

The National Neurosurgery Quality and Outcomes Database (N2QOD) Pilot: Interim Analysis of Registry Feasibility, Data Integrity, and the Safety and Effectiveness of Surgical Spine Care

A. Asher; C. Watridge; P. Mummaneni; J.Knightly; G. Oetting; N. Theodore; O. Gottfried; S. Khairi; M. Schmidt; M. Boakye; S. Kalkanis; D.Rabin; T. Ryken; G. Balturshot; J. Chaddock; D. Fassett; R. Reeder; C. Miller; T. Briggs; T. Grahm; B. Guthrie; J. F. Harrington; C. Shaffrey; E. Elowitz; K. Foley; S. Glassman; M. McGirt

Introduction: Quality measurement registry platforms lie at the center of all emerging evidence-driven reform models and will be used to inform all decision makers in healthcare delivery. In March of 2012, the National Neurosurgery Quality and Outcomes Database (N2QOD) Lumbar Pilot was launched with the aim to develop a valid and accurate platform to measure the safety and effectiveness of spine care in everyday neurosurgical and orthopedic practice.

Methods: A nationwide, web-based, prospective, longitudinal one-year outcomes registry (lumbar spine module) was introduced into neurosurgery and orthopedic spine practices. Baseline, peri-operative, 3- and 12-month medical record and patient-reported outcomes data was entered prospectively into the REDCap™ web-based portal. Figure 1. For quality control and validation, automated missing data reports were communicated to sites weekly, targeted site education was instituted when needed, and 10% of sites audited.

Results: Registry characteristics at 11 mos: 33 practice-groups; 175 surgeons; 47 hospitals in 24 states. 3550 patients were enrolled (1857 at 3 month f/u). Cumulative missing data was <2% (of > 400,000 independent variables). Three-month loss to follow-up: 9%. Site audits demonstrated data extraction accurate in all (56/56) cases; patient inclusion accurate in 97%. Surgical morbidity given in Table 1. Significant improvement in pain, disability, and quality of life was reported by patients three months post-operatively, however, marked variation in outcome was observed at the individual patient level(Figures 2&3). Risk-adjusted expected benchmarks for safety and effectiveness varied significantly between sites based on their unique patient populations(Figures 4&5).

Conclusions: A prospective, nationwide, outcomes registry is feasible with high degree of data integrity. Early results suggest that the safety and effectiveness of surgical spine care cannot be accurately measured in the real world setting without robust risk adjustment of disease-specific clinical variables. The N2QOD registry platform generates accurate evidence for informed decision support for patients, surgeons, and policy makers.

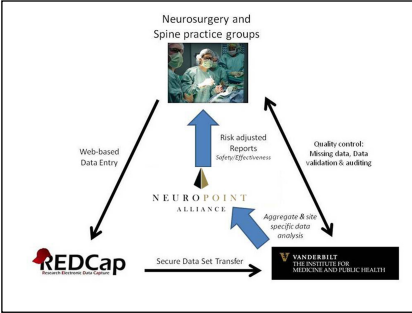


Figure 1. Schematic of N2QOD data collection and reporting. Data are entered in standardized format into N2QOD using the REDCap™ web-based portal. The Vanderbilt Institute of Medicine and Public Health (VIMPH) performs data analysis and modeling and prepares site-specific reports of risk-adjusted safety and effectiveness. Quality control is a two-way process between sites and VIMPH data coordinators. VIMPH provides weekly missing data reports and performs data validation auditing, while site coordinators provide feedback to VIMPH on the needs of process improvement via weekly coordinator teleconferences and a monthly practice-based learning network.

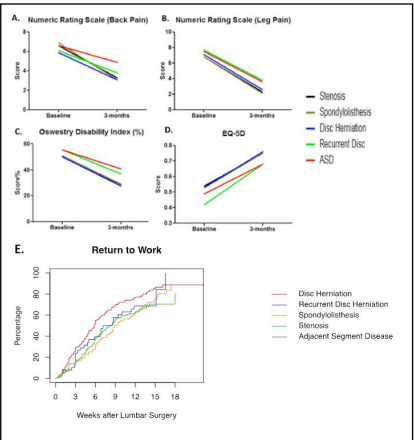


Figure 2. Real-world effectiveness of surgical spine care for five common lumbar diagnoses three months postoperatively. A significant improvement in A) back pain, B) leg pain, C) functional disability, and D) quality adjusted life year gain were reported by patients undergoing surgery in a real-world setting for all five disease states. E) Overall, 75% of patients planning to return to word (RTW) returned to their pre-operative job by three months after surgery. Median time to return to work varied as a function of diagnosis. Twelve month outcomes are pending.

