Clinical outcomes research in spine surgery: what are appropriate follow-up times?

Presented at the 2018 AANS/CNS Joint Section on Disorders of the Spine and Peripheral Nerves

Oliver G. S. Ayling, MD, MSc,1 Tamir Ailon, MD, MPH, FRCSC,1 Greg McIntosh, MSc,2 Alex Soroceanu, MD, MPH, FRCSC,3 Hamilton Hall, MD, FRCSC,4 Andrew Nataraj, MD, MSc, FRCSC,5 Christopher S. Bailey, MD, FRCSC,6 Sean Christie, MD, FRCSC,7 Alexandra Stratton, MD, MSc, FRCSC,8 Henry Ahn, MD, FRCSC,4 Michael Johnson, MD, FRCSC,9 Jerome Paquet, MD, FRCSC,10 Kenneth Thomas, MD, MSc, FRCSC,3 Neil Manson, MD, FRCSC,11 Y. Raja Rampersaud, MD, FRCSC,4 and Charles G. Fisher, MD, MHSc, FRCSC1

1Department of Surgery, Vancouver General Hospital/University of British Columbia, Vancouver, British Columbia; 2Canadian Spine Society, Markdale, Ontario; 3Department of Surgery, University of Calgary, Alberta; 4Department of Surgery, University of Toronto, Ontario; 5Department of Surgery, University of Alberta, Edmonton, Alberta; 6Division of Orthopaedics, University of Western Ontario, London, Ontario; 7Department of Surgery, Dalhousie University, Halifax, Nova Scotia; 8Department of Surgery, University of Ottawa, Ontario; 9Departments of Orthopedics and Neurosurgery, University of Manitoba, Winnipeg, Manitoba; 10Department of Surgery, Laval University, Quebec City, Quebec; and 11Department of Surgery, Canada East Spine Centre, Saint John, New Brunswick, Canada

OBJECTIVE There has been a generic dictum in spine and musculoskeletal clinical research that a minimum 2-year follow-up is necessary for patient-reported outcomes (PROs) to adequately assess the therapeutic effect of surgery; however, the rationale for this duration is not evidence based. The purpose of this study was to determine the follow-up time necessary to ensure that the effectiveness of a lumbar surgical intervention is adequately captured for three lumbar pathologies and three common PROs.

METHODS Using the different PROs of pain, physical function, and mental quality of life from the Canadian Spine Outcomes and Research Network (CSORN) prospective database, the authors assessed the time course to the recovery plateau following lumbar spine surgery for lumbar disc herniation, degenerative spondylolisthesis, and spinal stenosis. One-way ANOVA with post hoc testing was used to compare scores on the following standardized PRO measures at baseline and 3, 12, and 24 months postoperatively: Disability Scale (DS), visual analog scale (VAS) for leg and back pain, and SF-12 Mental Component Summary (MCS) and Physical Component Summary (PCS).

RESULTS Significant differences for all spine pathologies and specific PROs were found with one-way ANOVA (p < 0.0001). The time to plateaued recovery after surgery for lumbar disc herniation (661 patients), lumbar stenosis (913 patients), and lumbar spondylolisthesis (563 patients) followed the same course for the following PRO measures: VAS for back and leg pain, 3 months; DS, 12 months; PCS, 12 months; and MCS, 3 months. Beyond these time points, no further significant improvements in PROs were seen. Patients with degenerative spondylolisthesis or spinal stenosis who had undergone fusion surgery plateaued at 12 months on the DS and PCS, compared to 3 months in those who had not undergone fusion.

