**CV-BIOGRAPHY MARGRETH GROTLE**

Prof Margreth GROTLE is appointed professor of Health Science at the Oslo Metropolitan University (100% FTE since January 2013) and at Department of Research and Innovation, Division of Clinical Neuroscience, Oslo University Hospital (20% FTE). She is an Honorary Professor at the Primary Care Musculoskeletal Research Centre, Keele University, UK, and at the Centre for Pain Impact, Neuroscience Research Australia (NeuRA), Sydney. Since 2013 she has been head of the Master Programme in Physiotherapy at OsloMet and established a research group on Musculoskeletal Health (the MUSK Health research group) in 2015 (www.muskhealth.com). The research group has further developed and in 2021 Grotle was appointed to lead a Centre for Musculoskeletal health at OsloMet (https://uni.oslomet.no/cim/). CIM is a collaboration between musculoskeletal health researchers and computer scientist and aims to develop innovative health interventions by applying Artificial Intelligence. Grotle has more than 20 years of experience in musculoskeletal research, in particular spinal disorders, but also osteoarthritis and rheumatic diseases.

## Abstract Margreth Grotle/ 231108 Swespine Årsmöte Stockholm


### *ABSTRACT*

**Background and objective:** Lumber disc herniation surgery can effectively reduce pain and disability. However, approximately 30-45% of the patients receiving lumbar surgery report to not have a clinical important improvement one year after surgery. One way to improve the rates of successful surgery outcomes is to improve the selection of patients to surgery. However, most prediction models developed so far have not sufficient accuracy to be implemented in clinical practice. One of our major questions in the AID-Spine project, in which we use data from the Norwegian Registry for Spine Surgery (NorSpine) linked to public health and welfare registers, is whether and to what extent Machine Learning (ML) methods can provide more accurate prediction models for successful/non-successful outcomes after surgery due to disc herniation or spinal stenosis.

**Step 1; defining success/non-success:** First, we identified the optimal definition of treatment success at 12 months in the NorSpine data (2007 to 2022) for patients undergoing surgery due to disc herniation or spinal stenosis. The three major outcomes - Oswestry Disability Index (ODI), Numeric Rating Scale (NRS) for back pain and NRS leg pain – were operationalized with study-specific calculations of the cut-offs for success, using a dichotomized anchor response at the 7-point Global Perceived Effect scale with the cut-off for success/non-success set between patients responding “much improved” versus “slightly improved”. The thresholds were arrived at using the anchor-based predictive modeling method, adjusted for the proportion of improved patients (Terluin B et al *J Clin Epidemiol*. 2015;68(12):1388-96, Terluin B et al *J Clin Epidemiol*. 2017;83:90-100). The cut-off scores for patients undergoing disc herniation surgery were: ODI=22 points, NRS back pain=2 points, and NRS leg pain=4 points improvement from baseline to 12 months. The cut-off scores for

patients undergoing spinal stenosis surgery were lower for the ODI and NRS leg pain: ODI=14 points, NRS back pain=2 points, and NRS leg pain=3 points improvement from baseline to 12 months. By using these cut-offs, treatment non-success was experienced by 33% (ODI), 27% (NRS back pain), and 31% (NRS leg pain) in disc herniation surgery. For spinal stenosis treatment non-success was experienced by 45% (ODI), 31% (NRS back pain), and 41% (NRS leg pain).

**Step 2; ML model development and validations, example from the disc herniation sample:** Analysis included 22,707 surgical cases (ODI model) for disc herniation. The ML models were trained for model development and internal-external cross-validation applied over geographical regions to validate the models. Model performance was assessed through discrimination (*C*-statistic) and calibration (calibration slope and intercept). In the internal-external cross-validation, the selected machine learning models showed consistent discrimination and calibration across all five health regions of Norway. The *C*-statistic ranged from 0.81 to 0.84 (pooled random-effects meta-analysis estimate 0.82 [95% CI 0.81 to 0.84]) for the ODI model. Calibration slopes (point estimates 0.94 to 1.03; pooled estimate 0.99 [95% CI 0.93 to 1.06]) and calibration intercepts (point estimates -0.05 to 0.11; pooled estimate 0.01 [95% CI -0.07 to 0.10]) were also consistent across regions. Next step, which is ongoing work, is to externally validate these ML models in SweSpine and DanSpine data.

**Conclusion: T**he ML prediction models showed good discrimination and calibration to distinguish between success and non-success in disability and pain 12 months after lumbar disc herniation surgery. The accuracy was poorer in the prediction models for spinal stenosis surgery. The models have potential to inform patients and clinicians about individual prognosis and aid in surgical decision-making.