least 24-month follow-up data were included. The average age was 42 ±8.2 years (range 21-61) and mean follow-up 7.8 years (range 2-10). The majority received two level TDA, except two patients (1.6%) who received three level TDA. The two three-level TDA's were at the levels L3-4, L4-5 and L5-S1, while most two levels (n=110, 90.2%) were at L4-5 and L5-S1; the remainder (n=10, 8.2%) being at L3-4 and L4-5. Implants used were Charité (DePuy Spine, Raynham, MA, USA) in 119 patients (240 levels) and InMotion (DePuy Spine, Raynham, MA, USA) in 3 patients (6 levels). Improvement in pain and disability scores were both clinically and statistically significant (p<0.001) and this improvement was sustained in those patients over the course of their follow-up. 92% of patients reported good or excellent satisfaction with treatment at final review.

Conclusions: Multi-level TDA constructs for MLDDD demonstrate favourable and sustained clinical outcomes at mid- to long-term follow-up.

Key Words: artificial disc, arthroplasty, back pain, degenerative disc disease, total disc replacement, lumbar spine, back pain, multi-level disc arthroplasty, bisegmental, motion preservation.

Level of Evidence: 4

Key Points:

- Multi-level lumbar disc arthroplasty surgery appears to be a suitable option for individuals with multi-level symptomatic degenerative disc disease refractory to conservative management, when appropriate diagnosis, patient selection, surgical technique and rehabilitation methods are followed.
- The majority of patients showed favourable clinical outcomes at midterm follow-up.
- 92% of patients reported good to excellent satisfaction over the duration of the study.
- The majority of patients had reduction in disability scores from severe to minimal at latest follow-up.

INTRODUCTION

'Spinal pain' or 'non-specific low back pain' are symptoms influenced by structural, biomechanical, biochemical, medical, psychosocial and compensable factors that can result in dilemmas of diagnosis and management of such complexity that treatment may be rendered ineffective. Distinct from 'low back pain', degenerative disc disease (DDD) causing discogenic pain is a specific diagnosis and, therefore, can be treated non-operatively or, when conservative care fails, operatively.[1] The diagnosis is made from a combination of a clinical history, physical examination, radiological investigations, such as magnetic resonance imaging (MRI), and discriminating provocative discography with post-discography computed tomography scans. [2,3] Other authors have also found electrophysiological studies,[4] MR spectroscopy [5,6] and SPECT scanning[7] adjunctive in supporting a diagnosis.

Basic science studies have confirmed the validity of the model of internal disc disruption (IDD) and the DDD cascade, which can result in discogenic pain from biomechanical, chemical and neural factors .[2] Surgical solutions for multilevel degenerative disc disease (MLDDD) aim to stabilize the painful motion segments by removal of part or all of the sensitised discs. The benefit of the anterior lumbar approach is its ability to allow complete disc resection via a rectus splitting retroperitoneal approach, thus avoiding injury to the dynamic stabilizers. This allows the disc height and lordosis to be restored anatomically through parallel distraction techniques. Static or dynamic stabilization devices can be inserted; however, determining the best device has been associated with contentious debate over several decades with the options for MLDDD from an anterior approach including multi-level anterior lumbar interbody fusion (ALIF), hybrid fusion with lumbar total disc arthroplasty (TDA) surgery [8] and multi-level TDA. [9] The complexity of treating MLDDD with fusion techniques escalates the technical skills required of the surgeon, increases the risk

of adverse events and produces challenges such as pseudarthrosis and adjacent motion segment degeneration (AMSD), rotatory instability and sagittal imbalance.[10, 11]

L-TDA is now an established technology, which has clinical equipoise in reducing pain and improving function and a relatively reduced incidence of AMSD compared with fusion. [7,12 13 14] Conflicting results have been reported for multi-level TDA, with reports of comparatively higher levels of complications, post-operative pain and inferior outcomes to single level TDA. [15 16] However, others have found that complications arising from multi-level TDA are often related to previous surgeries [17] and equivalent [18] or even superior outcomes[19] have been reported when compared with single level TDAs. The technique of performing multilevel TDA requires an anterior lumbar surgery skillset, including training in retro- and trans-peritoneal approaches, adequate skills in vessel mobilization/repair, disc clearance, intervertebral tension balancing and, finally, obtaining both rotatory and coronal stability. Progression to surgical competence in multi-level TDA generally evolves from prior mastery of single-level TDA techniques and where a volume-performance threshold exists.[20]

The aim of this case series was to assess the efficacy of multi-level TDA in the treatment of symptomatic MLDDD through analysis of patient reported outcome measures (PROMs) and patient satisfaction. It is hypothesised that patients who are carefully selected and appropriately treated will achieve favourable outcomes over the mid- to long-term with multi-level TDA.

MATERIALS AND METHODS

Patients with symptomatic MLDDD who underwent multilevel lumbar TDA between April 1999 and January 2009 were identified and their PROMs analyzed. Patients with subsequent revision procedures were excluded from the analysis. This study was approved by the Bond University Human Research Ethics Committee (0000015881).