CONCLUSIONS Specific health dimensions follow distinctly different recovery plateaus, indicating that a 2-year postoperative follow-up is not required for all PROs to accurately assess the treatment effect of lumbar spinal surgery. Ultimately, the clinical research question should dictate the follow-up time and the outcome measure utilized; however, there is now evidence to guide the specific duration of follow-up for pain, physical function, and mental quality of life dimensions.

https://thejns.org/doi/abs/10.3171/2018.8.SPINE18715

KEYWORDS spine surgery; patient-reported outcomes; follow-up duration; registry
Historically, a minimum 2-year follow-up after spine surgery has been the gold-standard duration to assess the treatment effect as measured by patient-reported outcomes (PROs);\textsuperscript{4,6,10,13} however, the rationale for the 2-year follow-up is not evidence based. Long-term follow-up data for various lumbar pathologies, including lumbar disc herniation (LDH), lumbar spinal stenosis (LSS), and lumbar degenerative spondylolisthesis (LDS), suggest that outcomes after lumbar surgery often reach a plateau before 2 years-highlighted by studies that have demonstrated stability of outcomes beyond 2 years.\textsuperscript{17} Although a recovery plateau can be reached before 2 years, studies have also demonstrated that PROs assessed as early as 3 months are not reliable indicators of long-term outcome.\textsuperscript{2,10}

Despite these studies of both short- and long-term outcomes, the duration of follow-up required to sufficiently assess the time to a recovery plateau after lumbar spine surgery has not been directly investigated. Theoretically, the appropriate follow-up time should depend on the research question being asked, the health dimension being assessed, as well as the intervention and pathology being studied. Given the increasing implementation and importance of clinical registries and the high cost of obtaining long-term follow-up data in clinical studies,\textsuperscript{10} it is vital to ensure that the duration of follow-up is tailored to the pathology and the outcome dimension to ensure that treatments are effectively studied and reported. Precise determination of the required follow-up duration also enables appropriate patient counseling and monitoring of their recovery.

The purpose of this study was to determine the follow-up time necessary to ensure that the effectiveness of lumbar surgical intervention is adequately captured for three lumbar spine pathologies and three commonly captured PROs. Specifically, patients with LDH, LSS, and LDS were evaluated for 1) pain (visual analog scale [VAS] for leg and back pain), 2) physical function (Disability Scale [DS] and SF-12 Physical Component Summary [PCS]), and 3) mental health–related quality of life (QOL; SF-12 Mental Component Summary [MCS]).

**Methods**

**Study Population**

Consecutive patients who were over 18 years old, undergoing elective lumbar spinal surgery, entered into the Canadian Spine Outcomes and Research Network (CSORN) prospective database between June 2013 and March 2016, and eligible for 2-year follow-up were analyzed in a multicenter ambispective review. Patients undergoing surgical treatment of LSS, LDH, or LDS were eligible for inclusion. Patients were excluded if the surgery was performed for adjacent-segment disease, infection, trauma, or malignancy.

CSORN comprises a group of more than 50 neurosurgical and orthopedic spine surgeons from 18 tertiary care academic and nonacademic hospitals across Canada that prospectively collects data on patients with spinal conditions. The network’s database serves as a national registry created to answer research questions and to facilitate the implementation of best practices.

A national database research coordinator audits data quality and performance and sends reports to each contributing hospital site coordinator on a quarterly basis. Reports track data completion and follow-up rates to facilitate internal data validation at each site. A national privacy and security framework was created for CSORN and includes a governance structure, standard operating procedures, training processes, physical and technical security, and privacy impact assessments. This model ensures the privacy and security of personal health information. Written informed consent is obtained from all participating patients. Patient identification is anonymized to ensure that patients in the network cannot be individually identified. All participating sites obtained Research Ethics Board (REB) approval prior to any data collection. Decisions regarding data collection, storage, and analysis are independent of any particular company or commercial interest.

**Data Collection**

Local research coordinators enrolled the patients at each site. Collected patient characteristics included age, sex, body mass index, smoking status, preoperative education status, work status at baseline assessment, pending insurance claims, marital status, time with lumbar condition, type of surgery (discectomy, decompression, fusion), and medical comorbidities (depression, diabetes, hypertension, osteoarthritis, systemic vascular disease, cerebrovascular disease, cardiac disease, and peripheral vascular disease). Each site has a dedicated research coordinator responsible for conducting follow-ups. Each site coordinator has access to the data input component of the registry in order to directly enter PRO responses for each patient. For logistical and practical reasons, each follow-up time has a “window of opportunity” to contact patients, calculated from the date of surgery: for the 3-month follow-up, the window is ± 15 days; for the 12-month follow-up, ± 30 days; and for the 24-month follow-up, ± 45 days. Patients who are not contacted or who do not respond within these windows are categorized as “lost to follow-up.” Patients are given the option of completing the follow-ups in person on paper, by phone, by regular mail, or via a secure online portal.

