Comparative Effectiveness Between Primary and Revision Foraminotomy for the Treatment of Lumbar Foraminal Stenosis

EMILY HU, JIANNING SHAO, ARBAZ MOMIN, MAXWELL Y. LEE, HEATH P. GOULD, ROY XIAO, COLIN M. HAINES, DON K. MOORE, THOMAS E. MROZ and MICHAEL P. STEINMETZ

Int J Spine Surg 2020, 14 (4) 511-517
doi: https://doi.org/10.14444/7067
http://ijssurgery.com/content/14/4/511

This information is current as of December 9, 2020.

**Email Alerts**  Receive free email-alerts when new articles cite this article. Sign up at: http://ijssurgery.com/alerts
Comparative Effectiveness Between Primary and Revision Foraminotomy for the Treatment of Lumbar Foraminal Stenosis

EMILY HU, MD,1 JIANNING SHAO, BA,1,2 ARBAZ MOMIN, BA,2 MAXWELL Y. LEE, BA,2 HEATH P. GOULD, MD,1 ROY XIAO, MD,1 COLIN M. HAINES, MD,1 DON K. MOORE, MD,1 THOMAS E. MROZ, MD,1 MICHAEL P. STEINMETZ, MD1

1Cleveland Clinic Center for Spine Health, Cleveland, Ohio, 2Cleveland Clinic Lerner College of Medicine, Cleveland, Ohio

ABSTRACT

Background: Foraminotomy has demonstrated clinical benefit in patients with lumbar foraminal stenosis (LFS), as evidenced by several small retrospective investigations. However, there is a subset of patients who have recurrent symptoms following the operation and therefore require revision surgery. Yet, despite this phenomenon, the relative efficacy of revision foraminotomy (RF) is not well elucidated due to limited literature on the quality of life (QOL) outcomes and cost associated with primary foraminotomy (PF) and RF.

Purpose: To compare the effectiveness of PF and RF in terms of QOL outcomes and relative costs.

Study Design/Setting: This is a retrospective cohort study conducted at a single tertiary-care institution. The patient sample consisted of patients undergoing foraminotomy for the treatment of LFS between 2008 and 2016. The primary outcome measure was improvement in postoperative QOL, as measured by EuroQol 5-Dimensions (EQ-5D), and secondary outcome measures included Pain Disability Questionnaire (PDQ) and Patient Health Questionnaire-9 (PHQ-9) perioperative cost as well as minimum clinically important difference (MCID).

Methods: A retrospective chart review was conducted to identify individuals who underwent PF or RF for LFS and to collect clinical, operative, and demographic data. QOL scores (EQ-5D, PDQ, and PHQ-9) were collected between 2008 and 2016, and perioperative financial data were extracted via the institution’s cost utilization engine. Paired t tests were used to assess changes within treatment groups, and Fisher exact tests were used for intercohort comparisons.

Results: Five hundred seventy-nine procedures were eligible: 476 (82%) PF and 103 (18%) RF. A significantly higher proportion of males underwent RF than PF (71% versus 59%, \( P = .03 \)), and PF was done on a significantly higher number of vertebral levels (2.2 versus 2.0, \( P = .04 \)). There were no other significant differences in demographics. Preoperatively, mean PDQ-Functional scores (50 versus 54, \( P = .04 \)) demonstrated significantly poorer QOL in the RF cohort. Postoperatively, EQ-5D index showed significant improvement in both the PF (0.547 → 0.648, \( P < .0001 \)) and the RF (0.507 → 0.648, \( P < .0001 \)) cohorts. Similarly, total PHQ-9 improved significantly in the PF cohort (7.84 → 5.91, \( P = .001 \)) and in the RF cohort (8.55 → 5.53, \( P = .02 \)), as did total PDQ (PF: 77 → 63, \( P < .0001 \); RF: 85 → 70, \( P = .04 \)). QOL scores were also compared between groups preoperatively and postoperatively, and the only significant difference between PF and RF was observed in the preoperative PDQ-Functional score (49.7 versus 54.3, \( P = .04 \)). The proportion of patients achieving MCID was not significantly associated with cohort. Finally, perioperative cost did not differ significantly between cohorts (PF: $13,383 versus RF: $13,595, \( P = .82 \)).

