Comparative outcomes and cost-utility after surgical treatment of focal lumbar spinal stenosis compared with osteoarthritis of the hip or knee—part 1: long-term change in health-related quality of life

Y. Raja Rampersaud, MD, FRCSC, Stephen J. Lewis, MD, FRCSC, J. Roderick Davey, MD, FRCSC, Rajiv Gandhi, MD, MS, FRCSC, Nizar N. Mahomed, MD, MPH

*Division of Orthopaedic Surgery, Department of Surgery, Toronto Western Hospital, University Health Network, University of Toronto, 399 Bathurst St, Toronto, Ontario, Canada M5T-2S8

Arthritis Program, Toronto Western Hospital, University Health Network, University of Toronto, Toronto, Ontario, Canada

Spinal Program, Krembil neuroscience center, Toronto Western Hospital, University Health Network, University of Toronto, Toronto, Ontario, Canada

Received 6 February 2013; revised 2 December 2013; accepted 4 December 2013

Abstract

BACKGROUND CONTEXT: It is well accepted that total hip and knee arthroplasty (THA/TKA) for osteoarthritis (OA) is associated with reliable and sustained improvements in postoperative health-related quality of life (HRQoL). Although several studies have demonstrated comparable outcomes with THA/TKA after surgical intervention for lumbar spinal stenosis (LSS), the sustainability of the outcome after LSS surgery compared with THA/TKA remains uncertain.

PURPOSE: The primary purpose of this study is to assess whether improvements in HRQoL after surgical management of focal lumbar spinal stenosis (FLSS) with or without spondylolisthesis are sustainable over the long term compared with that of THA/TKA for OA.

STUDY DESIGN: Single-center, retrospective, longitudinal matched cohort study of prospectively collected outcomes, with a minimum of 5-year follow-up (FU).

PATIENT SAMPLE: Patients who had primary one- to two-level spinal decompression with or without instrumented fusion for FLSS and THA/TKA for primary OA.

OUTCOME MEASURES: Postoperative change from baseline to last FU in Short-Form 36 physical component summary (PCS) and mental component summary (MCS) scores among groups was used as the primary outcome measure.

METHODS: An age, sex-matched inception cohort of primary one- to two-level spinal decompression with or without instrumented fusion for FLSS (n = 99) was compared with a cohort of primary THA (n = 99) and TKA (n = 99) for OA and followed for a minimum of 5 years. Linear regression was used for the primary analysis.
same site) requiring 12 operations, respectively (p<.01). The average time to first revision was 56/65/43 months, respectively. Mean postoperative PCS (p<.0001) and MCS (p<.02) scores improved significantly and were durable for all groups at the last FU. The mean changes from baseline PCS/MCS scores to last FU were 8.5/6.4, 12.3/7.0, and 8.3/4.9 for FLSS, THA, and TKA, respectively. Adjusting for baseline age, sex, body mass index, PCS score, and MCS score, there was a strong trend in favor of greater sustained change in the PCS score of THA over FLSS (p=.07) and TKA (p=.08). No difference was noted for change in PCS score between FLSS and TKA (p=.95). No differences were noted for change in MCS score among all three cohorts (p>.1).

**CONCLUSIONS:** Significant improvements in HRQoL after surgical treatment of FLSS with or without spondylolisthesis and hip and knee OA are sustained for a mean of 7 to 8 years, with a minimum of 5-year FU. Despite a higher revision rate, patients undergoing surgery for FLSS can expect a comparable long-term average improvement in HRQoL from baseline compared with their peers undergoing TKA and to a lesser extent THA.

© 2014 The Authors. Published by Elsevier Inc. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/).

**Keywords:** Long term; Spinal stenosis; Osteoarthritis; Hip; Knee; Surgery; Health-related quality of life

---

**Introduction**

Musculoskeletal disorders such as low back pain and osteoarthritis (OA) are among the leading causes of years lived with disability and have a major impact on functional capacity [1,2]. Consistent with the increasing prevalence of these conditions within the aging population, the demand for surgical treatment of OA of the hip, knee, and spine has been on the rise [3–8]. As the global population continues to age, this trend has serious implications for the future use of health-care resources in these three common surgical populations [9].