**Outcome Measures**

Research coordinators, unaware of the study hypothesis, collected PRO measures at baseline and 3, 12, and 24 months postoperatively. The primary outcome measure was the time point at which each postoperative PRO reached a plateau. “Recovery plateau” was defined as the first follow-up point beyond which no further significant difference in a given PRO was identified. PROs were assessed with the DS, which is a modified Oswestry Low Back Pain Disability Questionnaire;\textsuperscript{2} the VAS for back and leg pain;\textsuperscript{3} and the SF-12 PCS and MCS.\textsuperscript{15}

**Statistical Analysis**

Baseline and demographic factors were tabulated, presented as frequency distributions or means ± standard deviations, and analyzed with the Fisher exact tests, Student t-test, or Mann-Whitney U-tests, as appropriate. One-
way ANOVA with Tukey’s Honest Significant Difference (HSD) post hoc test was employed to determine the time point at which recovery reached a plateau postoperatively.

We determined that a plateaued recovery had been reached when the mean of at least two consecutive postoperative time points was not significantly different as determined by the Tukey HSD post hoc test (e.g., if PROs at 12 and 24 months were not significantly different from one another but were both significantly different from the 3-month score, then the plateau was reached at 12 months). Postoperative PRO scores had to be significantly different from preoperative scores to be considered as reaching a plateau. A p value < 0.05 was used to determine significance.

Results

Patient Demographic Data

A total of 661 patients with a diagnosis of LDH, 913 with LSS, and 563 with LDS were identified from the CSORN database. Demographic and clinical data are summarized in Table 1. The follow-up rates at each time point were as follows: 3 months: LDH 89.1%, LSS 91.5%, LDS 91.7%; 1 year: LDH 75.6%, LSS 83.6%, LDS 85.3%; and 2 years: LDH 74.3%, LSS 80.2%, LDS 83.7% (Supplemental Fig. 1 and Supplemental Tables 1–5).

Lumbar Disc Herniation

The PRO for physical function after surgery for LDH plateaued by 1 year. Changes in the DS score were significant over time as determined by ANOVA (F[3,2226] = 287.3, p < 0.0001), and all postoperative DS scores were significantly lower than baseline on the Tukey post hoc test (p < 0.01; Fig. 1A). The recovery plateau according to the DS was reached at 1 year after surgery (baseline 50.9 ± 14.9 vs 1 year 23.6 ± 19.5). The PCS scores were also significantly different over time on the ANOVA (F[3,2175] = 252.3, p < 0.0001; Fig. 1D), with scores at all postoperative time points significantly higher than baseline (p < 0.001). Plateaued recovery according to the PCS was reached at 1 year postoperatively (baseline 32.0 ± 8.0 vs 1 year 44.8 ± 9.4).

The PRO for pain after surgery for LDH reached a plateau by 3 months. VAS leg pain scores were significantly different over time as determined by ANOVA (F[3,2203] = 391.6, p < 0.0001; Fig. 2A), with scores at all postoperative time points significantly lower than that at baseline (p < 0.01) and reaching a plateau at 3 months postoperatively (baseline 7.5 ± 1.9 vs 3 months 3.5 ± 2.9). VAS back pain scores followed the same trend similar to that for the leg scores (ANOVA, F[3,2203] = 269.9, p < 0.0001), with significantly lower postoperative scores (p < 0.01) that plateaued at 3 months (VAS back baseline 6.7 ± 2.3 vs 3 months 3.6 ± 2.4; Fig. 2D).