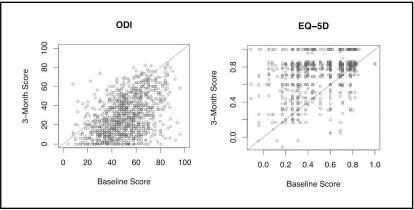


Figure 3. Variation in surgical effectiveness in real world care. While lumbar surgery was highly effective at improving disability and quality of life at the group mean level, marked variation in early (3mo) outcome was observed at the patient level despite standardized diagnosis criteria. Identifying the clinical, radiographic, and surgical characteristics of the patients not responding to surgical intervention will allow for improvement in patient selection and surgical care delivery, ultimately facilitating practice based learning and quality improvement. *Patients not improving or worsening 3months after lumbar surgery are represented by plots on or above the bisecting diagonal line for ODI and by plots on or below the bisecting diagonal line for EQ5D.*

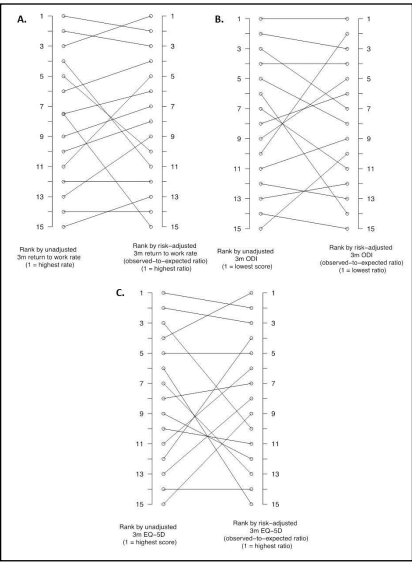


Figure 4. The fifteen highest patient accruing N2QOD centers ranked by Quality of care: Percentage of patients returning to work (A), 3mo postoperative low-back physical disability (ODI) (B), and 3mo postoperative quality of life (EQ-5D: QALY) (C). Ranking early effectiveness of care (Unadjusted) not accounting for the patient and disease factors influencing outcome was highly inaccurate when compared to risk-adjusted ranking of effectiveness of care (Observed/Expected). These clinical risk factors are not available in current administrative data reporting systems. *Factors outside of the provider's control influencing effectiveness of surgery included age, race, gender, BMI, smoking history, diabetes, CAD, osteoporosis, anxiety disorder, depression, education level, major surgery history, symptom type, symptom duration, worker's compensation claim, occupation type, ASA grade, # surgical levels, arthrodesis, lumbar diagnosis, and baseline level of pain, disability, and health state.*

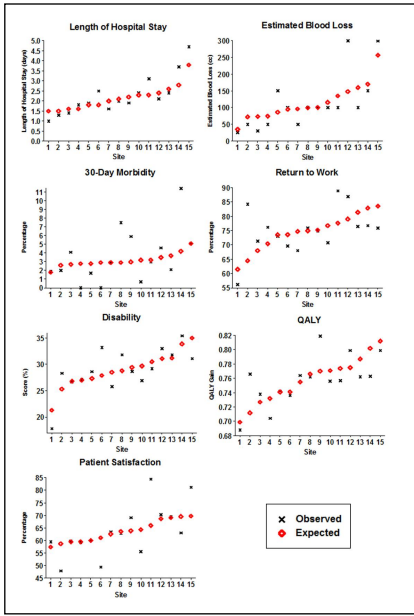


Figure 5. Risk-adjusted expected benchmarks of quality (safety and effectiveness) of elective lumbar surgery varies widely across spine centers based on each center's unique patient and disease populations. Expected norms (benchmarks) represent the mean outcomes of the multi-center center network risk-adjusted for each site's unique patient mix. Many sites performed above their quality benchmark while others performed below their benchmark. Registries failing to collect robust clinical variables that confound outcome are unable to determine accurate norms of surgical spine outcomes and are unable to assess the comparative effectiveness or relative value of treatments accurately. Site number does not represent the same center between graphs. ** Variables included in the regression model included age, race, gender, BMI, smoking history, diabetes, CAD, osteoporosis, anxiety disorder, depression, education level, major surgery history, symptom type, symptom duration, worker's compensation claim, occupation type, ASA grade, # surgical levels, arthrodesis, lumbar diagnosis, and baseline level of pain, disability, and health state.*