Standard surgical treatment of end-stage symptomatic hip and knee OA consists of total hip and knee arthroplasty (THA/TKA), respectively; the surgical management of symptomatic OA of the spine (ie, lumbar spinal stenosis [LSS] with or without degenerative spondylolisthesis) are decompression and decompression with fusion [4]. Primary THA and TKA have proven to be the effective and durable treatments for hip and knee OA, in terms of pain relief, improved functional status, quality of life (QoL), and overall patient satisfaction [10–12]. Furthermore, primary THA and TKA have proven to be among the most cost-effective surgical interventions, both musculoskeletal and otherwise [13–15]. Over the past decade, there has been increasing recognition and awareness of the challenges posed by the scarcity of health-care resources against virtually unlimited health-care needs and increasingly expensive modes of treatment [9,16]. As a result, decision makers at all levels are under greater pressure to justify their resource allocation and priority-setting decisions. Specific to the surgical management of degenerative illnesses, issues of surgical wait time, budget impact, cost-effectiveness, and utilization are an ongoing focus of health policy [5–8,17–23]. Subsequently, societal demand combined with sustainable cost and clinical effectiveness of THA and TKA have played a pivotal role in the general acceptance of these procedures by all stakeholders (patients, health-care providers, surgeons, payers, and government agencies) [24].

In contrast, until more recently, the data supporting the long-term cost- and clinical effectiveness of surgery for LSS are limited and inconsistent regarding durability [25–32]. Recent work from our group and others has demonstrated that surgical management of patient with focal lumbar spinal stenosis (FLSS) (one to two level) with or without degenerative lumbar spondylolisthesis (DLS) results in comparable improvement in health-related quality of life (HRQoL) compared with TKA at 2 years after surgery [32–37]. In addition, we have also demonstrated that from the perspective of the hospital, the estimated lifetime cost-utility of spinal surgery for the aforementioned LSS population is also comparable with both THA and TKA [38]. The current two-part study presents long-term clinical and cost-utility data from our previously published spine, hip, and knee cohorts [36,38]. The primary objective of part one, presented herein, was to assess whether improvement in HRQoL after surgical management of FLSS with or without DLS is sustainable over the long term compared with THA/TKA for OA.

**Methods**

**Study design**

This is a single-center, retrospective, longitudinal matched cohort study of prospectively collected outcomes, with a minimum of 5-year follow-up (FU).

**Patient population**

As described in Rampersaud et al. [36], the inception cohort was determined after treatment and independently (by research individuals not involved in patient care) selected from databases containing prospectively collected outcome measurements of consecutive patients who received THA and TKA for OA and decompression with or without fusion for LSS at our institution (academic tertiary care hospital) over a 4-year period from January 2000 to December...
and the goal of surgery was to relieve the leg symptoms)
relieved by postural change/rest, no or tolerable mechanical
back pain, anatomy favorable to facet-sparing (ie, undercut-
tting) decompression, and no obvious dynamic instability
with or without static spondylolisthesis (up to Grade I).
Patients who had greater than Grade I spondylolisthesis
or dynamic instability demonstrated on supine to standing
or flexion-extension imaging, who had facet anatomy that
precluded adequate decompression, or who had concomi-
tant mechanical back pain resulting from one- or two-
level disease that was felt to be intolerable by the patient
received decompression with instrumented fusion. Patients
who underwent both decompression and fusion had a con-
ventional open midline approach with complete laminecto-
my and segmental pedicle screw fixation with the use of
local bone and iliac crest bone graft (in the majority).

For the hip and knee patients, an uncemented THA
was performed using the Hardinge (lateral) approach and
a cemented TKA using a medial parapatellar approach, re-
spectively, was used in all patients.

Subsequent surgeries performed on the same site, for hip
and knee OA, were considered as a revision. Contralateral
or adjacent joint replacement was not included as a revi-
sion. Spine revision was defined as same site surgery and/
or adjacent site surgery.

Data collection

Data included the following patient characteristics: age,
sex, body mass index (BMI), and American Society of Anes-
thesiologists (ASA) physical status classification. The
preoperative and postoperative (1- and 2-year) Medical Out-
comes Study Short-Form 36 (SF-36) general health survey
scores were assessed. As part of prospective registries and
ongoing quality assurance initiatives, validated disease-
specific (hip and knee OA Index of Western Ontario and
McMaster Universities/LSS Oswestry Disability Index)
and generic outcome (SF-36) measures have been prospec-
tively collected within the division of orthopedic surgery
since 1998 (TJA) and 2000 (spine). The SF-36 has been
shown to be valid and responsive to a wide variety of med-
ical diagnoses. Furthermore, it has shown appropriate
responsiveness compared with disease-specific outcome
measures for both degenerative spinal pathologies (Oswes-
try Disability Index) and hip and knee OA (OA Index of
Western Ontario and McMaster Universities) [39–41].

