Van den Eerenbeemt KD, Ostelo RW, et al. Total disc replacement surgery for symptomatic degenerative lumbar disc disease: a systematic review of the literature. Eur Spine J 2010;19: 1262-1280.

Design: Systematic review of randomized clinical trials

PICOS:

- **P**atient population: Adults with symptomatic lumbar degenerative disc disease
- **I**ntervention: Lumbar total disc prostheses (Charité, ProDisc, Maverick, Acroflex)
- **C**ontrol intervention: Lumbar interbody fusion
- **O**utcomes: Three separate questions were formulated:
    - What is the course of DDD complaints and/or symptoms following total disc replacement surgery? (clinical course)
    - What is the effectiveness of total disc replacement compared to other treatments? (effectiveness)
    - What is the safety of total disc replacement surgery? (safety)
- **S**tudy types: Randomized clinical trials for efficacy compared to lumbar fusion, prospective cohort studies for studies of safety and for clinical course of complaints and symptoms following surgery

Study search and selection:

- Databases were MEDLINE, EMBASE, the Cochrane Library, searched from 1973 through October 2008
- Two authors independently assessed articles for inclusion eligibility and for risk of bias
- Bias of randomized trials was assessed using Cochrane risk of bias tool, which emphasizes randomization, concealment of allocation, blinding, intention-to-treat analysis, selective outcome reporting, co-interventions, baseline similarity of groups, and timing of assessment of outcome

Results:

- 1962 references were identified in the literature search; for clinical course, 21 articles (16 studies) were relevant; for effectiveness, 16 articles (only 3 studies) were relevant; for safety, 7 overview articles were relevant
- For clinical course of complaints following surgery, there were 6 studies of the Charité, 8 of the ProDisc, 1 for Maverick, and 1 for Acroflex
    - For Charité, there was data for up to 11.3 years of follow-up, reporting good or excellent clinical results in 90% of patients at 11.3 years; other cohort studies reported improvements on VAS, Oswestry, and patient satisfaction with studies having follow-up times ranging from 6 months to 6.9 years
    - For ProDiscI, 55 patients were available for followup after an average of 8.7 years, with 82.6% reporting satisfaction or complete satisfaction with their results

    - ProDiscII (ProDisc-L), the second generation ProDisc, had follow-up range from 3 months to 2 years, with a majority (79-100%) satisfaction and VAS improvements from 40-62 points and Oswestry improvements from 21% to 48%
    - For Maverick at 2 years follow-up, low back pain improved by an average of 44 points with Oswestry improvements of 20.7%
    - For Acroflex, there was mechanical device failure, and no randomized trial had been carried out
- For effectiveness compared to lumbar fusion, there were three randomized trials
    - The Charité trial, designed as a non-inferiority trial, enrolled 304 patients, and had data for follow-up times of 2 years and 5 years comparing it with anterior interbody fusion with a BAK cage
        - Pain intensity and functional improvements did not differ significantly at 2 years between Charité and BAK cage fusion, and overall success was non-inferior for the Charité (57.1%) and the fusion (46.5%); patient satisfaction was better for the Charité group (73.7%) than for fusion (53.1%)
        - At 5 years of follow-up, the Charité was non-inferior to fusion on a composite score of clinical success (57.8% vs. 51.2%); however, there was high loss to follow-up at 5 years, and the authors rated this as conferring a high risk of bias on the results
    - The ProDisc-L study, which compared the artificial disc with circumferential fusion (anterior lumber fusion done with femoral ring allograft plus posterolateral fusion done with iliac crest autograft plus pedicle screws) in 236 randomized patients, was considered to have a high risk of bias
        - For the composite success score, which combines 10 outcomes required by the FDA, the 2 year success rate was greater for ProDisc (54.3%) compared to fusion (40.8%); however, there were no group differences on mean functional improvement or on mean pain scores
    - The Flexicore trial reported on only 76 patients out of 401 who were randomized; the authors did try to draw conclusions due to the high risk of bias
- For safety of disc replacement, there was a very wide range of reported complication rates, with wide discrepancies between rates reported in the literature and rates reported to the FDA
    - For example, the Charité published complication rates were 29.1% for the artificial disc and 50.2% for fusion; however, in the FDA report, the rates were 181.9% and 189.6% respectively
    - For the ProDisc, the rates for the artificial disc and fusion were 7.3% and 6.3%; in the FDA report, the rates were 255.5% and 270.7%
    - Reoperation rates at the index level ranged from 3.7% to 11.4% for disc replacement and between 5.4% to 26.1% for fusion groups

Authors' conclusions:

- The quality of reporting on outcomes was often poor, preventing adequate interpretation of the effects of disc replacement for the low back
- There is low quality evidence (based on one study with low risk of bias) that there are no clinically important differences between Charité and BAK cage fusion at 2 years
- There is very low quality evidence (based on one study with high risk of bias) that there are no clinically important differences between ProDisc and circumferential fusion at 2 years, but the results were contradictory
- Nothing can be said about the effectiveness of Flexicore compared to lumbar fusion
- The ProDisc trial compared the artificial disc with a fusion procedure using autograft; since the Oswestry and pain VAS scores may be influenced by graft site pain, the comparison of ProDisc with fusion may have been made against a form of fusion likely to have less favorable outcomes than with allograft
- Because there is still uncertainty that fusion is better than conservative treatment of low back pain, there is no certainty that the artificial discs are any better than conservative treatment; there are no studies comparing the two
- The definition of success did not include opioid use
- There are concerns with long term success of artificial discs; there may be wear debris leading to osteolysis; there may be migration and extrusion leading to serious complications such as vascular damage
    - Because the artificial discs are implanted in fairly young patients, the survival of the disc prosthesis needs to be about 40 years
- It is recommended that disc replacement be done only in the setting of prospective scientific studies

Comments:

- Although a meta-analysis was not appropriate, the general conclusion that the artificial disc should be used only in the setting of clinical studies is well-argued
- The ProDisc study (Zigler 2007)is rated as having a high risk of bias, but it is not clear that this is the case
    - Table 2, which rates the methodological quality of the studies, does not credit Zigler with allocation concealment or with an acceptable dropout rate
    - However, Zigler did report that the randomization was held by the sponsor and disclosed to the site only after individual patient enrollment
    - Zigler also reported having followup of 98.2% at 24 months, with complete data suitable for calculating success as 91% in the ProDisc group and 88.5% in the fusion group
    - If these two criteria are credited to Zigler, the study changes from a high risk to a low risk of bias

    - The ProDisc study is described as having “contradictory” results; the only discrepancy is that while the overall success rate advantage of ProDisc was statistically significant, the comparison on Oswestry and VAS scores were not statistically significant
    - This has more to do with the variance in the scores than with any real contradiction between the different results of the study; it should not be construed to imply that the study results actually contradicted one another
    - However, the comparison of ProDisc with a fusion procedure in which the control patients had iliac crest autograft for the posterolateral fusion does create a separate risk of a bias in favor of the ProDisc
- It is highly relevant that fusion has evidence of superiority to conservative treatment only when there is spondylolisthesis with instability; since spondylolisthesis was an exclusionary criterion for these arthroplasty studies, the comparison applies only to fusion without instability

Assessment: Systematic review supports good evidence that the Charité disc is non-inferior to allograft fusion with the BAK cage for single level disease, and some evidence that the ProDisc is non-inferior to circumferential fusion with iliac crest autograft for single level disease, but that there is no evidence that the artificial disc is superior to nonoperative treatment

Reference:

Zigler J, Delamarter R, 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 2007;32(11): 1155-1162.