Conclusions: Both PF and RF produced significant improvement in nearly all measures in patients with LFS. There was no significant difference in cost between PF and RF, but both PF and RF showed postoperative QOL improvements as compared with preoperative scores, indicating that RF remains a reasonable treatment option for patients with recurrent symptoms of LFS.

Lumbar Spine
Keywords: lumbar, foraminotomy, primary, revision, stenosis

INTRODUCTION

Foraminal stenosis can be caused by multiple etiologies, the most notable of which include disc degeneration and loss of disc height, ultimately leading to perturbation of facet joint biomechanics, motion, and anatomical abnormalities (ie, hypertrophy).1 The specific clinical presentation, however, depends on the spinal levels involved. For instance, lumbar foraminal stenosis (LFS) can present with
leg pain at rest, back pain, paresthesia, radiculopathy, and weakness, all of which may lead to an impaired quality of life (QOL) and diminished functional capacity.\textsuperscript{2–4} To address these symptoms, one can employ either conservative treatments such as non-steroidal anti-inflammatory drugs, or surgical management, especially for patients who are refractory to the former.\textsuperscript{5}

A number of different surgical procedures are performed for LFS, ranging from simple decompression (eg, foraminotomy) to decompression with lumbar fusion. Although several studies have established primary foraminotomy (PF) as a viable treatment option, the effectiveness of revision foraminotomy (RF) has not been well-described.\textsuperscript{6–10} Furthermore, neither the cost of primary nor RF has been studied. In the present study, we compare the change in overall QOL, mental health status, and disability in patients who underwent PF and RF. We also performed a comprehensive financial analysis of PF and RF, including the costs incurred during the preoperative and postoperative periods.

**METHODS**

A retrospective cohort study was conducted to identify individuals who underwent PF or RF for the treatment of LFS. Consecutive patients from a single tertiary-care institution, between 2008 and 2016, were identified by Common Procedural Technology codes (63030 and 63047). The minimum follow-up was defined as 1 month, and a revision was defined as a foraminotomy procedure after a prior operation on the same side, at the same level. Patients who had prior surgeries other than foraminotomies (eg, discectomy, laminectomy) were not excluded. Selection of surgical procedure occurred at the discretion of the treating surgeons. All patients younger than 18 years of age were excluded on the basis of skeletal immaturity, as were patients with a spinal malignancy, infection, or history of acute trauma.

**Data Collection**

Clinical, operative, and demographic data were collected from the electronic medical record. A global Charlson Comorbidity Index (CCI) score was calculated for each subject enrolled in this study and was queried from institutional databases.\textsuperscript{11} Spine surgeons made the diagnosis of symptomatic LFS based on history, physical examination, and imaging. QOL data were collected preoperatively and at the last available postoperative follow-up visit.

**QOL Data**

Preoperative and postoperative QOL data were collected prospectively from an institutional database of patient-reported health status measures that include validated questionnaires administered at each outpatient visit. Questionnaires collected for this study include the EuroQol 5-Dimensions (EQ-5D), Patient Health Questionnaire-9 (PHQ-9), and Pain Disability Questionnaire (PDQ).\textsuperscript{12–19}

The minimum clinically important difference (MCID) for each QOL measure was obtained from the literature. The MCID for EQ-5D index following spine surgery was established in a population of patients receiving anterior cervical discectomy and fusion, and has been reported to be 0.1.\textsuperscript{13,14} The MCID for the PHQ-9 has been reported to be 5.\textsuperscript{20} The MCID for the PDQ has been reported to be 20.\textsuperscript{21} Both current PHQ-9 and PDQ MCIDs were established in more general populations experiencing depression and chronic pain, respectively.\textsuperscript{20,21}