Assessment of mental QOL with the MCS after surgery for LDH revealed significantly different scores on the ANOVA (F[3,2175] = 56.7, p < 0.0001), and all postoperative scores were significantly higher than that at baseline (p < 0.01). The recovery plateau was reached at 3 months postoperatively (MCS baseline 45.7 ± 8.5 vs 3 months 50.0 ± 7.8; Fig. 3A). Specific mean values for each time point are presented in Supplemental Table 2.

Lumbar Spinal Stenosis

Physical function scores plateaued at 12 months after surgery for LSS. The DS scores were significantly different over time (ANOVA, F[3,3209] = 83.68, p < 0.0001), and all postoperative DS scores were significantly different from baseline (p < 0.01). The recovery plateau as measured by the DS was reached at 12 months after surgery (baseline 47.8 ± 15.5 vs 3 months 27.9 ± 19.9; Fig. 1C). PCS scores were also significantly different over time on the ANOVA (F[3,3165] = 488.7, p < 0.0001). The recovery plateau on the PCS was reached at 12 months after surgery (baseline 31.9 ± 8.3 vs 12 months 39.0 ± 8.9, p < 0.01; Fig. 1F).

After surgery for LSS, VAS leg pain scores differed significantly over time (ANOVA, F[3,3165] = 488.7, p < 0.0001), with all postoperative scores lower than baseline (p < 0.01; Fig. 2C). Recovery plateau was reached at 3 months (baseline 7.5 ± 2.0 vs 3 months 3.2 ± 2.8). VAS back pain scores followed the same trend as the leg scores (ANOVA, F[3,3165] = 432.4, p < 0.0001) and plateaued at

### Table 1. Summary of patient and operative characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>LDH Group</th>
<th>LSS Group</th>
<th>LDS Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>661</td>
<td>913</td>
<td>563</td>
</tr>
<tr>
<td>Age (± SD)</td>
<td>45.27 (12.84)</td>
<td>64.27 (11.69)</td>
<td>60.26 (13.61)</td>
</tr>
<tr>
<td>Male (%)</td>
<td>352 (53.25)</td>
<td>525 (57.50)</td>
<td>221 (39.25)</td>
</tr>
<tr>
<td>Female (%)</td>
<td>309 (46.75)</td>
<td>388 (42.50)</td>
<td>342 (60.75)</td>
</tr>
<tr>
<td>BMI (± SD)</td>
<td>27.65 (8.01)</td>
<td>28.7 (8.35)</td>
<td>27.37 (8.13)</td>
</tr>
<tr>
<td>Nicotine (%)</td>
<td>223 (33.34)</td>
<td>222 (24.32)</td>
<td>147 (26.11)</td>
</tr>
<tr>
<td>Some college education (%)</td>
<td>221 (33.43)</td>
<td>268 (29.35)</td>
<td>212 (37.66)</td>
</tr>
<tr>
<td>Currently working (%)</td>
<td>259 (39.18)</td>
<td>200 (21.91)</td>
<td>173 (30.73)</td>
</tr>
<tr>
<td>Insurance claim (%)</td>
<td>94 (14.22)</td>
<td>73 (7.99)</td>
<td>56 (9.95)</td>
</tr>
<tr>
<td>Married (%)</td>
<td>481 (72.77)</td>
<td>675 (73.93)</td>
<td>388 (68.92)</td>
</tr>
<tr>
<td>&gt;2 yrs w/ condition (%)</td>
<td>127 (19.21)</td>
<td>319 (34.94)</td>
<td>322 (57.19)</td>
</tr>
<tr>
<td>Surgery (%)</td>
<td>325 (49.17)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Discectomy</td>
<td>20 (3.03)</td>
<td>334 (36.58)</td>
<td>72 (12.79)</td>
</tr>
<tr>
<td>Decompression &amp; fusion</td>
<td>169 (25.57)</td>
<td>53 (5.81)</td>
<td>23 (4.08)</td>
</tr>
<tr>
<td>Discectomy &amp; decompression</td>
<td>75 (11.35)</td>
<td>349 (38.23)</td>
<td>339 (60.21)</td>
</tr>
<tr>
<td>Decompression &amp; fusion</td>
<td>29 (4.39)</td>
<td>138 (15.12)</td>
<td>105 (18.65)</td>
</tr>
<tr>
<td>Fusion</td>
<td>0</td>
<td>0</td>
<td>12 (2.13)</td>
</tr>
<tr>
<td>Depression</td>
<td>84 (12.71)</td>
<td>107 (11.72)</td>
<td>103 (18.29)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>39 (5.9)</td>
<td>181 (19.82)</td>
<td>72 (12.79)</td>
</tr>
<tr>
<td>HTN</td>
<td>77 (11.65)</td>
<td>267 (29.24)</td>
<td>199 (35.35)</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>80 (12.10)</td>
<td>318 (34.83)</td>
<td>204 (36.23)</td>
</tr>
<tr>
<td>Systemic vascular disease</td>
<td>13 (1.97)</td>
<td>66 (7.23)</td>
<td>51 (9.06)</td>
</tr>
</tbody>
</table>