Outcome measures

The postoperative change from baseline to last FU in
SF-36 physical component summary (PCS) and mental
component summary (MCS) scores among groups was used
as the primary outcome measure. Because of the dramatic
differences in resource needs and surgical intensity of de-
compression alone versus decompression and fusion, we al-
so performed a subgroup analysis regarding change from

2003. Before receiving a surgical assessment, each patient
had undergone at least 6 months of conservative care. Con-
servative care entailed typical care as per the referring
physician (analgesia, formal and informal exercise, activity
modification, and assistive devices where applicable). At
the time of the index surgery, the typical wait time for re-
ferral to surgical consultation in the study populations
was over 1 year and the average surgical wait time (time
for decision to proceed with surgery to surgery) was 6 to
12 months for all three cohorts.

Inclusion criteria for LSS were intermittent neurogenic
claudication resulting from one- or two-level spinal stenosis
(i.e., FLSS), with or without DLS. Exclusion criteria for this
group included other causes of spinal stenosis (congenital,
posttraumatic, and degenerative scoliosis); multilevel sur-
gery (more than two levels); previous surgeries at the symp-
tomatic (previous discectomy was accepted) or adjacent
level; or multilevel coronal and sagittal plane deformity.
One- and two-level LSS were chosen a priori to compare with
total joint arthroplasty (TJA) as they represent the most com-
mon cases of surgery for LSS and have surgical parameters
similar to those of TJA (morbidity and hospital stay). Patients
in the THA and TKA groups had mechanical pain resulting
from primary OA of the hip or knee. Exclusion criteria in-
cluded secondary causes of OA (posttraumatic), inflamma-
tion, and previous surgery other than knee arthroscopy.

All patients represent a pragmatic cohort, where the deci-
sion to offer and accept surgery was an elective decision be-
tween the patient and the usual practice of the treating
surgeon. All consecutive patients who underwent surgery
for the generic LSS diagnosis (n = 220) during the study pe-
riod were assessed for the inclusion and exclusion criteria;
99 patients met the previously noted inclusion criteria and
were independently matched to a larger cohort of hip
(n = 248) and knee (n = 260) OA patients fitting the inclusion
criteria. Nine patients (all without DLS) in whom complete
FU data were obtained subsequent to inception cohort anal-
ysis and thus not reported in 2008 publication are included
in this analysis (did not have 2-year data but had long-
term data). The cohorts were matched in a blinded (blinded
to outcome data) manner according to age (within 2 years),
sex, and date of surgery. Matching to within 1 month of sur-
gical date was performed to eliminate any possible effects
from the potential variance of health-care delivery or other
unknown system confounders. If multiple matches were
found, a best-fit (age and date of surgery) principle was used.

Surgery

Surgical management of FLSS entailed a decompression
alone or decompression with fusion. A midline anatomy-
preserving decompression was used for patients receiving
decompression alone. This was chosen for patients with
leg-dominant symptoms (based on consensus between pa-
tient and surgeon that leg symptoms are greater and/or
more functionally significant than axial back symptoms

Data included the following patient characteristics: age,
sex, body mass index (BMI), and American Society of Anes-
thesiologists (ASA) physical status classification. The
preoperative and postoperative (1- and 2-year) Medical Out-
comes Study Short-Form 36 (SF-36) general health survey
scores were assessed. As part of prospective registries and
ongoing quality assurance initiatives, validated disease-
specific (hip and knee OA Index of Western Ontario and
McMaster Universities/LSS Oswestry Disability Index)
and generic outcome (SF-36) measures have been prospec-
tively collected within the division of orthopedic surgery
since 1998 (TJA) and 2000 (spine). The SF-36 has been
shown to be valid and responsive to a wide variety of med-
ical diagnoses. Furthermore, it has shown appropriate
responsiveness compared with disease-specific outcome
measures for both degenerative spinal pathologies (Oswes-
try Disability Index) and hip and knee OA (OA Index of
Western Ontario and McMaster Universities) [39–41].