**Cost Data**

Financial data were extracted using the institution’s cost utilization engine. These costs were generated based on patient-level resource utilization and include the total cost of delivering care in the inpatient and outpatient settings across the health system. The timeframe for financial data collection extended from 30 days before surgery to 1 year after discharge. The data were then subclassified into 3 episodes of care: (1) preadmission period, defined as the 30 days prior to the operation; (2) admission period; and (3) postdischarge period, defined as the 365 days following discharge. Costs included all resource expenditures for imaging, rehabilitation, laboratory diagnostics, inpatient medications, and healthcare encounters (ie, outpatient, inpatient, emergency room, home health visits). To present costs from the payer’s perspective, all costs were based upon institutional cost-to-charge ratios established by Medicare and presented in 2016 United States dollars.

Data were analyzed using the JMP Pro 12.0 statistical software (SAS Institute, Inc, Cary, North Carolina; 2012).\textsuperscript{22} Change in QOL within cohorts was compared using paired \( t \) tests and Wilcoxon signed rank tests for parametric and nonparametric continuous variables, respectively. Fisher exact were
used to compare categorical variables between cohorts. Absolute change in EQ-5D/PHQ-9/PDQ (defined as postoperative value minus preoperative value) was modeled, as well as absolute difference in cost between PF and RF, and was reported as regression coefficients ($b$). Percentage of patients exceeding the MCID was also reported as odds ratios (ORs).

## RESULTS

Five hundred seventy-nine patients were eligible for inclusion in this study. Four hundred seventy-six patients underwent PF (82%); 103 patients underwent RF (18%). Mean ages were 63.8 years in the PF group and 62.6 years in the RF group ($P = .35$). All demographic data are presented in Table 1. A significantly greater proportion of males underwent RF compared with PF (71% versus 59%, $P = .03$). For RF patients, the average number of foraminotomies per patient was 4.89 (standard deviation 3.54), and the average number of foraminotomies per level was 1.55. Additionally, PF was performed on a significantly higher number of vertebral levels compared with RF ($2.2 \pm 1.0$ versus $2.0 \pm 0.9$, $P = .04$). There were no statistically significant differences between treatment groups with regard to age, marital status, race, or CCI.

### QOL Intercohort Comparison Analysis

QOL metrics of PF and RF at the preoperative visit and at the last postoperative visit are presented in Table 2. Mean postoperative follow-up was 17.5 months in the PF cohort and 18.4 months in the RF cohort ($P = .71$).

EQ-5D index scores were compared preoperatively and postoperatively between treatment groups. Mean preoperative EQ-5D index was $0.547$ in the PF group and $0.507$ in the RF group ($P = .16$). Mean postoperative EQ-5D index at the last available follow-up visit was $0.648$ in the PF group and $0.648$ in the RF group ($P = .99$). The proportion of patients achieving an MCID on the EQ-5D instrument was 47% in PF and 59% in RF (PF versus RF: OR 1.64; $P = .25$).

Preoperative and postoperative PHQ-9 scores were also compared between cohorts. The mean preoperative PHQ-9 score was $7.84$ in the PF group and $8.55$ in the RF group ($P = .45$), whereas the mean postoperative PHQ-9 score at the last available follow-up visit was $5.91$ in the PF group and $5.53$ in the RF group ($P = .63$). The proportion of patients achieving an MCID on the PHQ-9 instrument was 25% in PF and 21% in RF (PF versus RF: OR 0.80; $P = .42$).

