# Measurement of outcome in lumbar spine surgery
*Validity and interpretability of frequently used outcome measures in the Swespine register*

## Akademisk avhandling

Som för avläggande av medicine doktorsexamen vid Sahlgrenska akademin, Göteborgs universitet kommer att offentligen försvaras i R-aulan, Sahlgrenska universitetssjukhuset, Mölndal, fredagen den 24 april, 2020, klockan 09.00 av

### Catharina Parai, leg. läkare

Fakultetsopponent:
**Teja Lund MD PhD**
Helsinki University, Finland

## Avhandlingen baseras på följande delarbeten

I.   Catharina Parai, Olle Hägg, Bengt Lind, Helena Brisby. The value of patient global assessment in lumbar spine surgery, an evaluation based on more than 90,000 patients.
     Eur Spine J (2018) 27:554–563.

II.  Catharina Parai, Olle Hägg, Bengt Lind, Helena Brisby. Follow-up of degenerative lumbar spine surgery – PROMS stabilize after 1 year: an equivalence study based on Swespine data.
     Eur Spine J. 2019 Sep;28(9):2187-2197.

III. Catharina Parai, Olle Hägg, Bengt Lind, Helena Brisby. ISSLS prize in clinical science 2020: the reliability and interpretability of score change in lumbar spine research.
     Eur Spine J. 2019 Nov 23.

IV.  Catharina Parai, Olle Hägg, Carl Willers, Bengt Lind, Helena Brisby. Characteristics and predicted outcome of patients lost to follow-up after degenerative lumbar spine surgery.
     Submitted.

**SAHLGRENSKA AKADEMIN**
**INSTITUTIONEN FÖR KLINISKA VETENSKAPER**

GÖTEBORGS
UNIVERSITET

# Measurement of outcome in lumbar spine surgery
*Validity and interpretability of frequently used outcome measures in the Swespine register*

**Catharina Parai,** Dept. of Orthopaedics, Inst. of clinical sciences, Sahlgrenska Academy, University of Gothenburg, Sweden, 2020

## Abstract

BACKGROUND. The purpose of elective lumbar spine surgery is mainly to reduce pain and to improve physical function and quality of life. The quality and results of the interventions are monitored in the Swedish spine register, Swespine. The large quantities of data offer unique opportunities to improve quality of care, decrease costs and enable benchmarking. For register-based data to be useful, however, the quality must be high, the variables must be carefully selected to ensure relevant data collection, and the logistics of data collection should be workable.

AIM. The overall purpose of the thesis was to find ways to simplify the assessment of patient-reported outcome without a loss in scientific credibility.

STUDY POPULATION. The main study population was obtained from the Swespine register and included patients operated in the lumbar spine in the period 1998–2015 for either disc herniation (n: 30,102), spinal stenosis (n: 50,194), isthmic spondylolysis/spondylolisthesis, or degenerative disc disorder. The two latter diagnoses were treated as a single entity (n: 13,836). A test-retest study was performed on 182 individuals obtained from two spine-care hospitals (2017–2019). Analyses on non-respondents were computed using Swespine data from 2008–2012 that were linked to hospital data, Statistics Sweden, the National Patient Register, and the Social Insurance Agency (n: 21,961).

METHODS. The usefulness of the single-item retrospective outcome measure GA as an overall PROM (Patient-Reported Outcome Measure) was tested in correlation analyses with symptom-specific (i.e VAS), disease-specific (i.e ODI), and generic PROMs (i.e EQ-5D, SF-36). The capability of GA as a discriminator of treatment success was explored in ROC curve analyses. The level of treatment success was defined for each of the Swespine PROMs with different lumbar conditions. The proportion that achieved these scores one year after the operation was compared with the proportion at two years. PROM retest reliability was tested on a symptom stable population. The SDC at the 95% confidence level was computed. The retrospective measurements were tested using weighted kappa. Regression analyses were conducted to identify variables associated with non-response. The output was used to predict outcomes for patients with the characteristics of the non-respondent population.

RESULTS AND CONCLUSIONS. High correlations were seen between GA and VAS, and also the ODI, indicating that GA can replace these tools in effectiveness studies. The correlations were better for final scores than for changes in score, indicating present-state bias and/or recall bias. Correlations with EQ-5D were lower, indicating that GA works less well as a discriminator of quality of life. The ROC curve analyses support the use of GA as a reference criterion in the interpretation of VAS and ODI scores. A tough cut-off signifying a considerable improvement is encouraged. The change in a PROM score needed to achieve treatment success (i.e. the MIC value) varied somewhat between the degenerative conditions tested; thus, the ODI MICs were 14–22 points, the $VAS_{BACK}$ MICs were 20–29 mm; the $VAS_{LEG}$ MICs were 23–39 mm; and the EQ-5D MICs were 0.10–0.18. The proportion of patients who reached these levels at the one-year follow-up was similar to the proportion at the two-year follow-up. Thus, collection of PROM data in Swespine on the latter occasion is not necessary. The retest reliability for the PROMs tested was similar or lower than previously reported. In general, the SDC estimates exceeded the MIC values, thereby complicating the interpretation of score changes, as the PROMs were not sensitive enough to detect score changes considered important. Being lost to follow-up was associated with male sex, younger age, smoking, lower disposable income, and lower education, and with being born outside the EU. Non-respondents were predicted to have a somewhat worse outcome than respondents.

**Keywords:** spine register, disc herniation, spinal stenosis, degenerative disc disorder, patient-reported outcome measure, Global Assessment, minimal important change, smallest detectable change, retest reliability, non-response to follow-up, attrition, measurement of change.