Outcome measures

The postoperative change from baseline to last FU in
SF-36 physical component summary (PCS) and mental
component summary (MCS) scores among groups was used
as the primary outcome measure. Because of the dramatic
differences in resource needs and surgical intensity of de-
compression alone versus decompression and fusion, we al-
so performed a subgroup analysis regarding change from
baseline PCS score for FLSS patients undergoing decompression alone or decompression and fusion compared with H-OA and K-OA.

Institutional ethics approval

The study was approved by the Institutional Research Ethics Board (#04-0283-BE).

Statistical analysis

Statistical analysis was performed by an independent biostatistician. SAS 9.3 (SAS Institute Inc., Cary, NC, USA) statistical package was used to perform the analyses. Summary statistics to assess the baseline comparability of the groups with respect to demographics were used. Unadjusted comparisons among groups were performed using one-way analysis of variance (for continuous variables) and t test or Pearson chi-square test (for categorical variables).

Univariate analyses were conducted to identify factors that were different among cohorts or significantly related to outcomes. Age, gender, BMI (different distribution among cohorts at last FU), and baseline PCS and MCS scores were all entered into simple linear regression models to adjust for the relationship of outcomes comparing spinal surgery with arthroplasty and subgroup analyses of specific procedures. Follow-up time was not shown to be associated with any of the outcomes in the univariate tests (p = .78 for PCS score change and p = .65 for MCS score change).

Results

Patient baseline characteristics

Baseline patient demographics are presented in Table 1. At baseline, the three groups (FLSS/H-OA/K-OA) were equally matched with respect to mean age, sex, BMI, and ASA physical status (p > .1). Within the FLSS group, 62 patients had decompression alone (n = 30, one level; n = 32, two level) and 28 had decompression and instrumented fusions (n = 12, one level; n = 16, two level). Thirty-eight patients within the FLSS group had a concomitant degenerative spondylolisthesis (n = 21, decompression alone; n = 17, decompression and fusion). For the FLSS cohort, there were no differences in subgroups with respect to baseline patient demographics and SF-36 PCS/MCS scores for those who had decompression (D) alone, those who had decompression and fusion (DF), and those with and without DLS (p > .1 for all groups).

The mean FU in months was 80.5 ± 16.04, 94.6 ± 16.62, and 80.6 ± 16.84 for the FLSS, H-OA, and K-OA cohorts, respectively. The range of FU was 5 to 10 years for all three cohorts. The number of remaining participants in each cohort at the last FU with a minimum of 5 years for the FLSS, H-OA, and K-OA cohorts were 78 (79%), 91 (92%), and 84 (85%), respectively. Within the 21 lost to FU in the FLSS cohort, 7 were deceased and 2 with confirmed terminal diseases; of the 8 lost to FU for the H-OA cohort, 3 were deceased; and of the 15 lost to FU for the K-OA cohort, 3 were deceased and 2 in long-term care facilities (Figure). The mean time in months at which patients were lost to FU was similar among groups (FLSS 38.1 ± 11.5, H-OA 31.5 ± 11.0, and K-OA 30.0 ± 10.9). Only two patients in each of the FLSS and H-OA cohorts were lost after the 5-year FU period. Compared with the baseline demographic and outcome data of the last FU cohorts, there was no difference in age in months at which patients were lost to FU was similar among groups (FLSS 38.1 ± 11.5, H-OA 31.5 ± 11.0, and K-OA 30.0 ± 10.9). Only two patients in each of the FLSS and H-OA cohorts were lost after the 5-year FU period. Compared with the baseline demographic and outcome data of the last FU cohorts, there was no difference in age and sex, BMI, and baseline PCS or MCS score to those lost to FU within each cohort. In the FLSS cohort, the lost to FU group was significantly older than those followed (61 + 11.7 vs. 72 + 8.7 years for those lost to FU, p < .001).

Unadjusted univariate analysis of FU cohorts

Analysis of the baseline data of the patients with a minimum of 5-year FU demonstrated no difference in age at surgery (p = .59), baseline PCS (p = .22) or MCS (p = .22) scores, or male/female distribution (p = .95). The BMI for the FU cohorts, however, was significantly different among cohorts (28.4 ± 5.5, 29.8 ± 6.2, and 32.0 ± 6.6 for FLSS, H-OA, and K-OA, respectively, p < .001). As demonstrated in Table 2, the mean postoperative PCS (p < .0001) and MCS (p < .02) scores improved significantly for all groups. For our primary outcome, there was no difference in the mean PCS score among cohorts at last FU (p = .37); however, the mean change in PCS score (p = .03) from baseline was significantly different in favor of the H-OA cohort. There was no significant difference in the mean (p = .46) or the mean change from baseline (p = .18) MCS score among cohorts.