All participants suffered chronic low back pain (≥12 months) and had been refractory to active non-operative treatment, including physical therapy and rehabilitation programs. Clinical indications for TDA have been demonstrated in the Food and Drug Administration Investigational Device Exemption (FDA IDE) studies that have published 5-year data.[21,22,23] These indications were followed, with the exception being that the primary indication was multi-level rather than single-level symptomatic DDD. A diagnosis of discogenic low back pain, with or without radicular pain, was established through clinical history, examination and diagnostic imaging, which included a combination of standing lumbar radiographs, MRI and provocative discography, with post-discography fine cut CT scan. Because of the high sensitivity and specificity of MRI, it remains an excellent tool for assessing disc morphology, but should be used in conjunction with discography when planning surgical treatment.[24] The general principles outlined by the International Association for the Study of Pain (IASP) were followed when utilizing discography as an investigative tool. Patients whose discographic results that were non-concordant where not offered surgery. Electrophysiological studies (needle electromyography and nerve conduction studies) were performed to confirm the presence or absence of radiculopathy, myopathies, peripheral neuropathies and degenerative neurological conditions. In patients with complex vascular anatomy, a CT angiogram was obtained. Surgery was offered to patients who had a diagnosis of discogenic pain confirmed without contraindications to TDA, who had exhausted non-operative modalities, and where the pathology was having significant effect upon their social, recreational and employment activities.[25]

Contraindications to TDA included ≥grade II facet arthropathy[25], spondylolisthesis, significant scoliosis (>20 degrees), active infection, tumors, severe atherosclerosis or anomalies of the lumbar vessels, pregnancy and diagnostic inconsistency. Obesity and involvement in workers' compensation or other litigation were regarded as relative

contraindications. Surgery was not offered in the presence of overt psychological derangement or maladaptive pain behavior.

Surgery was performed via a midline rectus split with a left or right sided retroperitoneal approach. At each level, in turn, the disc space was prepared for TDA by discectomy, disc space distraction and annuloplasty. After appropriate trialing, the prosthesis was then inserted and position confirmed in the coronal and sagittal planes by fluoroscopy. The annulus was repaired and an anterior longitudinal ligament reconstruction with synthetic ligament performed to reduce segmental coronal or rotatory instability.[26] Protheses used were Charité (DePuy Spine, Raynham, MA, USA) in 119 patients (240 levels) and InMotion (DePuy Spine, Raynham, MA, USA) in 3 patients (6 levels). A peri-operative physiotherapy based rehabilitation program was instituted routinely, which emphasized neural stretching, flexibility, improved dynamic stabilizer exercise tolerance, dynamic muscle strengthening and aerobic fitness.

Participants were required to complete PROMs including Oswestry Disability Index (ODI), Roland-Morris Disability Questionnaire (RMDQ) and self-rated indication of pain using a Visual Analogue Score (VAS) for back (VAS-B) and leg (VAS-L) pain (0-100 point scale) prior to- and at regular intervals post-surgery. Patient satisfaction was also assessed with a four-scale written questionnaire (excellent, good, satisfactory and poor). These outcomes were recorded pre- and post-surgery at 3, 6 and 12 months and yearly thereafter. The PROMs were analyzed by a research team independent of the surgical practice. Radiographic analysis was also completed at each follow-up visit to confirm the movement and alignment of the TDA and exclude complications (eg subsidence, subluxation, heterotopic ossification & AMSD). Routine standing anterior/ posterior lateral and flexion/extension radiographs where taken at 3 months, 6 months and 12 months post operatively. Additional radiographs, CT scans and/or MRI scans were obtained as needed.

Statistical analyses were performed using R Statistical Software Version 3.3.2. The VAS-B, VAS-L, ODI and RMDQ at baseline (prior to surgery) and at multiple time-points from 3 to 120 months after surgery were summarized using medians and IQRs due to skewness of the raw outcomes. The change scores for ODI and RMDQ approximated a normal distribution and are therefore reported using mean differences (95%CI) and *p*-values obtained from paired *t*-tests. However, most of the change from baseline scores for the VAS outcomes also displayed extreme skewness, which was not corrected by transformations. Hence, the median difference (Hodges-Lehmann estimator) and the corresponding 95% confidence intervals are reported, along with the *p*-value obtained from the sign test. Our research group chose to report summary measures (mean or median) according the nature of the data (symmetry of distribution). At times, this has been overlooked in spine research when considering pain reduction, specifically distribution or change in distribution of VAS scores.[27,28] Given the nature of the current data set, the median provided the most appropriate summary statistic, comparable to other studies where a mean is used (assuming a symmetrical distribution), given that both are considered to be the typical value according to the nature of the data.