HTN = hypertension; SD= standard deviation.

Ayling et al.
3 months (VAS back baseline 7.0 ± 2.3 vs 3 months 3.4 ± 2.4, p < 0.01; Fig. 2F).

Mental QOL scores according to the MCS were also significantly different over time (ANOVA, F[3,3096] = 44.3, p < 0.0001). Plateau on the MCS was reached at 3 months after surgery (baseline 48.4 ± 8.4 vs 3 months 51.7 ± 7.5, p < 0.01; Fig. 3C).

**Lumbar Degenerative Spondylolisthesis**

Changes in physical function scores, as measured with the DS, following surgery for LDS were significant over time (ANOVA, F[3,2009] = 170.0, p < 0.0001), and all postoperative scores were significantly lower than baseline (p < 0.01; Fig. 1B). Recovery plateau on the DS was reached at 1 year after surgery (baseline 45.9 ± 14.4 vs 1 year 24.7 ± 18.8, p < 0.01). The recovery plateau for PCS scores was also reached at 1 year postoperatively (baseline 33.1 ± 8.34 vs 1 year 44.1 ± 9.6, p < 0.01). The changes in PCS scores were significantly different over time (ANOVA, F[3,1975] = 166.3, p < 0.0001), with all postoperative scores significantly higher than baseline (p < 0.01; Fig. 1E).

The pain dimension, as expressed by VAS leg and back scores, was significantly different over time (VAS leg: ANOVA, F[3,2003] = 378.4, p < 0.0001; VAS back: ANOVA, F[3,2003] = 370.9, p < 0.0001; Fig. 2B and E). Improvement in VAS leg and back scores plateaued at 3 months (VAS leg baseline 7.3 ± 2.1 vs 3 months 2.7 ± 2.8,
p < 0.01; VAS back baseline 6.9 ± 2.3 vs 3 months 3.0 ± 2.3, p < 0.01).

Mental QOL after surgery for LDS also improved (ANOVA, F[3,1975] = 27.6, p < 0.0001), and all MCS scores postoperatively were greater than baseline (p < 0.01). MCS scores plateaued at 3 months postoperatively (baseline 48.9 ± 8.7 vs 3 months 52.5 ± 7.3, p < 0.01; Fig. 3B).

