



Surgery for Lumbar Disc Herniation and Spinal Stenosis: Surgeon's Clinical Judgment versus Actual Improvement in Patient Reported Outcomes

Leah Carreon^{1,2}, Casper Fries Pedersen^{1,2}, Peter Döring^{3,4}, Mikkel Andersen^{1,2}, Søren Eiskjær^{3,4}.

Department of Orthopedic Surgery, Lillebaelt Hospital, Middelfart, Denmark¹. Department of Orthopedic Surgery, University of Southern Denmark, Odense Denmark². Department of Orthopedic Surgery, The Spine Research Group, Aalborg University Hospital, Aalborg, Denmark³. Department of Clinical Medicine, Aalborg University Hospital, Aalborg Denmark⁴.

Background

Several predictive models have been developed to predict outcomes in patients choosing to have surgery for lumbar disc herniation (LDH) based on patient demographics and baseline patient-reported outcome measures (PROMs). However, there has been no formal evaluation of clinical outcomes based solely on a surgeon's clinical judgment based on patient gestalt.

Methods

Patients with LDH referred to a tertiary spine clinic for surgical evaluation were evaluated. As the option for surgical treatment of LDH is preference based, the surgeon discussed the risks and benefits of surgical versus non-surgical treatment with the patient and the patient's family as part of a shared decision-making process. Once the patient opted for surgery, the surgeon was asked to provide an opinion on the outcomes of surgery for the specific patient with a series of seven

faces denoting the Global Perceived Effect (GPE) as "Very bad" (GPE1), "Bad" (GPE2), "Fairly bad"(GPE3), "No change"(GPE4), "Fairly good" (GPE5), "Good" (GPE6) and "Very Good" (GPE7). Standard demographic and surgical data was collected as well as PROMs (ODI, EQ5D, SF36, VAS Back and Leg Pain) prior to surgery and one year after surgery. Patients were then stratified based on the surgeon's clinical judgement and change in one year outcome measures were compared.

Results

Of 153 subjects enrolled, 110 (72%) had one year data available with 0 GP1, 1 GPE2, 4 GPE3, 5 GPE4, 36 GPE5, 48 GPE6 and 16 GPE 7. Only patients in GPE3 to GPE7 were included in the analysis. There was no difference in demographic or surgical parameters among the different GPE groups. Improvements in ODI, EQ5D and SF36PCS were greatest in the GPE7 followed by the GPE6 then by the GPE5. GPE5 and GPE4 had similar improvements, while GPE3 had less improvement than GPE4. Improvement in VAS back and leg pain was similar the GPE7, GPE6 and GPE5 group, with less improvement seen in the GPE4 and GPE3 groups.

Conclusion

The results highlight the relevance of a surgeon's Global Perceived Effect (GPE) based on patient gestalt in predicting clinical improvements after LDH surgery. Especially, the GPE appears to be effective in assessing improvements in ODI, EQ5D, and SF-36. However, its reliability is less evident in predicting improvements in VAS back and leg pain. This indicates that a surgeon's clinical judgment is valuable in certain domains of patient outcomes, but may not fully capture the nuances of pain-related improvements. Recognizing the significance of artificial intelligence (AI) in predictive modeling, it is crucial not to underestimate the value of clinical intuition. Combining AI with clinical intuition of experienced clinicians — "hybrid intelligence" — might have the potential to improve the accuracy of outcome predictions in lumbar disc herniation surgeries.