To account for multiple comparisons of the improvements in the actual scores, the reference *p*-value of 0.05 was adjusted using Bonferroni correction. Graphical representations of median changes in VAS and mean differences in ODI and RMDQ were plotted along with 95% CI and the corresponding minimum clinically important difference (MCID) for each outcome.[29]

RESULTS

In total, 122 patients (77 males, 45 females) operated on between April 1999 and January 2009 were included in this study. The average age was 42 ± 8.2 years (Range 21-61) and mean follow-up was 7.8 years. Two patients (1.6%) received three level TDA, whilst the

remainder of the cohort received two level TDA (98.4%). The two three-level TDAs were at the levels L3-4, L4-5 and L5-S1; the majority of two levels (n=110, 90.2%) were at L4-5 and L5-S1; the remainder (n=10, 8.2%) being at L3-4 and L4-5. A survival of 93.2% (122/131) of multi-level TDA constructs at final follow-up is considered satisfactory and the problems, surgical strategies and subsequent outcomes of the 9 cases of revision and re-operation after multi-level TDA will be discussed in a separate paper.

Table 1 shows the summary statistics for the VAS outcomes for both back and leg pain. At all stages of follow-up, a statistically significant difference from baseline can be seen (p<0.001). By 12 months, the median VAS-B had improved by 88.75% to a score of 9/100.

A total of 24 participants, comprising of 15.5% of the total sample, were lost to follow-up. More than half (n=14; 58.3%) of these patients reported a patient satisfaction score of *Excellent* or *Good* at the last recorded follow-up point which, on average, occurred at 79.7 months (6.6 years). The primary reason for loss to follow-up was non-compliance with completing questionnaires despite reminders. A total of 9 patients underwent index or adjacent segment revision (7.3%).

Table 1 shows the summary statistics for VAS-B and VAS-L. At all stages, a statistically significant difference from baseline can be seen (p<0.001). By 12 months, the median VAS-B scores had improved by 88.8% to a score of 9/100.

Table 2 displays the summary statistics for both the ODI and RMDQ. Statistically significant improvements from baseline are seen throughout the follow-up period (p<0.001) in both outcome measures. The average mean difference from baseline was 31.7 points on the ODI and 12.6 on the RMDQ.

Figure 1AB is a graphical representation of the change scores in both VAS-B and VAS-L over the follow-up period. The reference line in both graphs is the MCID. VAS-B and

VAS-L median score differences can be seen to remain above the MCID consistently during the follow-up period.

A graphical representation of the change from baseline for both the ODI and RMDQ can be seen in Figure 2AB. Again, at all time points, both measures are above the MCID for that specific outcome measure.

DISCUSSION

The aim of this study was to evaluate the PROMs of multilevel TDA for the treatment of symptomatic MLDDD and the efficacy of this technique is validated where appropriate methods of diagnosis, patient selection and technique are followed. Clinically relevant and statistically significant improvements in VAS-B from baseline measures were seen at all time points post-operatively, as all the pre- and post-operative differences were well above the MCID of 12 (p<0.001).[30]

Given long-term single level TDA studies have reported improvement in PROMs and low revision rates,[31, 32] there is increasing attention in the spinal community on the benefits of preserving motion, which facilitates the ability of patients to 'self-centre', thereby theoretically reducing the rate of AMSD. Multilevel lumbar TDA may have benefits over multi-level fusion in obtaining physiological positions required for activities of daily living as suggested by the studies on spinopelvic parameters.[33, 34] Multilevel-fusion in relative kyphosis (particularly in patients with type III and IV spines) can cause extensor muscle fatigue, persistent back and leg symptoms and increase AMSD, while multi-level fusion that increases lordosis (particularly with type I and II spines) can cause difficulty in the elderly with deep squatting positions that may be required for transferring onto chairs or toileting.

This is currently the largest multi-TDA cohort in the literature with the longest follow-up. Improvements in pain are similar to or greater than those reported in other studies. It is difficult to define reasons for this, beyond it being a product of the strict diagnostic criteria, patient selection, consistent surgical techniques and a structured physical therapy program. Back pain (VAS-B) improved post-operatively by 83.3% on average. Tropiano[35] reported an 84% improvement in back pain at final follow-up but, notably, the follow-up time was a mean of 1.4 years. Bertagnoli[36] also demonstrated back pain improvements of 75% at 2 years post-surgery. At an average of 4 years, patients in the study conducted by Trincat[9] had a 60% improvement in their back pain. Other improvements in the order of 56.8%,[10] 40.8%,[15] 37%[18] and 39.6%[37] in back pain have been reported in multi-level TDAs. A study conducted by Yue[38] showed similar pre-operative VAS-B scores (VAS-B of 77.3); however, the improvements were to 31.3 at 2 years (59.5%) and 28.7 at 5 years (62.9%).[38] In the current study, at the same time frames, back pain had improved by 89.4% and 86.9% respectively.