PDQ scores were compared in terms of composite score and subscores reflecting functional and psychosocial domains. These data are presented in Table 2. Mean PDQ-Functional scores were significantly poorer in the RF cohort preoperatively (49.7 versus 54.3, $P = .04$) and showed no significant difference postoperatively ($P = .13$). PDQ-Psychosocial scores showed no statistically significant difference between PF and RF either preoperatively 

### Table 1. Demographics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Primary</th>
<th>Revision</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>476</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>282 (59)</td>
<td>73 (71)</td>
<td>.03*</td>
</tr>
<tr>
<td>Age at surgery, mean ± SD, y</td>
<td>63.8 ± 11.0</td>
<td>62.6 ± 10.6</td>
<td>.35</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td>White</td>
<td>427 (90)</td>
<td>93 (90)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>38 (8)</td>
<td>9 (9)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>11 (2)</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
<td>.28</td>
</tr>
<tr>
<td>Single</td>
<td>53 (11)</td>
<td>9 (9)</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>335 (70)</td>
<td>83 (81)</td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>51 (11)</td>
<td>6 (6)</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>34 (7)</td>
<td>5 (5)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>3 (1)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>CCI, mean ± SD</td>
<td>1.0 ± 1.5</td>
<td>1.1 ± 1.6</td>
<td>.55</td>
</tr>
<tr>
<td>CCI ≥ 2, n (%)</td>
<td>125 (26)</td>
<td>29 (28)</td>
<td>.69</td>
</tr>
<tr>
<td>Number of vertebral levels, mean ± SD</td>
<td>2.2 ± 1.0</td>
<td>2.0 ± 0.9</td>
<td>.04*</td>
</tr>
<tr>
<td>Length of stay, mean ± SD, d</td>
<td>3.7 ± 4.6</td>
<td>4.0 ± 4.1</td>
<td>.63</td>
</tr>
<tr>
<td>Follow-up, mean ± SD, mo</td>
<td>17.5 ± 17.7</td>
<td>18.4 ± 17.7</td>
<td>.71</td>
</tr>
<tr>
<td>Perioperative costs, mean ± SD, $</td>
<td>3860 ± 4576</td>
<td>3181 ± 4382</td>
<td>.17</td>
</tr>
</tbody>
</table>

Abbreviation: CCI, Charlson Comorbidity Index.

Continuous variables reported as mean standard deviation; categorical variables reported as count (percent).

*Statistically significant, $P < .05$. 


by guest on December 9, 2020

International Journal of Spine Surgery, Vol. 14, No. 4

513
The proportion of patients achieving an MCID on the PDQ instrument was 42% in PF and 30% in RF (PF versus RF: OR: 0.58; \(P = .14\)).

**QOL Intracohort Improvement Analysis**

The change in QOL from the preoperative visit to the postoperative follow-up visit (average duration of 17.5 ± 17.7 months for primary foraminotomies and 18.4 ± 17.7 months for revision foraminotomies) was assessed for both PF and RF. These data are presented in Table 2. Mean EQ-5D index for the PF cohort improved significantly from 0.547 preoperatively to 0.648 postoperatively (\(P < .0001\)). Similarly, mean EQ-5D index for the RF cohort improved significantly from 0.507 preoperatively to 0.648 postoperatively (\(P < .0001\)).

PHQ-9 scores for the PF group decreased significantly from 7.84 preoperatively to 5.91 postoperatively (\(P = .001\)). PHQ-9 scores for the RF group decreased significantly from 8.55 preoperatively to 5.53 postoperatively (\(P = .02\)).

As in the intercohort QOL comparison, PDQ results were assessed in terms of total score as well as separated into functional and psychosocial components. All PDQ scores for the PF group improved significantly from the preoperative visit (Total: 77.3; Functional: 49.7; Psychosocial: 27.5) to the postoperative visit (Total: 62.9, \(P < .0001\); Functional: 41.4, \(P < .0001\); Psychosocial: 21.6, \(P = .03\)). PDQ-Total and Functional scores for the RF group improved significantly from the preoperative visit (Total: 85.1; Functional: 54.3) to the postoperative visit (Total: 69.8, \(P < .0001\); Functional: 45.7, \(P = .04\)). PDQ-Psychosocial scores for the RF group improved from the preoperative visit (30.8) to the postoperative visit (24.1), but this improvement did not reach statistical significance (\(P = .17\)).