**Time to Plateau After Fusion Versus Nonfusion Surgery**

We next assessed the impact of fusion surgery on the timing of PRO recovery after LSS and LDS. For both LSS and LDS, patients who did not undergo fusion surgery reached a recovery plateau by 3 months on physical function scores, compared to those with fusion who plateaued by 1 year (LSS: nonfusion DS baseline 44.6 ± 15.7 vs 3 months 27.0 ± 18.4, p < 0.01; fusion DS baseline 50.5 ± 14.9 vs 12 months 29.9 ± 19.5, p < 0.01; nonfusion PCS baseline 32.6 ± 8.4 vs 3 months 41.6 ± 8.7, p < 0.01; fusion PCS baseline 31.1 ± 8.1 vs 12 months 40.8 ± 9.3, p < 0.01; LDS: nonfusion DS baseline 44.1 ± 14.2 vs 3 months 25.2 ± 16.9, p < 0.01; fusion DS baseline 46.6 ± 14.6 vs 12 months 25.1 ± 18.9, p < 0.01; nonfusion PCS baseline 33.3 ± 8.3 vs 3 months 42.4 ± 8.9, p < 0.01; fusion PCS baseline 32.7 ± 8.2 vs 12 months 44.1 ± 9.3, p < 0.01; Figs. 4 and 5). All postoperative scores were significantly different from baseline as assessed by ANOVA (nonfusion LSS: DS, F[3,1355] = 90.5, p < 0.0001; PCS, F[3,1355] = 103.6, p < 0.0001; fusion LSS: DS, F[3,1765] = 120.7, p < 0.0001; PCS, F[3,1765] = 113.6, p < 0.0001; nonfusion LDS: DS, F[3,330] = 34.8, p < 0.0001; PCS, F[3,330] = 25.8, p < 0.0001; fusion LDS: DS, F[3,1575] = 126.1, p < 0.0001; PCS, F[3,1575] = 135.4, p < 0.0001).

The timing of recovery in the pain dimension and mental QOL did not differ between fusion and nonfusion cohorts and plateaued by 3 months for both PROs (LSS: nonfusion VAS leg baseline 7.4 ± 2.1 vs 3 months 3.3 ± 2.9, p < 0.01; fusion VAS leg baseline 7.5 ± 1.9 vs 3 months 3.1 ± 2.8, p < 0.01; nonfusion VAS back baseline 6.5 ± 2.4 vs 3 months 3.1 ± 2.5, p < 0.01; fusion VAS back baseline 7.4 ± 2.1 vs 3 months 3.5 ± 2.5, p < 0.01; nonfusion MCS baseline 50.2 ± 8.0 vs 3 months 53.1 ± 6.4, p < 0.01; fusion MCS baseline 47.4 ± 8.6 vs 3 months 50.8 ± 7.9, p < 0.01; LDS: nonfusion VAS leg baseline 7.4 ± 1.8 vs 3 months 2.8 ± 2.7, p < 0.01; fusion VAS leg baseline 7.4 ± 2.1 vs 3 months 2.8 ± 2.7, p < 0.01; nonfusion VAS back baseline 6.7 ± 2.4 vs 3 months 3.1 ± 2.4, p < 0.01; fusion VAS back baseline 7.0 ± 2.3 vs 3 months 3.0 ± 2.4, p < 0.01; nonfusion MCS baseline 50.6 ± 9.1 vs 3 months 53.8 ± 6.2, p < 0.01; fusion MCS baseline 48.9 ± 8.4 vs 3 months 52.0 ± 7.5, p < 0.01; Figs. 4 and 5). All postoperative scores were significantly different from baseline as assessed by ANOVA (LSS: nonfusion VAS leg, F[3,1355] = 199.2, p
< 0.0001; fusion VAS leg, F[3,1765] = 290.3, p < 0.0001; nonfusion VAS back, F[3,1355] = 153.3, p < 0.0001; fusion VAS back, F[3,1765] = 283.8, p < 0.0001; nonfusion MCS, F[3,1355] = 12.6, p < 0.0001; fusion MCS, F[3,1765] = 45.7, p < 0.0001; LDS: nonfusion VAS leg, F[3,330] = 67.5, p < 0.0001; fusion VAS leg, F[3,1575] = 295.7, p < 0.0001; nonfusion VAS back, F[3,330] = 50.0, p < 0.0001; fusion VAS back, F[3,1575] = 286.7, p < 0.0001; nonfusion MCS, F[3,330] = 3.2, p < 0.02; fusion MCS, F[3,1575] = 18.7, p < 0.0001).