The median pre-operative score for leg pain in the study was 54.5 VAS-L, which improved to an average of 2.6 (95.2% improvement). Although the percentage improvement was higher than for VAS-B, due to relatively lower baseline values for VAS-L, the absolute improvements were smaller (as in the Trincat et al. study) [9] and percentage reduction may be a better measure when comparing actual pain reduction in the two VAS outcomes. A large proportion of patients had little to no leg pain at baseline and were expected to experience little or no change at follow-up (approximately 20% of the patients scored from 0-20 VAS-L at baseline) and this is reflected in the lower change from baseline scores and 95% confidence intervals (Figure 1). However, these changes were still statistically and clinically significant (above MCID of 16) until 36 months.[39] Single level ALIF and TDA studies demonstrate that VAS-L can be improved and proven radiculopathy treated, with the

extrapolation from those results suggesting multi-level TDA can also affect VAS-L significantly and proportionately to VAS-B.

The ODI improved on average 31.7 points in this study, above the 18.8 points that is considered to be substantial benefit in taking patients from severe to minimal disability.[40, 41] Over the period observed, the ODI score for the cohort dropped from 48 to a mean post-operative value of 11 (77% improvement), which is greater than what has been previously reported. Comparative reports include improvements of 50%[9], 43.2%[15], 31.2%[37], 56.8%[19] and 58%[38]. The ODI improvements in the studies by both Tropiano[35] and Bertagnoli[36] are similar to the improvements seen here with 67% and 75% improvement respectively. However, these studies only involved 24 months and 14 months follow-up; thus, the effect of PROMs decay could not be assessed. Few other long-term studies have utilised the RMDQ as an outcome measure and, therefore, while the data presented here is favourable (78-94% improvement from baseline depending on timepoint chosen) they are included for comparative purposes with future studies that may also utilise this outcome measure.

Not all patients with symptomatic multiple level discogenic pain are suitable for multi-level TDA. Considerations for this include spinopelvic parameters, operative level, facet arthritis, bone density, the presence or absence of radiculopathy, and other comorbidities, as discussed previously. The evidence suggests that, when patients are appropriately selected, effective and durable results can be obtained.

This study is a prospective case series that supports the safety and efficacy of multi-level TDA with a clearly defined protocol and explicit inclusion and exclusion criteria. Patients were enrolled consecutively and the follow-up of clinical outcomes occurs on an annual basis indefinitely. In addition, the follow-up rate is high; thus, the validity of the treatment effect and the study protocol is robust. However, it is acknowledged that a case series does not have

a control group and can be prone to bias, thus limiting its generalisability to larger populations and surgeons at other institutions. An unconstrained TDA implantation at multiple levels is technically demanding whereas newer, more rotationally constrained, one piece prostheses may prove to be 'more forgiving' and thereby result in relatively better clinical and radiological outcomes for multi-level TDA.[26]

The authors recognise the importance of coronal and sagittal balance. The ability to fully understand global alignment and the types of spine via EOS™ has only been available recently, whereas the patient cohort in this study received treatment between 1999 and 2009. Since the advent of EOS™ imaging, much emphasis is placed on analysing the relationship between the pelvis and the spinopelvic parameters. When considering multi-level TDA, surgeons need to consider spinopelvic parameters and the type of spine just as much as the type of prosthesis (constrained or unconstrained).

Future studies should compare long-term clinical outcomes of single level TDA, multi-level TDA and hybrid construct surgery for the treatment of DDD.

CONCLUSION

This study suggests that multi-level TDA for MLDDD is associated with favourable and sustained clinical outcomes for the majority of patients. Provided diagnosis, patient selection, surgeon technique and rehabilitation are adequate, multi-level lumbar TDA is an effective management technique for individuals identified as being affected by more than one degenerative disc. To our knowledge, this represents the largest cohort and longest follow-up of multi-level lumbar TDA constructs in the literature.

REFERENCES

1. Bogduk N. Management of chronic low back pain. *Med J Aust.* 2004;180(2):79-83.
2. Sehgal N, Fortin JD. Internal disc disruption and low back pain. *Pain Physician.* 2000;3(2):143-157.
3. Lee YC, Zotti MG, Osti OL. Operative Management of Lumbar Degenerative Disc Disease. *Asian Spine J.* 2016;10(4):801-819.
4. Cho SC, Ferrante MA, Levin KH, Harmon RL, So YT. Utility of electrodiagnostic testing in evaluating patients with lumbosacral radiculopathy: An evidence-based review. *Muscle Nerve.* 2010;42(2):276-282.
5. Brayda-Bruno M, Tibiletti M, Ito K, et al. Advances in the diagnosis of degenerated lumbar discs and their possible clinical application. *Eur Spine J.* 2014;23 Suppl 3:S315-323.
6. Zuo J, Joseph GB, Li X, et al. In-vivo Intervertebral Disc Characterization using Magnetic Resonance Spectroscopy and T(1ρ) Imaging: Association with Discography and Oswestry Disability Index and SF-36. *Spine (Phila Pa 1976).* 2012;37(3):214-221.
7. Malham GM, Parker RM, Ballok ZE, Goss B, Diwan AD, Uribe JS. Bone scans are reliable for the identification of lumbar disk and facet pathology. *Global Spine J.* 2015;5(1):23-30.
8. Scott-Young M, McEntee L, Schram B, Rathbone E, Hing W, Nielsen D. The Concurrent Use of Lumbar Total Disc Arthroplasty and Anterior Lumbar Interbody Fusion: The Lumbar Hybrid Procedure for the Treatment of Multi-Level Symptomatic Degenerative Disc Disease- A Prospective Study. *Spine (Phila Pa 1976).* 2017:1.