**Cost Analysis**

Financial data are presented in Table 3. The median total cost of care was higher in the RF cohort than the PF cohort (PF: $22,465 [$14,183–$34,371]; RF: $22,539 [$13,729–$32,943]; \(P = .83\), but this difference did not reach significance. The median cost of the preoperative episode (PF: $1773 [$504–$7149]; RF: $1073 [$501–$3600]; \(P = .12\) was greater in the PF group, but again, this difference did not reach significance. The median cost of the 365-day postoperative episode (PF: $3317 [$544–$13,726]; RF: $4248 [$964–$12,495]; \(P = .04\) was greater in the RF group, but this difference did not reach significance. Finally, the difference between treatment groups with regard to the median cost of the hospital admission episode was not statistically significant (PF: $11,352 [$8127–$16,014]; RF: $11,434 [$8194–$17,118]; \(P = .98\)).
DISCUSSION

Foraminotomy is a commonly performed procedure for the treatment of LFS. Some patients undergo RF when the relief from the primary procedure has not been adequate. The efficacy of RF in comparison to that of PF remains unclear. In the present study, we compared preoperative and postoperative QOL, mental health status, and disability for PF and RF. The results of our study indicate that neither cost nor postoperative QOL was significantly different between PF and RF.

To our knowledge, no studies have examined the outcomes and cost of RF in comparison to PF. However, a small number of studies has evaluated the safety and efficacy of lumbar foraminotomy as a therapeutic option. The results of these investigations have consistently demonstrated that lumbar foraminotomy is associated with improvement of radicular symptoms and QOL. For instance, Ahn et al.\textsuperscript{10} showed that Oswestry Disability Index scores improved at 6 weeks, 1 year, and 2 years after endoscopic lumbar foraminotomy ($P < .001$). Further, the authors reported excellent or good results in 81.8\% of the patients, and symptomatic improvements in 93.9\%.\textsuperscript{10} Fu et al.\textsuperscript{23} compared traditional decompressive laminectomy to a window technique laminoforaminotomy and reported that patients in the latter group had fewer complications and superior outcomes with regard to back pain, leg pain, walking tolerance, and neurologic recovery (89\% versus 63\% with “good to excellent” results).

Although the literature suggests that foraminotomy is effective, symptomatic LFS can recur. In cases that are recalcitrant to nonoperative treatment, RF may be performed. By nature, revision lumbar surgeries tend to be more challenging and have been shown to have higher complications rates.\textsuperscript{24,25} The correct diagnosis must be confirmed and the most effective surgical plan—including determination of the appropriate levels, anatomical targets, and evaluation for scar tissue—must be developed.\textsuperscript{26,27} However, surgery for recurrent spine surgery can be performed safely with a reasonable degree of symptomatic relief if the aforementioned plan is well developed.\textsuperscript{28,29} In the present study, we sought to be the first to describe the outcomes of RF in comparison to PF.

When comparing preoperative and postoperative QOL outcomes, both PF and RF groups exhibited significant improvement in all measures, with the exception of the PDQ-Psychosocial measure in the RF group. RF patients had poorer preoperative PDQ-Functional scores as compared with PF patients. Postoperatively, there were no significant differences in any of the QOL measures between RF and PF, suggesting that RF achieves similar outcomes as PF regardless of preoperative status.

Within treatment groups, both the PF cohort and the RF cohort demonstrated statistically significant improvements in all QOL measures, with the exception of the PDQ-psyhosocial subscore for RF, which did not reach significance. Our data suggest that both PF and RF are effective in improving overall QOL, mental health, and disability in patients with radicular symptoms due to foraminal stenosis. However, further investigation is needed to compare the relative effectiveness of RF versus nonoperative treatment for patients who present with recurrent symptoms of LFS.