Discussion

We assessed the time course of PROs for pain, physical function, and mental health to reach a plateau following lumbar spine surgery for LDH, LSS, and LDS. The time course to the recovery plateau depends on the specific health dimension being measured. Physical function as measured by the DS and the PCS plateaued at 1 year after surgery for LDH, LSS, and LDS. Pain (VAS leg and back) and mental QOL (MCS) reached a recovery plateau at 3 months, regardless of the pathology. We also demonstrated that patients undergoing fusion for LSS and LDS plateaued at 1 year on measures of physical function, compared to 3 months in those without fusion.

Findings in the current study are supported by previous results from the Spine Patient Outcomes Research Trials (SPORTs).16,18,19 Although the original intent of these trials was not to assess the timing of recovery, investigators found that PROs after lumbar spine surgery for LDH, LSS, and LDS plateaued before 2 years postoperatively. Additionally, previous work using surgical registries to examine PRO scores after lumbar spine surgery demonstrated that DS scores continued to improve up to a year after surgery.2,10 In the present study, improvements in the scores for VAS (3.2–4.7) and DS (18.5–28.4) at the time of the recovery plateau are similar to those previously reported for lumbar pathology at 2 years and reflect the minimum clinically important differences previously determined for PROs.9

Although our current data are supported by previous studies, we are, to the best of our knowledge, the first to examine the specific time to PRO recovery of pain, function, and mental health after lumbar spine surgery, and there are some important differences to note. Previous studies assessing recovery trajectories have grouped all lumbar pathologies together, making it difficult to elucidate differences for specific pathologies, or they have assessed cases from a single center, limiting generalizability.10 In our study, patients undergoing surgery for lumbar pathologies reached a plateau at 1 year on measures of physical function (DS and PCS) but at 3 months on pain and mental QOL outcomes, regardless of pathology. Furthermore, previous studies aimed at assessing the timing of recovery have not examined follow-up past 1 year,2,10 which further limits the ability to draw conclusions regarding when recovery is truly achieved after surgery for lumbar pathologies. In contrast, the SPORT trials16,18,19 had long-term follow-up, but the trial investigators’ intent was to assess the adequacy of the treatment effect rather than the timing of recovery. Parker and colleagues10 asked if PROs at 3 months after lumbar spine surgery could predict scores at 1 year postoperatively and demonstrated that 3-month scores are not highly accurate predictors of long-term outcome. They noted that approximately 10% of patients who had not met a minimum clinically important difference at 3 months did eventually reach this milestone at 1 year. In light of the data presented in our current study, we believe that this finding likely reflects the fact that the time course to recovery depends on the specific PRO being utilized rather than the lumbar pathology being examined. Furthermore, a follow-up duration of 1 year will adequately capture the outcome in all patients who may not have fully recovered by 3 months.

The use of prospective registries in spinal surgery care is becoming increasingly prevalent.1,2,8,11,12,14 Such registries can provide outcome data that reflect real-world clinical practice in a way that randomized trials, which are heavily weighted toward efficacy and internal validity, cannot. While registries are viewed as a practical and more cost-
effective means to acquire data than controlled trials, they can still be very costly. Perhaps the largest driver of prospective registry cost is the duration of follow-up, which becomes increasingly expensive as more patients are enrolled. It is estimated that each additional year of follow-up can cost $160 (US$) per year per patient. Precisely targeting follow-up durations for specific outcome dimensions of lumbar pathologies will ensure that outcomes are evaluated in a time-efficient manner as well as a cost-effective manner, which has implications for the design of prospective registries and clinical trials.

The results of this study have important practical implications. The generic dictum that all clinical research requires 2 years of follow-up data, independent of the research question, outcomes, or pathology of interest, does not appear to be valid. Disseminating this fact is important not only to provide guidance for researchers, but also to facilitate the peer review process and ensure that clinical research is appropriately evaluated. Evidence-based pathology and outcome-specific recovery patterns are essential information for the clinician so they can better inform patients and monitor their recovery. Additionally, for all measured PROs and disease categories, 1 year appeared to be adequate to reach the plateau, though some did so earlier. Interestingly, we also demonstrated that pain and mental QOL plateaued earlier than physical function (DS, PCS). Furthermore, we also showed that patients who did not have a fusion operation for LSS or LDS reached a plateau by 3 months, as compared to 1 year in the patients who underwent fusion surgery. This finding suggests that a 1-year follow-up may be adequate to assess the efficacy of established surgical treatments for these lumbar spine conditions. Limiting the duration of follow-up may decrease the ability to provide surveillance of new implant devices or long-term complications, but this drawback may be mitigated by tracking complications in a registry rather than by ongoing QOL assessments. These data, building on previous work, provide a rationale for considerable cost savings in large patient registries and offer a welcomed reduction in the patient questionnaire burden.