9. Trincat S, Edgard-Rosa G, Geneste G, Marnay T. Two-level lumbar total disc replacement: Functional outcomes and segmental motion after 4 years. *Orthop Traumatol Surg Res.* 2015;101(1):17-21.

10. Röllinghoff M, Schlüter-Brust K, Groos D, et al. Mid-range outcomes in 64 consecutive cases of multilevel fusion for degenerative diseases of the lumbar spine. *Orthop Rev (Pavia).* 2010;2(1):e3.

11. Campbell PG, Yadla S, Malone J, et al. Complications related to instrumentation in spine surgery: a prospective analysis. *Neurosurg Focus.* 2011;31(4):E10.

12. Zigler J, Sachs B, Rashbaum R, Ohnmeiss D. Two level total disc replacement with Prodisc: results and comparison to one level cases. *Spine J.* 2005;5:S4-S5.

13. David T. Long-term results of one-level lumbar arthroplasty: minimum 10-year follow-up of the CHARITE artificial disc in 106 patients. *Spine (Phila Pa 1976).* 2007;32(6):661-666.

14. Ren C, Song Y, Liu L, Xue Y. Adjacent segment degeneration and disease after lumbar fusion compared with motion-preserving procedures: a meta-analysis. *Eur J Orthop Surg Traumatol.* 2014;24 Suppl 1:S245-253.

15. Siepe CJ, Mayer HM, Heinz-Leisenheimer M, Korge A. Total lumbar disc replacement: different results for different levels. *Spine (Phila Pa 1976).* 2007;32(7):782.

16. Siepe CJ, Heider F, Wiechert K, Hitzl W, Ishak B, Mayer MH. Mid- to long-term results of total lumbar disc replacement: a prospective analysis with 5- to 10-year follow-up. *Spine J.* 2014;14(8):1417.

17. Clavel P, Ungureanu G, Catalá I, Montes G, Málaga X, Ríos M. Health-related quality of life in patients undergoing lumbar total disc replacement: A comparison with the general population. *Clin Neurol Neurosurg.* 2017;160:119-124.

18. Hannibal M, Thomas DJ, Low J, Hsu KY, Zucherman J. ProDisc-L total disc replacement: a comparison of 1-level versus 2-level arthroplasty patients with a minimum 2-year follow-up. *Spine (Phila Pa 1976).* 2007;32(21):2322.

19. Schätz C, Ritter-Lang K, Gössel L, Dreßler N. Comparison of Single-Level and Multiple-Level Outcomes of Total Disc Arthroplasty: 24-Month Results. *Int J Spine Surg.* 2015;9:14.

20. Regan, John J, McAfee, Paul C, Blumenthal, Scott L, et al. Evaluation of Surgical Volume and the Early Experience With Lumbar Total Disc Replacement as Part of the Investigational Device Exemption Study of the Charité Artificial Disc. *Spine* 2006; 31(19):2270-2276

21. Zigler J, Gornet MF, Ferko N, et al. Comparison of Lumbar Total Disc Replacement With Surgical Spinal Fusion for the Treatment of Single-Level Degenerative Disc Disease: A Meta-Analysis of 5-Year Outcomes From Randomized Controlled Trials. *Global Spine J.* 2018 Jun;8(4):413-423.

22. Zigler JE, Delamarter RB. Five-year results of the prospective, randomized, multicenter, Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential arthrodesis for the treatment of single-level degenerative disc disease. *J Neurosurg Spine.* 2012 Dec;17(6):493-501

23. Guyer RD, McAfee PC, Banco RJ, et al. Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: five-year follow-up. *Spine J.* 2009 May;9(5):374-86

24. Manchikanti L, Soin A, Benyamin RM, et al. An Update of the Systematic Appraisal of the Accuracy and Utility of Discography in Chronic Spinal Pain. *Pain Physician.* 2018 Mar;21(2):91-110

25. Weishaupt D, Zanetti M, Boos N, Hodler J. MR imaging and CT in osteoarthritis of the lumbar facet joints. *Skeletal Radiol.* 1999;28(4):215-219.

26. Cunningham BW, Gordon JD, Dmitriev AE, Hu N, McAfee PC. Biomechanical evaluation of total disc replacement arthroplasty: an in vitro human cadaveric model. *Spine (Phila Pa 1976).* 2003;28(20):S110-117.