Adjusted analysis

For our primary outcome, adjusting for baseline age, sex, BMI, PCS score, and MCS score, there was no
statistically significant difference in the change from baseline PCS score or MCS score to last FU among groups (Table 3). A trend for greater change was noted for the H-OA compared with both the FLSS (p = .07) and K-OA (p = .08) cohorts.

Adjusted subgroup analysis demonstrated almost identical findings for the decompression-alone cohort as those noted previously for the entire FLSS cohort (thus data not presented). For the smaller decompression and fusion cohort, the superior findings related to H-OA were no longer significant (Table 4).

Table 2
Mean PCS and MCS scores at 5- to 10-year FU

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (change from baseline)</th>
<th>SD</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS score</td>
<td>FLSS (n=78) 39.93 (8.52)</td>
<td>10.37</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>H-OA (n=91) 41.98 (12.25)</td>
<td>11.66</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>K-OA (n=84) 39.74 (8.33)</td>
<td>11.86</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>MCS score</td>
<td>FLSS (n=78) 49.49 (6.3)</td>
<td>10.90</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>H-OA (n=91) 51.97 (6.98)</td>
<td>11.00</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>K-OA (n=84) 50.53 (4.9)</td>
<td>12.50</td>
<td>&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>

* p Values represent comparison from baseline within each cohort.

Revision surgery

At the last FU, the current revision rate (for definitions, see “Methods” section) was significantly higher for the FLSS cohort and lowest for the H-OA cohort. The number of patients who have undergone revision, including those lost to FU, for spine, hip, and knee are n=20 (20.2%) requiring 27 operations, n=3 (3%) requiring 5 operations, and n=8 (8.1%) requiring 12 operations, respectively (p<.01). The average time to first revision was 56/65/43...
months, respectively. Subgroup revision rates for the FLSS cohort are shown in Table 5.

Discussion

The results of this study confirm the positive, significant, and sustained impact on HRQoL in patients electing to undergo surgical intervention for spine, knee, and hip OA on HRQoL. Furthermore, this study demonstrates that the improvement from baseline HRQoL for the surgical management of FLSS previously demonstrated at 2 years is sustainable beyond 5 years and comparable with that of TJA for H-OA and K-OA. However, the reoperation rate for the FLSS cohort is significantly higher than that of THA and TKA [36]. These results have far-reaching implications from a direct patient care and health policy perspective. It is estimated that over the next 30 to 40 years, the proportion of seniors in Canada, the United States, and in most global industrialized countries (those older than 60–65 years) will double [42–44]. In addition, Perruccio et al. [45] have noted an increasing prevalence of arthritis that was not explainable by the aging population alone. Perruccio et al. [45] estimated that by 2021, the prevalence of arthritis in Canada would be increased 21% to 26% compared with 17.6% in 2003. In the United States, several authors have noted similar concerns and have reported on the rapidly increasing rate of surgery for degenerative conditions of the spine and TJA [4–8,17,18,20,21]. The implications of our study on health system funding and human resource allocation are substantial if we are to accommodate the increasing demand in the surgical care of Musculoskeletal conditions, particularly OA of the spine, hip, and knee [9,19].

In Canada, wait times can result in significant delays in obtaining surgical consultation and management for elective disorders in certain specialties and subspecialties [33,46,47]. Wait times for obtaining surgical services for common procedures such as total joint replacement have been particularly publicized [19,48–51]. In 2003, the Canadian First Ministers’ Accord on Health Care Renewal was struck and a commitment to achieve meaningful reductions in wait times in priority areas such as cancer, heart, diagnostic imaging, joint replacements, and sight restoration was publically made [50]. Since that time, significant, and often dramatic, improvements have been made in accessibility in these five top priority areas. Prioritization of hip and knee replacement was not only purely on the basis of patient demand but also from the consistent demonstration of comparative clinical effectiveness, durability, and cost-effectiveness of TJA [10]. In other words, hip and knee TJA has demonstrated significant health-care value [52].