There are several limitations to our study. First, we did not have 6-month postoperative outcome data available for analysis, and it is possible that measures of physical function (both DS and PCS) plateaued by this time. Second, all data came from centers in Canada, which has a universal healthcare system, and this possibly makes generalization to countries with different payer models less applicable. Nevertheless, within Canada, the inclusion of multiple centers, representing a broad spectrum of clinical practice patterns, enhances the generalizability of our study. Third, the rate of loss to follow-up at 1 and 2 years was up to 25%, and it is possible that this attrition could have altered our results. This follow-up is close to the 80% gold standard for clinical trials and is excellent for a registry design. The rate lost to follow-up at 3 months was approximately 10%, and at 1 year it was between 15% and 25%, depending on the specific lumbar pathology. Patients undergoing surgery for LDH were more likely to be lost to follow-up at 2 years than those being treated for LDS (about 25% vs about 15%). These trends likely represent real-world practices in follow-up patterns after lumbar spine surgery, in which patients who undergo fusion procedures tend to be followed up for a longer duration than those who undergo decompression alone. Fourth, the intent of our study was to assess the time course to the recovery of function and the improvement of pain and mental QOL after lumbar spine surgery and not specific clinical factors predictive of outcome; thus, specific factors such as age, smoking, comorbidities, or fusion were not adjusted for in order to assess overall trends in these patient populations. Last, we did not specifically assess other validated patient-reported measures of function or specific mental health dimensions (e.g., depression, anxiety, kinesiophobia); thus, we cannot say whether other specific PROs measuring the same or different dimensions would result in similar or unique findings. It would be useful for future studies from established prospective registries to also examine the time to the recovery plateau in PROs after lumbar surgery.

Conclusions

Specific health dimensions follow distinctly different recovery plateaus, indicating that a 2-year postoperative follow-up is not required for all PROs to accurately assess the treatment effect of established surgeries for lumbar spinal pathologies. The clinical research question should dictate the follow-up time and the outcome measure utilized; however, there is now evidence to guide the specific duration of follow-up for pain, physical function, and mental dimensions. This information will be useful in designing and maintaining prospective registries and in counseling patients and monitoring their recovery after lumbar spine surgery.

Acknowledgments

Dr. Fisher has received a grant to his institution from the Orthopaedic Research and Education Foundation.

References

7. McGirt MJ, Speroff T, Dittus RS, Harrell FE Jr, Asher AL:


Disclosures
Dr. Fisher has received consultant fees from Medtronic and NuVasive, has received royalties to his institution from a Medtronic research grant, and has received fellowship support to his institution from AOSpine and Medtronic for work outside of the present study. Dr. Johnson has received grant support to his institution from Stryker. Dr. Paquet has received educational grant support to his institution from Medtronic of Canada. Dr. Manson has been a consultant for and received non–study-related support from Medtronic. Dr. Rampersaud has been a consultant for and received royalties from Medtronic.

Author Contributions

Supplemental Information
Online-Only Content
Supplemental material is available with the online version of the article. 
*Supplemental Fig. 1 & Tables 1–5*. https://thejns.org/doi/suppl/10.3171/2018.8.SPINE18715.

Previous Presentations
This work was presented at the Annual Meeting of the AANS/CNS Section on Disorders of the Spine and Peripheral Nerves held in Orlando, FL, on March 14–17, 2018.

Correspondence
Charles G. Fisher: Vancouver General Hospital/University of British Columbia, Vancouver, BC, Canada. charles.fisher@vch.ca.