27. Geisler FH. Surgical treatment for discogenic low-back pain: lumbar arthroplasty results in superior pain reduction and disability level improvement compared with lumbar fusion. *SAS J.* 2007;1(1):12-19.

28. Ohnmeiss D. Numerical rating scales. In: Rihn JA V, AR, Albert TJ, ed. *Defining the value of spine care.* Jaypee Brothers Medical Publishers Ltd; 2012:77.

29. 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 Outcome Study questionnaire Short Form 36, and Pain Scales. The Spine Journal 2008;8(6):968-974

30. Copay AG, Eyberg B, Chung AS, et al.Minimum Clinically Important Difference: Current Trends in the Orthopaedic Literature, Part II: Lower Extremity: A Systematic Review. *JBJS Rev*. 2018 Sep;6(9):e2

31. Plais N, Thevenot X, Cogniet A, Rigal J, Le Huec JC. Maverick total disc arthroplasty performs well at 10 years follow-up: a prospective study with HRQL and balance analysis. *Eur Spine J.* 2018;27(3):720-727.

32. Lu SB, Hai Y, Kong C, et al. An 11-year minimum follow-up of the Charite III lumbar disc replacement for the treatment of symptomatic degenerative disc disease. *Eur Spine J.* 2015;24(9):2056-2064.

33. Roussouly P, Nnadi C. Sagittal plane deformity: an overview of interpretation and management. *Eur Spine J.* 2010;19(11):1824-1836.

34. Sparrey CJ, Bailey JF, Safaee M, et al. Etiology of lumbar lordosis and its pathophysiology: a review of the evolution of lumbar lordosis, and the mechanics and biology of lumbar degeneration. *Neurosurg Focus.* 2014;36(5):E1.

35. Tropiano P, Huang RC, Girardi FP, Marnay T. Lumbar Disc Replacement: Preliminary Results with ProDisc II After a Minimum Follow-Up Period of 1 Year. *Spine (Phila Pa 1976).* 2003;28:362-368.

36. Bertagnoli R, Yue JJ, Shah RV, et al. The treatment of disabling single-level lumbar discogenic low back pain with total disc arthroplasty utilizing the Prodisc prosthesis: a prospective study with 2-year minimum follow-up. *Spine (Phila Pa 1976).* 2005;30(19):2230.

37. Silvestre M, Bakaloudis G, Lolli F, Vommaro F, Parisini P. Two-level total lumbar disc replacement. *Eur Spine J.* 2009;18(Supplement 1):64-70.

38. Yue J, Zhang K, Bai H, et al. A Comparison of Patients Who Have Undergone 1-Level Versus 2-Level Prodisc Arthroplasty A Prospective Study With Minimum of 5-Year Follow-up. *Spine (Phila Pa 1976).* 2013;38(14):1194-1198.

39. Copay AG, Chung AS, Eyberg B, et al. Minimum Clinically Important Difference: Current Trends in the Orthopaedic Literature, Part I: Upper Extremity: A Systematic Review. *JBJS Rev.* 2018 Sep;6(9):e1.

40. Aunoble S, Meyrat R, Al Sawad Y, Tournier C, Leijssen P, Le Huec J. Hybrid construct for two levels disc disease in lumbar spine. *Eur Spine J.* 2010;19(2):290-296.

41. Glassman SD, Copay AG, Berven S, Polly DW, Subach BR, Carreon LY. Defining substantial clinical benefit following lumbar spine arthrodesis. . *J Bone Joint Surg Am.* 2010;90(9):1839-1847.

Figure 1. Median difference between pre- and post- surgery over time, and 95% confidence intervals for VAS back (a) and leg pain (b) scores in 122 patients.