Although the aforementioned results seem to support the use of these procedures, our MCID analysis focuses on how meaningful these improvements actually are. Based on the previously established MCID values, 59\% of patients in the RF cohort achieved the MCID for EQ-5D. RF patients performed the worst when measured by the PHQ-9, with only 21\% of patients achieving the MCID. RF and PF were not associated with any significant intergroup differences in MCID. Thus, it can be concluded that although neither PF nor RF should be expected to consistently produce a clinically meaningful improvement, both procedures may still...
Be an alternative option for patients who have exhausted all other options.

Finally, no prior studies have investigated the relative costs of PF and RF in the LFS population. In our study, RF was more expensive than PF when cumulative costs are considered across the 3 episodes of care: 1-month preadmission, admission period, and postdischarge up to 1 year, as detailed in Table 3. RF was slightly less expensive in the preadmission period and slightly more expensive in the postdischarge period compared with PF; however, neither difference was significant. The admission period was very similar in cost, with no significant difference between groups. The interquartile ranges of the preadmission and postdischarge episodes are notably large, most likely due to large outliers in the sample. Our financial data serve to determine the cost of foraminotomy in a large sample size. These results suggest that RF and PF are similar in cost, further supporting the case for RF as a reasonable treatment for intractable, recurrent symptoms of LFS.

Despite this evidence, the advantages and disadvantages of RF must be carefully weighed when recommending patients for reoperation. Larger, prospective studies are needed to explore the cost-effectiveness of LFS treatment, as well as to elucidate what the optimal patient populations are for foraminotomy and repeat foraminotomy.

Limitations

Our study is not without its limitations. The retrospective design creates the potential for measurement bias due to the absence of a consistent data collection protocol among the treating clinicians. Furthermore, several different spine surgeons practicing at a single tertiary-care medical center participated in the care of this patient cohort. The resulting lack of standardization with regard to surgical decision-making may affect the generalizability of the results. Aside from this variation among clinicians who participated in this study, there may also have been variation among the patients themselves. LFS can be diagnosed and treated at many different stages in the disease course, introducing the potential for lead-time bias. The specific operative levels may have varied between cohorts, which may also affect standardization.

Additionally, PF was performed more frequently than RF in our cohort. Although the relative frequencies at which these procedures were performed in our study population appear to be consistent with the national trends, the disparity in sample size between PF and RF is not ideal from a statistical analysis standpoint. The length of follow-up is another possible limitation, as our study did not have the ability to capture any further changes in QOL that may have occurred after the last available follow-up visit.

Finally, because the MCID cutoffs used in this study were established in other patient populations, they may not be ideal for use in the LFS population. Thus, further studies should be performed to establish MCID cutoffs in the LFS population for each of the 3 QOL measures.

CONCLUSION

Our study found no significant difference with regard to the proportion of patients achieving MCID in the PF and RF groups, suggesting similar efficacy between PF and RF. Both PF and RF produced significantly improved scores in almost all measures of QOL, mental health, and disability. Postoperatively, there were no significant differences in any of the QOL measures between RF and PF, suggesting that RF achieves similar outcomes to PF regardless of preoperative status. However, we did not observe any significant difference in cost between PF and RF. Ultimately, our study indicates that RF can serve as a possible treatment option in patients with recurrent LFS and has the potential to produce clinically significant improvements in QOL. Given the current lack of literature on this topic, however, future prospective studies across multiple institutions are necessary to further elucidate the cost-effectiveness and efficacy of RF in LFS.

REFERENCES


Disclosures and COI: The authors received no funding for this study and report no conflicts of interest.

Corresponding Author: Emily Hu, MD, 3403 Ormond Rd, Cleveland Heights, OH 44118. Phone: (973) 980-8762; Email: emilyhu38@gmail.com.

Published 28 August 2020

This manuscript is generously published free of charge by ISASS, the International Society for the Advancement of Spine Surgery. Copyright © 2020 ISASS. To see more or order reprints or permissions, see http://ijssurgery.com.