The results of this study combined with the current growing higher quality literature for LSS suggest that similar arguments that have been made for THA/TKA can now be made for the surgical management of LSS. It is estimated that LSS causing neurogenic claudication affects about 20% of people older than 65 years and about half of that group suffer serious restrictions in their daily routines [42–44]. In a recent study by Battie et al. [54], the authors demonstrated that associated health burden of LSS on HRQoL was significant and is about the same or greater than diabetes, heart disease, arthritis, or stroke. Furthermore, LSS is the commonest reason to undergo spine surgery in patients older than 65 years [4,5,18]. Superior and sustained comparative cost- and clinical effectiveness for surgical management compared with conservative management at 2 and 4 years postoperatively for LSS with or without DLS have

Table 4
Simple linear regression models (subgroup analysis): PCS score change from baseline for FLSS decompression and fusion (DF), FLSS (DF)/THA/TKA at 5- to 10-year FU

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reference</th>
<th>Coefficient</th>
<th>95% CI for coefficient</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS-OA change from baseline to 5- to 10-year FU</td>
<td>H-OA FLSS (DF)</td>
<td>2.61</td>
<td>−3.23 to 8.46</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>K-OA FLSS (DF)</td>
<td>−0.97</td>
<td>−6.92 to 4.98</td>
<td>.75</td>
</tr>
</tbody>
</table>

CI, confidence interval; FLSS, focal lumbar spinal stenosis; FU, follow-up; H-OA, hip osteoarthritis; K-OA, knee osteoarthritis; PCS, physical component summary; MCS, mental component summary; SD, standard deviation; THA/TKA, total hip and knee arthroplasty.

Note: Adjusted for baseline age, sex, body mass index, PCS score, and MCS score. All parameters were significantly related to the outcome (p<.02).

Table 5
Revision surgery for the FLSS cohort

<table>
<thead>
<tr>
<th>FLSS Subgroup</th>
<th>Primary</th>
<th>Revisions (number of surgeries)</th>
<th>Same site</th>
<th>Adjacent segment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td>DF</td>
</tr>
<tr>
<td>LSS with DLS (n=38)</td>
<td>D (n=21)</td>
<td>3 (3)</td>
<td>0</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean time to revision 44.3 (m)</td>
<td>7 (9)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean time to revision 43.1 (m)</td>
<td>7 (12)</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean time to revision 41.4 (m)</td>
<td>3 (3)</td>
<td>0</td>
</tr>
<tr>
<td>LSS without DLS (n=61)</td>
<td>D (n=51)</td>
<td>7 (12)</td>
<td>1 (1)</td>
<td>3 (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean time to revision 55.3 (m)</td>
<td>3 (3)</td>
<td>0</td>
</tr>
</tbody>
</table>

D, decompression; DF, decompression and fusion; DLS, degenerative lumbar spondylolisthesis; FLSS, focal lumbar spinal stenosis.

Note: Revisions=number of patients with (number of surgeries).

* One unrelated remote L1 fracture.
been recently demonstrated by the large multicenter Spine Patient Outcomes Research Trial studies confirming the finding of previous smaller studies [25–27,55,56]. The demonstration of comparable and sustainable improvement of HRQoL after the surgical management of FLSS compared with the benchmark set by TJA for OA enables patients, surgeons, and policy and decision makers to consider not only the comparative effectiveness within the spine literature but also two well accepted and familiar nonspine interventions commonly performed in the aging population. Our group has also demonstrated generalizability by affirming our 2-year results in a recent multicenter study [33]. The relatively consistent results achieved among hospital centers in our multicenter study suggest that unlike the surgical management of low back pain for degenerative disc disease, consistent surgical outcome for the management of FLSS is achievable on a national scale [57,58]. Furthermore, our early findings at 2 years postintervention are in line with several other similar studies that have been conducted in different countries [33–37]. The consistency across multiple studies affords a greater degree of confidence in our findings.