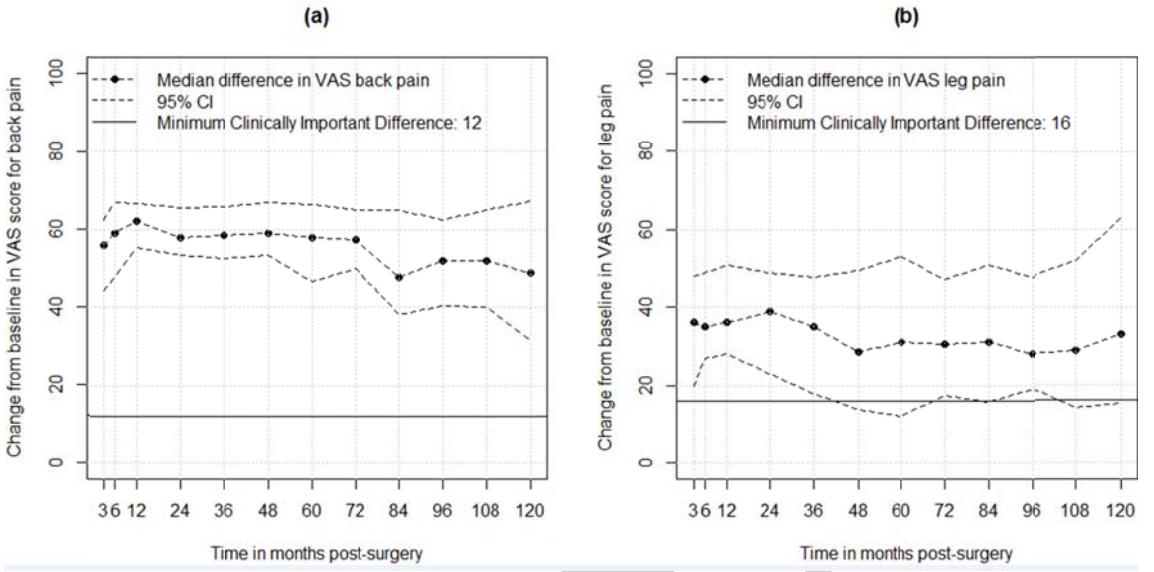

Figure 2. Mean difference between pre- and post- surgery over time, and 95% confidence intervals for ODI (a) and RMDQ (b) disability scores in 122 patients

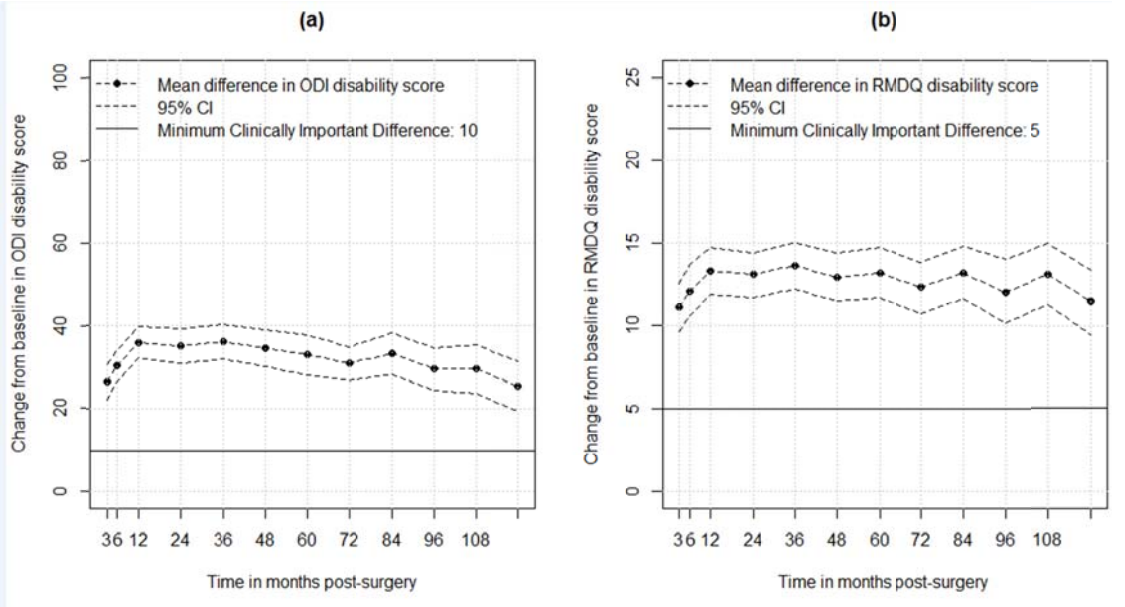

Table 1. VAS back and leg pain outcomes over time in 122 patients

VAS[1] outcome				Change from baseline			
Time (months) post-surgery	***n***	**Median**	**IQR**	***n***	**Median difference[2]**	**95% CI**	***p*-value[3]**
Back pain							
0 baseline	107	80.0	65.5 – 91.0				
3	85	18.0	6.0 – 33.0	79	56.0	44.3 to 62.5	<0.001
6	93	13.0	3.0 – 26.0	85	59.0	48.0 to 67.0	<0.001
12	101	9.0	1.0 – 26.0	88	62.0	55.3 to 66.7	<0.001
24	96	8.5	0.8 – 25.8	82	58.0	53.6 to 65.4	<0.001
36	96	10.5	0.0 – 31.0	82	58.5	52.6 to 65.8	<0.001
48	86	12.5	0.0 – 28.8	73	59.0	53.6 to 66.8	<0.001
60	78	10.5	0.0 – 35.5	68	58.0	46.8 to 66.2	<0.001
72	89	14.0	2.0 – 32.0	76	57.5	50.0 to 65.0	<0.001
84	73	11.0	2.0 – 49.0	62	48.0	38.0 to 65.0	<0.001
96	71	16.0	2.5 – 48.5	61	52.0	40.3 to 62.4	<0.001
108	61	14.0	2.0 – 53.0	51	52.0	40.0 to 65.0	<0.001
120	58	19.0	4.3 – 53.5	47	49.0	31.8 to 67.2	<0.001
Leg pain							
0 baseline	90	54.5	19.3 – 81.0				
3	76	3.0	0.0 – 18.0	69	36.0	19.9 to 48.1	<0.001
6	86	3.0	0.0 – 20.5	74	35.0	27.0 to 48.9	<0.001
12	93	2.0	0.0 – 11.0	73	36.0	28.0 to 51.0	<0.001
24	94	2.0	0.0 – 9.0	67	39.0	23.0 to 49.0	<0.001
36	94	3.0	0.0 – 17.0	67	35.0	18.0 to 48.0	<0.001
48	85	1.0	0.0 – 13.0	60	28.5	13.8 to 49.5	<0.001
60	75	2.0	0.0 – 15.0	53	31.0	12.0 to 53.1	<0.001
72	88	4.0	0.0 – 22.0	62	30.5	17.5 to 47.2	<0.001
84	71	3.0	0.0 – 25.0	50	31.0	15.6 to 50.9	<0.001

