

Hellum C, Johnsen LG, et al. Surgery with disc prosthesis versus rehabilitation in patients with low back pain and degenerative disc: two year follow-up of randomised study. BMJ 2011;342:d2786.

Design; Randomized clinical trial

Population/sample size/setting:

- 172 patients (91 women, 81 men, mean age 41) treated for degenerative lumbar discs at 5 centers in Norway
- Eligibility criteria were age 25-55 with at least 1 year of low back pain, six months of physical therapy or chiropractic treatment without satisfactory resolution of symptoms, a score of at least 30 on the Oswestry disability index, and disc degeneration confined to L4-L5 and L5-S1
 - o Degeneration defined by at least 40% reduction in disc height, Modic changes type I or II or both, grade 3 or 4 signal intensity in the disc
 - o Discs independently classified by 2 observers with disagreement resolved by a third observer when necessary
- Exclusion criteria included nerve root involvement (but not facet joint degeneration), spinal stenosis, disc protrusion, spondylolysis, arthritis, history of fracture of L1-S1, osteoporosis, or generalized chronic pain (in Appendix 1)

Main outcome measures:

- Randomization was to either rehabilitation (n=86) or disc replacement surgery (n=86)
- Rehabilitation was based on a model previously used by Brox et al in studies of spinal fusion surgery, and involved a team of physical therapists and specialists in multidisciplinary treatment
 - o Treatment was outpatient done in groups lasting about 60 hours over 5 weeks
 - o Program consisted of group lectures and individual discussions of topics such as anatomy and physiology of the back, normal reactions, coping strategies, training of abdominal and multifidus muscles, and challenging patients' thoughts about limitations of physical activities
 - o Follow-up consultations were done 6 weeks, 3 months, 6 months, and 1 year after the intervention
- Surgery consisted of replacement of the degenerative disc with the ProDisc II by surgeons who had performed at least 6 disc prostheses
 - o Patients were not referred for postoperative physical therapy, but could be referred for PT at 6 weeks after surgery if required
- Primary outcome was the Oswestry at baseline, 6 weeks, 3 months, 6 months, 1 year, and 2 years
 - o At the 2 year follow-up, 2 independent observers evaluated patients on a 5-test back performance scale and on the Prolo scale, which consists of functional and economic parts which are added together into a single score from 2 (worst) to 10 (best)

- Patients were instructed not to tell these examiners which intervention they had received, and had tape placed over the abdominal wall to conceal surgical scars
- Several secondary outcomes were measured, including pain, EQ-5D health status scores, and SF-36 scores; a net back-to-work rate was also calculated as a secondary outcome
- Among patients randomized to surgery, 33% underwent surgery at 2 disc levels
- Both groups improved their Oswestry scores between baseline and 2 years: the surgery group by a mean of 20.8 and the rehabilitation group by an average of 12.4; the mean difference between groups was 8.4 in favor of surgery in this primary, intention-to-treat analysis
 - A per protocol analysis was done as a secondary measure; the treatment group difference here was 8.1 in favor of surgery
 - Another secondary analysis compared the proportion of patients in the two groups with improvements of at least 15 points on the Oswestry; this favored the surgery group (70%) over the rehab group (47%)
- Several secondary outcomes (back pain, SF-36 physical function) also favored surgery, but return to work was not different between groups
- In the 2-year examiner-blinded back performance scale, there was no group difference, but the Prolo sum score favored surgery by a mean difference of 0.9 points on the scale of 2 to 10
- Adverse outcomes did occur with surgery; the most serious was when the disc inlay was dislodged, requiring a second operation in which the left iliac artery was damaged, leading to amputation of the left leg
 - 2 patients had an additional fusion and 2 had partial resection of the spinous process
- There were significant dropouts in both groups at the 24 month follow-up: 20% in the surgical group and 24% in the rehab arm
 - 5 patients crossed over from rehab to surgery, but no patient crossed over from surgery to rehab

Authors' conclusions:

- For the main outcome, surgery had an 8.4 point advantage over rehab on the Oswestry score; this is less than the difference of 10 points that the study was designed to detect
- There is no general consensus on what magnitude of change in the Oswestry index is clinically important; there is a need for such a consensus
- Although there were numerous withdrawals, a questionnaire was sent to patients who dropped out; the 9 surgery dropouts had a reduction in Oswestry of 30.2 points; the 6 rehab dropouts had a reduction of 11.8 points; the 11 patients who withdrew without any treatment had no change
 - The difference in compliance between groups could lead to an underestimate of the true effect of surgery

- Both groups could have placebo effects: the surgery group from the operation, and the rehab group from frequent contact with physical therapists and other practitioners
- There was considerable improvement in the rehab group, suggesting that it is reasonable to consider a rehab program before surgery in patients with degenerative disc disease of the lumbar spine

Comments:

- A few points are not clear; the text states that no patient crossed from surgery to rehabilitation, but Table B of Appendix 3 , which lists patients lost to follow-up after randomization, shows two patients randomized to surgery who received rehab instead
- Also not clear is which follow-up values were adjusted for baseline status; at the end of the Methods section, the authors state that “significantly different baseline scores were not adjusted for in the longitudinal model. Each outcome variable was adjusted for the baseline values of the variable.” Since the outcomes were measured in a longitudinal fashion, the nature of the adjustments is not clear
- The primary outcome and the unplanned analyses handled missing data differently; the primary outcome used the last observation carried forward (LOCF) , while the unplanned analyses used mixed models
 - o Mixed models, in contrast to LOCF, use all the available data, and, since most current software programs have mixed models, LOCF seems not to be the best choice, since it is an imputation method and may introduce bias which could be partially controlled with other available methods of analysis (Mallinckrodt et al 2003)
 - o The direction of such potential bias is difficult to estimate, but is suspected to increase reported group differences
 - o Because there is a fairly large amount of missing follow-up data for both groups, there is considerable uncertainty regarding the treatment effect
- The mean treatment effect in the unplanned mixed model (6.9 Oswestry points) was also smaller than the planned analysis, and it is reasonable to conclude that neither analysis is pointing to a large advantage of surgery over rehabilitation
- The authors’ conclusions that a trial of rehabilitation should be considered before disc replacement is reasonable

Assessment: Adequate for some evidence that disc replacement has a small advantage over a multidisciplinary rehabilitation program for lumbar degenerative disc disease, and that rehabilitation is a reasonable first choice of intervention for lumbar DDD.

Reference:

Mallinckrodt CH, Sanger TM et al. Assessing and Interpreting Treatment Effects in

Longitudinal Clinical Trials with Missing Data. *Biol Psychiatry* 2003;53:754–760.