There are several notable strengths to this study. First, it still remains one of the first studies to provide direct comparison of four common surgical interventions in a growing population. Second, we used a validated patient-reported generic outcome measure (SF-36) that has shown adequate sensitivity across our study populations [39–41]. Third, the pragmatic nature of our matched inception cohort enables a greater degree of generalizability. Fourth, we have achieved an excellent response rate from all three cohorts at a minimum of 5-year (5–10 years) FU, ranging from 79% for the FLSS cohort and 92% for the H-OA cohort. Fifth, and most importantly, the long-term of minimum 5-year FU (mean of 7 years for FLSS and K-OA and 8 years for H-OA) enables a direct assessment of the relative clinical and surgical durability of these interventions. Although limited, the long-term literature for the surgical treatment of LSS is consistent with the findings of the current study. In 11 studies (6 prospective studies, 2 of which were randomized controlled trials) with a mean FU ranging from 5 to 13 years, outcomes and patient satisfaction were relatively maintained in the majority of patients (53–92%) [56,59–68]. The revision rates for the LSS population varied widely from 1% to 23%, with the majority reporting a revision rate of less than 10%. The revision rate in our FLSS cohort is certainly in the upper end of those reported and, however, is consistent with the rates reported by several prospective studies and one large administrative database study [26,27,62,63,69]. Furthermore, all same site (n=7 patients) and adjacent segment procedures (n=13 patients) were included in our revision rate, with the majority (65%) of revisions indicated for adjacent segment disease. The results of our H-OA and K-OA cohorts are consistent with the long-term clinical and surgical outcomes of TJA [70–77].

The limitations of this study are primarily methodological in nature and related to the retrospective selection of our consecutive FLSS inception cohort from prospective databases. The potential confounding effects of patient selection biases and difference in surgical technique cannot be accounted for but reinforces generalizability. Although the SF-36 allows for valid comparisons among different disease states or interventions, it nevertheless has a deficiency in orthopedic conditions, particularly in mental and social subcomponents, where it is less responsive to change and is subject to significant floor or ceiling effects [78]. This is possibly why there is significantly less difference in change for the MCS among our cohorts. Furthermore, determination of the PCS and MCS is dependent of some shared subscales and consequently is not independent. Another potential confounder that has been shown to impact the change in SF-36 after surgery is the presence or absence of medical or other MSK comorbidities [73,79]. Although surgical and anesthetic screening will have excluded patients with more severe cardiovascular or respiratory conditions, the effects of medical or psychosocial comorbidities common to these surgical populations were not controlled in this study. Furthermore, we did not collect specific data nor adjust for the likely development of new comorbidities or other joint dysfunction over the time duration of our study. Differential development of additional comorbidities among cohorts is likely and may influence the relative changes in SF-36 scores. However, controlling for other variables including age, sex, baseline BMI, ASA, and baseline SF-36 scores adjust for factors relevant to the patient and surgeon at the time of choosing surgical intervention. Furthermore, we also did not adjust for revision status; therefore, the negative impact of a higher revision rate for the FLSS cohort is reflected in the current data. Consequently, we feel that these data are valid from the perspective of the overall long-term HRQoL outcome regardless of future health events including revision surgery that may influence the HRQoL.

Another potential limitation of this study relates to biases and assumptions related to the comparison of different patient populations. First, our combination of spinal subdiagnoses (LSS with or without DLS) and surgical intervention (decompression alone or with instrumented fusion) from a policy perspective represents the common spectrum of LSS patients; thus, we felt it more relevant to assess the population as a whole. However, our spine cohort mixture (with or without DLS and decompression vs. decompression and fusion) is consistent with those presented in the aforementioned long-term studies. Consequently, we are confident that our spine cohort is representative of the typical patient undergoing surgery for FLSS. Second, although we are comparing relative change in outcomes from baseline among the spine, hip, and knee cohorts, outcomes may have been differentially impacted by possible differences in the degree of failure of nonoperative treatment among groups. However, because of the prolonged wait
times for consultation and eventual surgical intervention during that period of time in our health-care system, further improvement with conservative care was less likely in the present study. Specifically, as demonstrated by other Canadian studies [46,49], hip, knee, and spine patients are more likely to deteriorate while waiting for elective surgery rather than improve. These patients represent individuals who have failed to gain any benefit from further conservative care and elected to endure a significant wait (6–12 months) to have surgery. Furthermore, to reduce surgical burden and manage wait-list, any available methods to optimize nonoperative treatment before and after the decision to proceed with surgery are typically pursued.

Conclusions

For patients failing prolonged conservative care, significant improvement in HRQoL after elective surgical treatment of FLSS (including those with and without degenerative spondylolisthesis) and hip and knee OA is sustained to a mean of 7 to 8 years with a minimum of 5-year FU. Despite a higher revision rate, patients undergoing surgery for FLSS can expect a comparable long-term average improvement in HRQoL from baseline compared with their peers undergoing TKA and to a lesser extent THA.

Acknowledgments

The study was supported by unrestricted funds from the Toronto General and Western Foundation.

References