96	72	3.0	0.0 – 25.8	51	28.0	19.0 to 48.0	<0.001
108	62	3.0	0.0 – 21.5	42	29.0	14.1 to 51.9	<0.001
120	58	2.0	0.0 – 16.0	37	33.0	15.1 to 62.9	<0.001

[1]The Visual Analogue Scale (VAS) is scored on a 0 (no pain) to 100 (worst imaginable pain) scale.
[2]The median difference is the Hodges-Lehmann estimator. A positive median difference indicates an improvement or reduction in pain score from baseline (prior to surgery). [3]The *p*-value is the result of the sign test. Significance is achieved when $p<0.004$ using Bonferroni correction, as applied to multiple comparisons within each type of pain outcome. All differences from baseline were statistically significant.

Table 2. ODI and RMDQ disability outcomes over time in 122 patients

Disability outcome				Change from baseline			
Time (months) post-surgery	***n***	**Median**	**IQR**	***n***	**Mean difference[2]**	**95% CI**	***p*-value[3]**
ODI[1]							
0 baseline	122	48.0	34.5 – 60.0				
3	81	16.0	10.0 – 28.0	81	26.4	22.1 to 30.7	<0.001
6	93	12.0	4.0 – 26.0	93	30.3	26.6 to 34.0	<0.001
12	101	8.0	0.0 – 20.0	101	35.9	32.1 to 39.6	<0.001
24	96	8.0	2.0 – 22.5	96	35.0	30.9 to 39.1	<0.001
36	96	7.0	0.0 – 24.0	96	36.1	32.0 to 40.2	<0.001
48	86	10.0	0.5 – 25.5	86	34.6	30.2 to 38.9	<0.001
60	78	9.0	0.5 – 24.0	78	32.9	28.1 to 37.6	<0.001
72	90	10.0	2.0 – 24.0	90	30.9	26.8 to 34.9	<0.001
84	73	10.0	0.0 – 24.0	73	33.3	28.2 to 38.3	<0.001
96	72	11.0	2.0 – 28.8	72	29.5	24.4 to 34.6	<0.001
108	62	12.0	0.0 – 26.0	62	29.5	23.7 to 35.4	<0.001
120	58	15.0	2.0 – 28.3	58	25.4	19.4 to 31.3	<0.001
RMDQ[4]							
0 baseline	101	18.0	13.0 – 20.0				
3	76	4.0	1.0 – 9.0	73	11.1	9.6 to 12.5	<0.001
6	86	2.0	0.0 – 6.0	77	12.1	10.6 to 13.7	<0.001
12	94	1.0	0.0 – 5.0	82	13.3	11.9 to 14.7	<0.001
24	94	1.0	0.0 – 5.0	77	13.1	11.7 to 14.4	<0.001
36	95	1.0	0.0 – 5.0	77	13.6	12.2 to 15.0	<0.001
48	86	1.0	0.0 – 5.8	71	12.9	11.5 to 14.4	<0.001
60	78	1.0	0.0 – 6.0	63	13.2	11.7 to 14.7	<0.001
72	90	1.0	0.0 – 6.0	72	12.3	10.7 to 13.8	<0.001
84	73	1.0	0.0 – 7.0	60	13.2	11.6 to 14.8	<0.001

96	72	1.0	0.0 – 10.3	58	12.0	10.1 to 14.0	<0.001
108	62	1.0	0.0 – 6.0	51	13.1	11.3 to 15.0	<0.001
120	58	1.5	0.0 – 6.0	51	11.5	9.5 to 13.4	<0.001

[1]The Oswestry Disability Index (ODI) is scored on a 0 (none) to 100 (worst) disability. [2]A positive mean difference indicates an improvement or reduction in disability index from baseline (prior to surgery).
[3]The *p*-value is the result of the paired *t*-test. Significance is achieved when $p<0.004$ using Bonferroni correction, as applied to multiple comparisons within each disability measure. All differences from baseline were statistically significant. [4]The Roland-Morris Disability Questionnaire (RMDQ) is scored on a 0 (none) to 24 (worst) disability.